### CONTENTS

What is the Carroll Round? vii  
Notes on Paper Submissions and Conference Participation vii  
Acknowledgements viii  
A Brief History of the Carroll Round x  
Introduction: Why I Support the Carroll Round xv  

#### Carroll Round Proceedings

<table>
<thead>
<tr>
<th>Title</th>
<th>Author(s)</th>
<th>Institution</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>An Analysis of Spatial-Based Signalling Games</td>
<td>Thomas BumBerger</td>
<td>Cambridge University</td>
<td>3</td>
</tr>
<tr>
<td>The Indian Inflation Puzzle: Is the Reserve Bank of India Using the Right Price Index for its Headline Inflation Measure?*</td>
<td>SAGIIAS SEN</td>
<td>London School of Economics</td>
<td>19</td>
</tr>
<tr>
<td>On the Other Side of the Tracks: Neighborhood Effects and Segregation in Morocco*</td>
<td>Russell Morton</td>
<td>Princeton University</td>
<td>20</td>
</tr>
<tr>
<td>Political Risk and Stock Market Returns: Evidence from International Equity Markets and Options Markets*</td>
<td>Jonathan Pedder</td>
<td>Dartmouth College</td>
<td>21</td>
</tr>
<tr>
<td>Inequality and Unrest: Could Local Inequality Increase the Marginal Propensity to Strike?</td>
<td>Alyssa HuBERTs</td>
<td>Georgetown University</td>
<td>22</td>
</tr>
<tr>
<td>Do U.S. Border Enforcement Operations Increase Human Smuggling Fees Along the U.S.-Mexico Border?*</td>
<td>Kyle CoOMBS</td>
<td>Macalester College</td>
<td>40</td>
</tr>
<tr>
<td>Federal Reserve Stimulus Effects on Mortgage and Treasury Rates</td>
<td>Benjamin Lee SHOESMITHSY</td>
<td>University of North Carolina Wilmington</td>
<td>41</td>
</tr>
<tr>
<td>Heterogeneity in Conditional Cash Transfer Effects: Evidence from a Randomized Trial in Malawi*</td>
<td>Johnny HUYNH</td>
<td>Pomona College</td>
<td>57</td>
</tr>
<tr>
<td>Motivation and Work Performance: The Effects of Prosocial and Intrinsic Motivation on Teacher Performance in Rural China*</td>
<td>Katherine LoOlSLEY</td>
<td>The London School of Economics</td>
<td>58</td>
</tr>
</tbody>
</table>

* denotes abstracts only

iv
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry and Finance in Interwar India</td>
<td>59</td>
</tr>
<tr>
<td>Russell Black</td>
<td></td>
</tr>
<tr>
<td>Oxford University</td>
<td></td>
</tr>
<tr>
<td>A Macro-financial Model for Credit Risk Stress testing:</td>
<td>79</td>
</tr>
<tr>
<td>The Case of Latvia</td>
<td></td>
</tr>
<tr>
<td>Victoria Pilinko</td>
<td></td>
</tr>
<tr>
<td>Andrei Romanescu</td>
<td></td>
</tr>
<tr>
<td>Stockholm School of Economics</td>
<td></td>
</tr>
<tr>
<td>Can I Stay a BIT Longer? The Pernicious Effect of Bilateral Investment Treaties on Political Survival*</td>
<td>102</td>
</tr>
<tr>
<td>SoumyaJit Mazumder</td>
<td></td>
</tr>
<tr>
<td>Georgetown University</td>
<td></td>
</tr>
<tr>
<td>The Self-Employed and Saving: New Evidence from Great Britain</td>
<td>103</td>
</tr>
<tr>
<td>Joshua Walker</td>
<td></td>
</tr>
<tr>
<td>Lancaster University</td>
<td></td>
</tr>
<tr>
<td>A Mechanism Design Approach to Increasing the Efficiency of Aid Distribution*</td>
<td>119</td>
</tr>
<tr>
<td>Matthew Klein</td>
<td></td>
</tr>
<tr>
<td>University of Chicago</td>
<td></td>
</tr>
<tr>
<td>The Challenge of National Treatment in International Trade: Evidence from WTO Dispute Settlement</td>
<td>120</td>
</tr>
<tr>
<td>Heather Hediger</td>
<td></td>
</tr>
<tr>
<td>Georgetown University</td>
<td></td>
</tr>
<tr>
<td>The Impact of Inequality and Ethnic Fragmentation on Educational Finance Evidence from the US, 1970-2000*</td>
<td>140</td>
</tr>
<tr>
<td>Jill Ni</td>
<td></td>
</tr>
<tr>
<td>Georgetown University</td>
<td></td>
</tr>
<tr>
<td>Dawn Chan</td>
<td></td>
</tr>
<tr>
<td>Georgetown University</td>
<td></td>
</tr>
<tr>
<td>Liberal Objectives: Does the Ideological Distance between the United States and Developing Countries Affect how much Aid they Receive?</td>
<td>158</td>
</tr>
<tr>
<td>Chris Stromeyer</td>
<td></td>
</tr>
<tr>
<td>Georgetown University</td>
<td></td>
</tr>
<tr>
<td>Investigating Risk Tolerance Stability using the NLSY79*</td>
<td>176</td>
</tr>
<tr>
<td>Robert Denz</td>
<td></td>
</tr>
<tr>
<td>University of Virginia</td>
<td></td>
</tr>
<tr>
<td>Liberal Objectives: Does the Ideological Distance between the United States and Developing Countries Affect how much Aid they Receive?</td>
<td>158</td>
</tr>
<tr>
<td>Investigating Risk Tolerance Stability using the NLSY79*</td>
<td>176</td>
</tr>
<tr>
<td>Robert Denz</td>
<td></td>
</tr>
<tr>
<td>University of Virginia</td>
<td></td>
</tr>
</tbody>
</table>

* denotes abstracts only
# Table of Contents

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment Benefit Increases and Mortality*</td>
<td>177</td>
</tr>
<tr>
<td>Eric S. Aldenhoft</td>
<td></td>
</tr>
<tr>
<td>University of Maryland</td>
<td></td>
</tr>
<tr>
<td>Subnational Politics and Foreign Direct Investment in India</td>
<td>178</td>
</tr>
<tr>
<td>Meredith Strike</td>
<td></td>
</tr>
<tr>
<td>Georgetown University</td>
<td></td>
</tr>
<tr>
<td>Market Reactions to the US Shale Gas Boom*</td>
<td>200</td>
</tr>
<tr>
<td>Natalya Langbord</td>
<td></td>
</tr>
<tr>
<td>Yale University</td>
<td></td>
</tr>
<tr>
<td>Democracy and the Effects of Negative Exogenous Income Shocks: Evidence from individual and household data in sub-Saharan Africa</td>
<td>201</td>
</tr>
<tr>
<td>Samun Knight</td>
<td></td>
</tr>
<tr>
<td>Oberlin College</td>
<td></td>
</tr>
<tr>
<td>Democracy on Demand? Local Government Representation and Access to BPL Cards in India</td>
<td>217</td>
</tr>
<tr>
<td>Brian Goggin</td>
<td></td>
</tr>
<tr>
<td>Georgetown University</td>
<td></td>
</tr>
<tr>
<td>An Econometric Approach to the Dollarization in Ukraine</td>
<td>230</td>
</tr>
<tr>
<td>Nikhil Kalathil</td>
<td></td>
</tr>
<tr>
<td>Oberlin College</td>
<td></td>
</tr>
<tr>
<td>Does Schooling Pay? Evidence from China</td>
<td>243</td>
</tr>
<tr>
<td>Vincent La</td>
<td></td>
</tr>
<tr>
<td>Dartmouth College</td>
<td></td>
</tr>
<tr>
<td>Driving Style as a Game: Behavioral Responses to Road Safety Factors</td>
<td>259</td>
</tr>
<tr>
<td>Rachel Seymansi</td>
<td></td>
</tr>
<tr>
<td>Georgetown University</td>
<td></td>
</tr>
<tr>
<td>Appendix A: Thirteenth Annual Carroll Round Presentation Schedule</td>
<td>273</td>
</tr>
<tr>
<td>Appendix B: Past Speakers</td>
<td>275</td>
</tr>
<tr>
<td>Appendix C: Former Carroll Round Steering Committees</td>
<td>276</td>
</tr>
<tr>
<td>Appendix D: Members of the Advisory Panel</td>
<td>278</td>
</tr>
<tr>
<td>Appendix E: Past Participants</td>
<td>279</td>
</tr>
</tbody>
</table>

* denotes abstracts only
What is the Carroll Round?

The Carroll Round is an international economics conference for undergraduate students held annually at Georgetown University in Washington, D.C. It takes the format of a professional academic conference at which students present their original research in international economics (broadly defined) that are typically honors theses. The goal of the Carroll Round is to foster the exchange of ideas among the leading undergraduate economics students by encouraging and supporting the pursuit of scholarly innovation. To date, over 250 students from universities and colleges in North America, Western and Eastern Europe, and Australia have participated, making the Carroll Round the premier conference of its kind. The conference also provides opportunities for participants to interact with prominent academic and policy economists. Alumni have moved on to top Ph.D., J.D., M.B.A., and other graduate programs, positions at the Federal Reserve, World Bank, and other public institutions, and major private corporations.

Notes on Paper Submissions and Conference Participation

The Carroll Round Proceedings is a publication of synopses and full-length papers from the Carroll Round Undergraduate International Economics Conference at Georgetown University. We do not accept paper submissions from the general public. If you are interested in presenting at the conference, please log on to our website: http://carrollround.georgetown.edu. All undergraduate students who have written or are in the process of writing original work in the field of international economics (broadly defined) are encouraged to apply.
**ACKNOWLEDGEMENTS**

Beginning in 2001, the Carroll Round has evolved from a dream—an ambitious, powerful dream—into the premier conference for undergraduate research in international economics. This dream, however, could neither have been realized nor continue to grow into the future without the supports of myriad donors, faculty, speakers, session chairs, and Georgetown University itself. For this reason, the Carroll Round Steering Committee would like to thank those who have dedicated much of their own time and energy to make this conference a remarkable achievement year after year.

An exciting announcement can be made here, that the Carroll Round Endowed Program Fund has now reached the minimum to allow it to be established as a formal permanent endowment the payouts from which will support part of the annual conference operation. This endowment was the idea of Ms. Marianne Keler who with her husband Mr. Michael Kershaw have been the primary contributors. They designed it to be a fund to which anyone can contribute, and over the past five years, several alumni of the Carroll Round have joined this effort. We intend to continue to grow this endowment. Ms. Keler has been the primary contributor and advocate of the Carroll Round since its establishment in 2001. We thank Mr. and Ms. Keler from the bottom of our hearts for looking after us from birth, and pledge that we will continue to pursue excellence in undergraduate learning and research at Georgetown so that we can in turn become contributors to our global society.

Our deepest gratitude will forever to go to Mr. Yunho Song, a Georgetown graduate and long-time supporter of the conference, who designated his endowment fund to partly finance the Carroll Round. The first committee approached Mr. Song in 2001, and since then, his support has continuously kept the Carroll Round in operation. For his support and extremely generous financial contributions, we and all future participants of the Carroll Round are truly indebted.

The Carroll Round has been gifted with many other individuals committed to its cause. We would like to recognize Mr. Mario Espinoso, Mr. Oleg Nodelman, Ms. Colleen Murphy, Ms. Sarah Osborne, Mr. and Mrs. Kenneth Kunkel, Mr. Jon Skillman and Ms. Lu- anne Selk, Mr. Geoffrey Yu, and former Carroll Round Steering Committee members Mr. James Arnold, Mr. Stephen Brinkmann, Ms. Amanda Delp, Dr. Andrew Hayashi, and Mr. Scott Pedowitz. The first five conferences were funded significantly by the Sallie Mae Corporation and we are most grateful for their foresight to support what has continued to grow. We express our gratitude to the Kanzanjian Foundation, which provided the startup funds without which it would have been impossible to develop the Carroll Round Proceedings.

Within Georgetown, the conference is indebted to Ms. Gail Griffith, Mr. Richard Jacobs, Mr. Mohamed Abdel-Kader and Ms. Carma Fauntleroy, whom have made tireless efforts to advocate the Carroll Round cause from our campus’ advancement department. We are also deeply appreciative of the efforts of Ms. Christine Smith, Mr. Thomas Esch, Ms. Cara Sodos, Ms. Reema Ghazi, Dr. Venilde Jeronimo, Ms. Katerina Kulagina, and Ms. Elizabeth Franzino from years past.

An exciting announcement can be made here, that the Carroll Round Endowed Program Fund has now reached the minimum to allow it to be established as a formal permanent endowment the payouts from which will support part of the annual conference operation. This endowment was the idea of Ms. Marianne Keler who with her husband Mr. Michael Kershaw have been the primary contributors. They designed it to be a fund to which anyone can contribute, and over the past five years, several alumni of the Carroll Round have joined this effort. We intend to continue to grow this endowment. Ms. Keler has been the primary contributor and advocate of the Carroll Round since its establishment in 2001. We thank Mr. and Ms. Keler from the bottom of our hearts for looking after us from birth, and pledge that we will continue to pursue excellence in undergraduate learning and research at Georgetown so that we can in turn become contributors to our global society.

Our deepest gratitude will forever to go to Mr. Yunho Song, a Georgetown graduate and long-time supporter of the conference, who designated his endowment fund to partly finance the Carroll Round. The first committee approached Mr. Song in 2001, and since then, his support has continuously kept the Carroll Round in operation. For his support and extremely generous financial contributions, we and all future participants of the Carroll Round are truly indebted.

The Carroll Round has been gifted with many other individuals committed to its cause. We would like to recognize Mr. Mario Espinoso, Mr. Oleg Nodelman, Ms. Colleen Murphy, Ms. Sarah Osborne, Mr. and Mrs. Kenneth Kunkel, Mr. Jon Skillman and Ms. Lu- anne Selk, Mr. Geoffrey Yu, and former Carroll Round Steering Committee members Mr. James Arnold, Mr. Stephen Brinkmann, Ms. Amanda Delp, Dr. Andrew Hayashi, and Mr. Scott Pedowitz. The first five conferences were funded significantly by the Sallie Mae Corporation and we are most grateful for their foresight to support what has continued to grow. We express our gratitude to the Kanzanjian Foundation, which provided the startup funds without which it would have been impossible to develop the Carroll Round Proceedings.

Within Georgetown, the conference is indebted to Ms. Gail Griffith, Mr. Richard Jacobs, Mr. Mohamed Abdel-Kader and Ms. Carma Fauntleroy, whom have made tireless efforts to advocate the Carroll Round cause from our campus’ advancement department. We are also deeply appreciative of the efforts of Ms. Christine Smith, Mr. Thomas Esch, Ms. Cara Sodos, Ms. Reema Ghazi, Dr. Venilde Jeronimo, Ms. Katerina Kulagina, and Ms. Elizabeth Franzino from years past.
We would also like to give special recognition to former steering committee members, beyond those already mentioned, who have contributed very generous portions of their post-collegiate income to the Carroll Round after graduation. Among them, we would especially like to thank Ms. Meredith Ballotta, Ms. Stacey Droms, Mr. Brandon Feldman, Ms. Daphney Francois, Ms. Yasmine Fulena, Mr. Edward Hedke, Mr. Christopher Griffin, Ms. Rebecca Heide, Mr. Dennis Huggins, Ms. Cindy Jin, Mr. Michael Karno, Ms. Anna Kis, Mr. Dan Leonard, Mr. J. Brendan Mullen, and Mr. Shuo Tan. Beyond the financial viability of the Carroll Round, the conference also enjoys the grace of many proponents on Georgetown University’s campus to ensure its continuing and vibrant existence. In particular we would like to remember the support of the late Dean Carol Lancaster of the School of Foreign Service, as well as thank her successor Dean James Reardon-Anderson.

We thank the help of SFS Dean’s Office members Dean Kendra Billingslea, Mrs. Denisse Bonilla-Chaoui, Dr. Dan Powers, Mr. Beau Boughamer, Mr. Benjamin Zimmerman, Mr. Franz Hartl, and Ms. Rebecca Ernest. We would like to recognize former Dean Robert Galucci, who supported the development of the Carroll Round from its inception.

The Carroll Round has been fortunate for the last fourteen years to enjoy the substantive quality of the brightest economics undergraduates from across the world. We are particularly grateful to those students whose names are all listed in the Appendix section. We would also like to recognize the session chairs who take the time to read participants’ papers and critique their presentations at the conference. We would like to thank the 2014 session chairs for their contributions to the conference: Professors Robert Cumby, Raj Desai, Christopher Griffin, Rodney Lude- ma, Carlo Prato, Steven Radelet, Charles Udomsaph, and James Vreeland.

We thank the past Carroll Round Steering Committees, which have shaped and directed the development of the conference into its current state today. Their names are all listed in the Appendix section. We are also indebted to the contributions of the Carroll Round Advisory Panel for their assistance in developing a long-term vision for the Carroll Round and for grounding where the next decade may take this institution.

Finally, though not least importantly, we would like to express our ever-growing gratitude to Dean Mitch Kaneda, the Carroll Round Faculty Advisor. Without his support, time, and passion, this endeavor would not be possible.
A BRIEF HISTORY OF THE CARROLL ROUND

(Revised April 2015)

Each year when April is on the horizon, I realize how the Carroll Round is at once completely recognizable as the successor to the first conference weekend and unlike anything my friends on the inaugural committee imagined. Accepted paper quality has increased exponentially, and the weekend’s highlights are the students’ masterful presentations as much as the keynote speeches. None of these advances would be possible without the extraordinary work of the Georgetown students who organize the Round and of course the global contingent that descends on the nation’s capital each year. Other alumni and I remain awestruck by the effort, dedication, and commitment of each successive participant group. Despite the need always to look ahead, reviewing one’s origins is equally important. During the first year, the ingenuity and dedication of a stellar group of Georgetown students, combined with the contributions of remarkable young scholars from around the country, showed how strong undergraduate economics—and the work of undergraduate economists—can be.

The conference’s birthplace, as many know by now, was an Oxford pub called the Radcliffe Arms. Even though that tale is completely true, the Carroll Round’s roots extend firmly to the Georgetown University campus. For it was there that an incredible team of friends and colleagues assembled and launched the event in 2001.

Throughout the 1999-2000 academic year, I had the great pleasure of meeting and learning alongside seven outstanding economics classmates. My first meaningful discussions about economics took place that year with fellow students Andrew Hayashi and Ryan Michaels. Andrew and I were both enrolled in Professor Mitch Kaneda’s International Trade class that semester, and Ryan suffered with me through Microeconomic Theory as well as the demanding Introduction to Political Economy. I remember feeling intimidated at first by their boundless knowledge of theory and their irrepressible enthusiasm for learning. Over time I realized the extent to which I was learning from them as much as our instructors; their insights often proved more valuable than the content of weekly lectures. I also became acquainted with a second group of classmates, including Bill Brady, Josh Harris, Kathryn Magee, Brendan Mullen, and Scott Pedowitz. By the spring, our paths all pointed to Europe: Bill, Kathryn, and Scott were on their way to the London School of Economics; Brendan had chosen the University of Bristol; and Josh was destined for Poland and Hungary. Andrew, Ryan, and I planned to spend our year abroad at the University of Oxford studying a mixture of philosophy, politics, and economics. Before departing in October 2000, I knew our shared plans were not the product of mere coincidence. Something special would emerge from the experience.

Having established initial ties at Georgetown, the three of us began meeting on a regular basis to discuss our latest tutorial sessions, grueling problem sets, the future of macroeconomics and, occasionally, the latest gossip about luminaries in the field. Whereas C.S. Lewis, J.R.R. Tolkien, and the other Inklings made The Eagle and Child pub their
intellectual home away from home, we adopted the Radcliffe Arms as our haven. Over pints and pub food, Andrew’s twin passions for game theory and philosophy emerged. The future of monetary policy and development began to vex Ryan’s thoughts, while I hoped to better understand the mechanisms of cooperation, and conflict, underlying international trade institutions.

Meanwhile at Pembroke College, I encountered a group of students from universities across the country also spending their junior years at Oxford. I naturally befriended the other economists in our contingent, but I also developed close relationships with physicists, biologists, literary scholars, and art historians. In the Junior Common Room, a student lounge of sorts for undergraduates, or over traditional English dinners in the dining hall, we shared stories about life at our respective universities and the latest research we were conducting at Oxford. As thesis and postgraduate plans matured during these conversations, I appreciated ever more my exposure to alternative experiences and approaches to scholarship. The year eventually came to an end, and I worried that these exciting connections would dissolve upon return to the United States.

One evening at the start of my final term in Oxford, I thought about the importance of this dialogue and my commitment to the study of international economics. I had a distressing feeling that undergraduates, especially in economics, were not afforded adequate opportunities to present their work in a serious setting. After all, I always felt privileged when Andrew, Ryan, and my fellow Pembrokians shared their original ideas with me. I thought that undergraduate economists from around the country deserved an event at which they could interact significantly with each other and the professional academic community. In March 2001, I composed a memo that outlined my solution: the Carroll Round. The following paragraph from that proposal captures my motivating thoughts:

As they prepare for careers in academia, public service, and business, undergraduate students throughout the country also have joined a momentous dialogue in collegiate, national, and global fora. Many are involved in independent research representing the next generation of critical thought in international relations. Others have enjoyed unique experiences through jobs and internship programs that expose them to the front lines of economic policy-making and statecraft. Young women and men also have championed vociferously environmental and labor-related causes through awareness and service programs. Clearly, these timely economic issues are assuming greater importance for the future of international relations and are reflected in the abundance of attendant student research, interest, and initiative. Therefore, I propose to coordinate and host, in association with Georgetown University’s School of Foreign Service and John Carroll Scholars Program, the next ‘round’ of economic and political discussion and debate—the Carroll Round.

I invited Andrew and Ryan to join me in this endeavor over pints at the Radcliffe Arms even though there was no guarantee they would think it a good idea. I was confident that if such rising stars believed in the concept, other students would join in time. Having worked out more substantive ideas over the summer, I finally was prepared to call upon the other economics celebrities in my class to collaborate on the project. Bill, Josh, Kathryn, Brendan, and Scott fortunately signed on and completed the senior circle. A few months
later we welcomed four more students: Cullen Drescher, Mark Longstreth, Waheed Sheikh, and future Chair Meredith Gilbert to encourage younger students and ensure continuity for the future.

With the unflagging assistance of then-John Carroll Scholars Program Director John Glavin, the proposal was circulated among university administrators. After gaining their initial support, I asked Mitch Kaneda, my most influential undergraduate teacher and a newly appointed Associate Dean of the School of Foreign Service, to review the proposal. Without hesitation—and somewhat to my surprise—he offered his assistance, embarking on an indefinite and irreplaceable stewardship of the Carroll Round. Former Dean Robert Gallucci and his staff also extended moral and financial support, which cemented our institutional place at Georgetown.

The first Carroll Round Steering Committee struggled through many difficult decisions regarding conference content, format, and funding. Should submitted papers be limited to topics in international economics? What elements must be included in submissions and presentations? How do we ensure that financial constraints do not prevent the best students from attending? Over marathon sessions in Healy Hall and at the Tombs, we developed a model for the Carroll Round that has largely remained intact. Development Officers shared our ideas with generous alumni who responded favorably and pledged individual donations. Little by little, our initial concepts materialized into reality. When School of Foreign Service alumna Marianne Keler ’76 convinced the Sallie Mae Fund to contribute $10,000 to the Carroll Round, we both gained a lead sponsor and secured the long-term future of the conference. Since that year, Marianne has been gracious in her support and instrumental in expanding our reach to new global partners, including the American University in Bulgaria.

After distributing colorful brochures, contacting the top departments in the country, and preparing the Hilltop for the event, applications streamed in during the spring. By late March, we had narrowed our list of invited students to thirty-two. Seniors traveled to Washington from as near as the University of Virginia and as far as Stanford University. The Committee was stunned by the participants’ and their home departments’ enthusiasm. Among the more notable responses, Illinois-Wesleyan University sent four young economists to the conference and soon after published a special Carroll Round edition of their undergraduate economics journal.

The first Carroll Round officially began on Friday April 5, 2002, and the proceedings came to a close two days later. Participants enjoyed an exclusive audience with Director of the National Economic Council Lawrence B. Lindsey in the beautiful Riggs Library before hurrying to the Federal Reserve for another private meeting with former Vice Chairmen Roger W. Ferguson and Donald L. Kohn. The two monetary policy experts shared candid stories about the effects of September 11, 2001 on the nation’s banking system and the various roles that the Federal Reserve plays in American economic activity.

Dr. John Williamson of the Institute for International Economics spoke about development...
issues over a splendid dinner at Cafe Milano, and Dr. Edwin M. Truman, former Assistant Secretary of the U.S. Treasury for International Affairs, closed the conference with words of wisdom to students considering careers in academia and policymaking.

A total of twenty-eight papers were presented over the weekend, showcasing the impressive work of men and women now at the forefront of academia, law, and business. Georgetown professors who served as panel discussants later remarked that the quality of some presentations met or surpassed the sophistication of recent graduate-level dissertations. Judging by their comments, the conference brought together some of the best young prospects in economics as they approached the frontiers of research.

I never imagined in March 2001 that the first Carroll Round would attain the heights realized one year later, or for that matter even exist. The event has grown since then in size and scope beyond my initial hopes. The participation of Nobel Laureates from John F. Nash, Jr., in 2004 to George Akerlof in 2015, as well as Susan Athey, the first female recipient of the John Bates Clark Medal, in 2008 mark special peaks in the evolution of the conference. Indeed, this historic slate of speakers could not be more finely tuned to the spirit of the Carroll Round. The groundbreaking work that each has contributed to the study of international economics, including numerous articles and books designed to influence lay readers and public policy decision-makers, serve as exemplars for other scholars and practitioners.

Looking to the Carroll Round’s future, I still hope that students from the developing world eventually will be able to attend. Regardless of their home institutions, I continue to enjoy meeting participants and learning about their research interests. As they share in the excitement of presenting their work and the occasional trepidation of fielding questions, I feel humbled to be among such gifted individuals. In fact, alumni from previous years have advanced to graduate study at Berkeley, Chicago, Cornell, Duke, MIT, Michigan, Minnesota, Northwestern, Oxford, Princeton, Yale, and Wisconsin as well as top government and finance positions around the country. Past participants now are tenure-track members of economics faculties including: Santosh Anagol at the University of Pennsylvania Wharton School, Andrew Hayashi at the University of Virginia School of Law, Ryan Michaels at the University of Rochester, Raphael Schoenle at the Brandeis University Economics Department and International Business School, and Tom Vogl at Princeton University. The cadre of former conference participants truly has grown into a professional and academic network unlike any other for young economists.

As always, I thank the Kazanjian Foundation for their generous support, which makes annual publication of the Carroll Round Proceedings possible. I also would like to extend my unwavering gratitude to the members of the inaugural Carroll Round Steering Committee without whom this history would have remained fiction. I have great respect and admiration for successive Chairs from Seth Kundrot in 2003 to Kristen Skillman in 2015. Those leaders and all in between ensure the success of the Carroll Round each year and deserve our appreciation.
The Carroll Round received a donation not long ago, much like the original Sallie Mae Fund contribution, which created an endowment for the conference thanks to the largesse of School of Foreign Service alumnus Yunho Song ’86. I distinctly remember meeting with him and some of my closest friends at the Tombs to discuss our fledgling project, uncertain that fall semester in 2001 whether it would ever see the light of day. He was instrumental then in making the Carroll Round a reality, and he now has solidified its place within the fabric of Georgetown and the School of Foreign Service. For that, all of us who have watched the conference grow extend our heartfelt gratitude. The spirit of his gift, though, should live on through us. Support from alumni, not just of the financial variety, maintains the conference’s vibrancy long after the proceedings conclude. I encourage each of you to return to Georgetown in April and to consider making any donations to the Carroll Round fund when possible.

Finally, and as always, I must thank Mitch Kaneda who has miraculously preserved my vision for the Carroll Round over the years and watched over past Committees as they built upon its initial success and joined the ranks of distinguished alumni. With his continued collaboration and the eagerness of future Georgetown students, the Carroll Round’s future will dwarf the accomplishments of its past, creating even more exciting opportunities for undergraduate economists to learn from the best in the field and, more importantly, from each other.

Christopher L. Griffin, Jr.
Georgetown Class of 2002
Carroll Round Founder
INTRODUCTION

WHY I SUPPORT THE CARROLL ROUND

Ever since coming to Georgetown University, I have been a strong proponent of advancing the work of undergraduate research at this university, especially in the use of quantitative methods. The undergraduates that I have taught at Georgetown and at Yale are among the brightest, most analytical individuals with which I have had the pleasure to work, and I believe that the research they have the potential to perform can rival some of the work coming out of graduate programs today. In past years of this conference, for example, I have repeatedly marveled at the rigorous analytical work, the creativity in flexibly adapting existing tools to answer new problems, and the ability to synthesize results into a compelling conclusion that nearly every presented paper demonstrates.

In recent years, an increasing number of universities and undergraduate programs have come to recognize what I long ago intuited, and now are offering an increasing number of resources to and outlets for undergraduate research in International Economics and International Political Economy. More and more colleges, for example, are requiring students who major in these fields to produce a senior thesis, and more and more colleges are publishing journals of undergraduate economic research or else hosting conferences for journal-quality material.

These programs, like me, have recognized that a quality undergraduate education should not only serve to teach people a laundry list of facts, or even transfer a discreet skill set that can be sold on the job market. Rather, we have come to recognize that a college education must also provide students with an inquisitive mind, one that can ask the kinds of questions that lead to the discovery of new, transformative knowledge. In short, colleges should teach their students how to be researchers.

The Carroll Round, then, struck me when I first arrived at Georgetown in 2009—and continues to strike me today—as an inspired conference truly ahead of its time. Started 15 years ago, this conference has already brought together over 600 participants and dozens of acclaimed economics scholars and practitioners, fostering the exchange of ideas among the field’s current and future leaders. These discussions can be truly extraordinary—I have been continually amazed by the depth of understanding that presenters offer regarding both their topics as well as of quantitative methodology, where they consistently use cutting edge research methods to analyze questions both old and new.

Moreover, I am always impressed by the quality of the peer critiques offered as fellow conference participants, all undergraduates, offer constructive, insightful suggestions on each other’s work. Truly, this conference provides students an opportunity to experience what it will mean for them to someday be thought leaders in this field.

Most importantly, however, I support the Carroll Round because of the opportunity that it provides me as a professor to reflect on the current state of undergraduate education. What I see at the Carroll Round consistently encourages me to continue to help students
to become not only consumers but also producers of knowledge in the fields of Economics and Political Economy. Best of all, the Carroll Round allows me to glimpse the future of the fields of International Economics and Political Economy. And the future is bright indeed.

James Vreeland
Professor of International Relations
School of Foreign Service
Georgetown University
1 Introduction

Increasing globalization changed the importance of advertisements. Consumers might know of goods and services offered within their region, but they might not be aware of products offered further away. Advertisements can reduce this information asymmetry. But signaling literature also suggests a more subtle value to advertising. Suppose the good is an experience good, i.e. a good whose quality cannot be verified at purchase. Then ads can help companies to truthfully reveal whether they have a high quality product. Ads act as a signal (Nelson 1970, 1974, 1978). While consumers might know about product quality in their region, consumers likely experience higher quality uncertainty for goods bought from further away. As such, the spatial dimension may be an important part of a firm’s signaling decision.

This paper expands on existing literature by introducing a spatial dimension to a signaling game. This paper first introduces an agent-based computational model in order to analyze signaling in a spatial context. This paper then introduces a broadcasting radius to capture a signal’s reach. This paper also analyzes if the value of signaling changes as the system becomes...

AN ANALYSIS OF SPATIAL-BASED SIGNALING GAMES
THOMAS BUMBERGER
Cambridge University

ABSTRACT
This dissertation uses an agent-based computational approach to analyze the effectiveness of spatial signals in revealing private information. Receivers only observe a signal if they are within a sender’s broadcasting radius. Signaling costs do not depend on private information, suggesting an interpretation of signals as advertisements. I find that there is little evidence of signaling having any value if receivers are isolated. However, once the system becomes more interconnected, there is a significant benefit if senders engage in signaling. They can successfully reveal their private information. The spatial structure, therefore, is important when assessing the effectiveness of signals.

1 Introduction

Increasing globalization changed the importance of advertisements. Consumers might know of goods and services offered within their region, but they might not be aware of products offered further away. Advertisements can reduce this information asymmetry. But signaling literature also suggests a more subtle value to advertising. Suppose the good is an experience good, i.e. a good whose quality cannot be verified at purchase. Then ads can help companies to truthfully reveal whether they have a high quality product. Ads act as a signal (Nelson 1970, 1974, 1978). While consumers might know about product quality in their region, consumers likely experience higher quality uncertainty for goods bought from further away. As such, the spatial dimension may be an important part of a firm’s signaling decision.

This paper expands on existing literature by introducing a spatial dimension to a signaling game. This paper first introduces an agent-based computational model in order to analyze signaling in a spatial context. This paper then introduces a broadcasting radius to capture a signal’s reach. This paper also analyzes if the value of signaling changes as the system becomes...
more interconnected. Value in this case is defined as whether agents are successful in revealing their private information.

Using simulations, I find little evidence that there is any value to signaling if consumers are isolated. However, if consumers are interconnected there is a significant benefit if firms engage in signaling. Firms can successfully reveal their private information resulting in consumers buying from high quality firms. These results suggest that, when making signaling decisions, analyzing the underlying spatial structure is important.

I proceed as follows. Section 2 places the model in the wider research context and reviews relevant literature. Section 3 focuses on the economic context of the game. The formal game is introduced in section 4. Section 5 motivates the agent-based computational approach and section 6 discusses the results. Finally, section 7 summarizes the main conclusions.

2 Literature

This dissertation relates to both signaling and spatial agent-based computational literature. There is currently little overlap between these two areas. Economic signaling games have not yet been explored using spatial agent-based computational methods. The analysis in section 6 makes a case for filling this gap.

2.1 Signaling Literature

Spence (1973) is the first to introduce the concept of signaling to explain why education can be a signal of innate ability. Unlike the model analyzed in this dissertation, signaling costs depend on private information. It is, however, not obvious, why advertising costs should be higher for low quality firms than for high quality firms.

Therefore, I follow a type of signaling game first described by Nelson. He analyzes how advertisements are related to the unobserved quality of experience goods and hence constitute signals. Signaling costs do not directly depend on unobserved quality. Indirectly, however, it is more costly for low quality firms to signal. Only firms with high quality products can expect repeat purchases and therefore it is easier to recoup signaling costs.

Milgrom (1986) formalize Nelson’s arguments in a model in which both the sales price and advertisement expenditure are signals. While I maintain the assumption that the cost of advertising is homogenous across types, I do not consider the signaling impact of pricing decisions. I focus on extending existing analysis by introducing a spatial dimension.

Empirically Brandts (1992) finds that in experiments using standard signaling games, players consistently learn to play the equilibrium implied by Cho’s (1987) intuitive criterion. Cooper (1997) find that in a repeated signaling game, experiments converge towards a unique equilibrium. An adaptive learning model best characterizes the cost of advertising in homogenous across types. I do not consider the signaling impact of pricing decisions. I focus on extending existing analysis by introducing a spatial dimension.

Milgrom (1986) formalize Nelson’s arguments in a model in which both the sales price and advertisement expenditure are signals. While I maintain the assumption that the cost of advertising is homogenous across types, I do not consider the signaling impact of pricing decisions. I focus on extending existing analysis by introducing a spatial dimension.

Empirically Brandts (1992) finds that in experiments using standard signaling games, players consistently learn to play the equilibrium implied by Cho’s (1987) intuitive criterion. Cooper (1997) find that in a repeated signaling game, experiments converge towards a unique equilibrium. An adaptive learning model best characterizes the con-
vergence phase and equilibrium selection depends on the observed history of the game. Real-life agents therefore seem to exhibit myopic rather than forward-looking responses. This suggests approaching signaling from an agent-based computational point of view. The computational implementation of myopic responses is simple and often even required due to performance limitations. The presented agents' behavioral rules are in line with an adaptive learning model and outcomes depend on the observed history.

Tropeano (2001) analyzes a simple signaling game from a spatial perspective. He finds that if firms can choose their location as a signal, informational asymmetries will lead to spatial agglomeration. This is, however, fundamentally different from the model analyzed in this dissertation. Here, a firm’s location is fixed. The firm can, however, choose the spatial reach of a signal. The spatial structure therefore relates to the reach of the signal as opposed to the location being the signal itself.

Note that the model presented in this dissertation also extends Zahavi (1975) and Grafen (1990)’s work on signaling in the context of theoretical evolutionary biology. While the narrative of this dissertation focuses on advertisements, the dynamics of the game could also describe mating rituals. For example, an interpretation of the signal could be a feature of an animal’s courtship behavior. A bird’s singing could be interpreted as a signal. Singing louder increases the signal’s geographic reach, while singing in a particular way affects the intensity of the signal.

2.2 Agent-Based Computational Literature

Schelling (1969, 1971) uses an agent-based computational approach to analyze a spatial game (without a computer). By using simple dynamic rules to model agents’ housing decisions, a spatial segregation of individual types emerges. This arises, despite none of the individual agents actually preferring to live in a segregated neighborhood. Given the non-linearity and heterogeneity introduced by a spatial arrangement, it is natural to follow an agent-based computational approach, as closed-form solutions are often not available. I follow Schelling’s basic approach of repeatedly applying simple dynamic rules on a spatial structure. However, I use a more complex spatial structure and different behavioral rules to capture the features of a signaling game.

Thus far, few agent-based models relate to product quality uncertainty. Jonker (2006) analyze trust in a trade network using agent-based simulations. The underlying game they analyze is, however, a trust and tracing game. Agents decide on whether to get the quality of a product verified by a trusted third party, whereas in a signaling game, product quality can be inferred from observable characteristics such as advertisements. Blasco (2010) analyze a simulation of agents who choose between a known technology and a new technology with an uncertain outcome. They study the flow of information through a network described by a Moore and a Kleinberg neighborhood. They find that a Kleinberg neighborhood, which results in a more interconnected system, leads to more people adopting the new technology compared to a Moore neighborhood. To analyze how the value of a spatial signal changes as the system becomes more interconnected, I use a similar spatial setup. Instead of a Moore neighborhood, convergence phase and equilibrium selection depends on the observed history of the game. Real-life agents therefore seem to exhibit myopic rather than forward-looking responses. This suggests approaching signaling from an agent-based computational point of view. The computational implementation of myopic responses is simple and often even required due to performance limitations. The presented agents’ behavioral rules are in line with an adaptive learning model and outcomes depend on the observed history.

Tropeano (2001) analyzes a simple signaling game from a spatial perspective. He finds that if firms can choose their location as a signal, informational asymmetries will lead to spatial agglomeration. This is, however, fundamentally different from the model analyzed in this dissertation. Here, a firm’s location is fixed. The firm can, however, choose the spatial reach of a signal. The spatial structure therefore relates to the reach of the signal as opposed to the location being the signal itself.

Note that the model presented in this dissertation also extends Zahavi (1975) and Grafen (1990)’s work on signaling in the context of theoretical evolutionary biology. While the narrative of this dissertation focuses on advertisements, the dynamics of the game could also describe mating rituals. For example, an interpretation of the signal could be a feature of an animal’s courtship behavior. A bird’s singing could be interpreted as a signal. Singing louder increases the signal’s geographic reach, while singing in a particular way affects the intensity of the signal.

2.2 Agent-Based Computational Literature

Schelling (1969, 1971) uses an agent-based computational approach to analyze a spatial game (without a computer). By using simple dynamic rules to model agents’ housing decisions, a spatial segregation of individual types emerges. This arises, despite none of the individual agents actually preferring to live in a segregated neighborhood. Given the non-linearity and heterogeneity introduced by a spatial arrangement, it is natural to follow an agent-based computational approach, as closed-form solutions are often not available. I follow Schelling’s basic approach of repeatedly applying simple dynamic rules on a spatial structure. However, I use a more complex spatial structure and different behavioral rules to capture the features of a signaling game.

Thus far, few agent-based models relate to product quality uncertainty. Jonker (2006) analyze trust in a trade network using agent-based simulations. The underlying game they analyze is, however, a trust and tracing game. Agents decide on whether to get the quality of a product verified by a trusted third party, whereas in a signaling game, product quality can be inferred from observable characteristics such as advertisements. Blasco (2010) analyze a simulation of agents who choose between a known technology and a new technology with an uncertain outcome. They study the flow of information through a network described by a Moore and a Kleinberg neighborhood. They find that a Kleinberg neighborhood, which results in a more interconnected system, leads to more people adopting the new technology compared to a Moore neighborhood. To analyze how the value of a spatial signal changes as the system becomes more interconnected, I use a similar spatial setup. Instead of a Moore neighborhood,
however, I use a von Neumann neighborhood. It is, unlike the Moore neighborhood, a proper subset of the Kleinberg neighborhood (cf. section 4) and as such, the more appropriate structure for simulating a counterfactual for increased interconnectedness. As such, the game outlined in section 4 constitutes an original model.

3 Economic Motivation

The following model is a game between few firms (senders) and many consumers (receivers). Firms produce a good with an inherent quality which does not change over time. Consumers try to buy the highest quality good, but can only observe the quality of the product ex-post, i.e. after they buy the good. The observed quality is, however, only correlated with the true quality. The good therefore is an experience good. This is a realistic assumption for a wide range of goods and services such as consumer electronics, health care, accounting services, et cetera.

Consumers learn over time by consuming the good and observing the outcomes. They change which firm to buy from accordingly. If, over a predefined period of time, a sender accrues higher signaling costs than it gains from its buyers, it goes out of business. This naturally limits the amount of signaling a firm can engage in. Following Cooper (1997)’s empirical observations, agents are not forward looking. Therefore there is no discounting of future payoffs and behavioral rules are only based on past not future information.

signaling in this context is a firm’s spending on advertisements. Ad campaigns usually do not directly depend on inherent product quality and as such the proposed cost function does not depend on it. In a stage game signaling is not more costly for low quality types. In a repeated game, however, a high quality type is more likely to recoup the advertisement spending through repeat purchases (cf. Milgrom (1986)). This implies that spending on advertisements is indirectly less costly for higher quality types.

I introduce spatial features to capture two real world phenomena. Firstly, signals such as advertisements have a geographical reach and it is costly to extend this reach. Billboards are only visibly within their immediate neighborhood. To reach more consumers, firms have to pay for billboards in many places. I therefore introduce the concept of a broadcasting radius, which limits the geographical reach of a signal. Only consumers within a firm’s broadcasting radius will observe the signal.

Secondly, consumers know more about products which are widely used within their social environment. The model uses the concept of neighborhoods which share outcome information. A consumer’s neighborhood consists of his directed links on the network of consumers. He can use outcome information from any of his immediate peers. This allows for social learning. An example of this would be consumers receiving recommendations for high quality products from friends, or reading online reviews of a product. These neighborhoods create network effects which diminish the value of a larger broadcasting radius. High quality firms benefit the most from these network effects. Not
only can they expect repeat purchases from existing customers, but they can now also expect purchases as a result of referral effects.

The reason this model limits a sender’s strategy space to choosing radius and signal – and not radius, signal and price – is for ease of exposition. Further work should extend the model by including pricing decisions as well because price can also act as a signal (cf. Milgrom (1986)). Two (strong) assumptions can justify this restriction. Consumers cannot infer quality from price and they do not face any budgetary constraints in their purchasing decisions. The model also abstracts from the cost of producing the good. This is valid if high quality firms have better technology which allows them to produce better products at the same cost as low quality firms.

4 Game

Consider the following repeated game:

Let $S = \{s_1, s_2\}$ be the set of senders competing for the attention of the set of receivers $R = \{r_1, r_2\}$. To model the spatial dimension, receivers are placed on a $\sqrt{N_R} \times \sqrt{N_T}$ two-dimensional lattice $T$ with connected borders. This means that $(0,0)$ is adjacent to $(1,0)$, $(0,1)$, $(\sqrt{N_R},0)$ and $(0,\sqrt{N_T})$. Geometrically this results in a torus and simplifies the analysis, as border cases do not have to be considered. Each receiver $r_i \in R$ occupies one space.

The $N_R$ senders are placed randomly on the lattice. Each $s_i \in S$ has an inherent quality $q^s_i \in Q = [0,1]$ which is drawn from a continuous uniform distribution. This inherent quality constitutes a sender’s private information in the game, i.e. an internal state. In order to reveal the private information, senders can decide to send a costly signal with intensity $n_{s_{ij}} \in M = [0,1]$ to all receivers within a broadcasting radius, $d^s_i \in N$. For this purpose, the definition of distance is the Manhattan distance, viz. the rectilinear distance, on the two-dimensional torus. In other words, it is the distance that results if one can only move up, down, left or right on the lattice. It is therefore the sum of the absolute difference of the x coordinates and the absolute difference of the y coordinates. It follows by inspection that the Manhattan distance $d(\cdot)$ between two points $(x,y)$ and $(x',y')$ – denoted in Cartesian coordinates – on $T$ is:

$$d(\cdot) = \min(|x-x'|, \sqrt{N_R-|x-x'|}) + \min(|y-y'|, \sqrt{N_T-|y-y'|})$$

By using Manhattan instead of Euclidean distances, distances represent the number of hops needed to get from one agent to another agent using the network implied by the spatial structure.

receivers observe a signal if they are within the chosen broadcasting radius of a sender. After observing the signal, they choose a sender to engage with. Receiver $r_i$ observes the quality of their chosen sender $s_j$ ex post as a binary random outcome variable $q^j_{ij}$ where $Pr(q^j_{ij} = 1|q^s_i) = q^s_i$ and $Pr(q^j_{ij} = 0|q^s_i) = 1 - q^s_i$. $q^s_i = 1$ is a good outcome, whereas $q^s_i = 0$ is a bad outcome. It follows, therefore, that the higher a sender’s inherent quality, the higher the probability of a good outcome for the receiver. Only can they expect repeat purchases from existing customers, but they can now also expect purchases as a result of referral effects.

- and not radius, signal and price – is for ease of exposition. Further work should extend the model by including pricing decisions as well because price can also act as a signal (cf. Milgrom (1986)). Two (strong) assumptions can justify this restriction. Consumers cannot infer quality from price and they do not face any budgetary constraints in their purchasing decisions. The model also abstracts from the cost of producing the good. This is valid if high quality firms have better technology which allows them to produce better products at the same cost as low quality firms.
upon engagement. After each receiver has observed their outcome, information about the choice of sender and the observed outcome becomes available in the local neighborhood of a receiver. Note that a receiver’s neighborhood does not change over time, but is determined at the start of the game. To model how different local neighborhoods affect the efficiency of signals, I consider the following concepts of neighborhoods. Figure 1 provides a visual representation of the neighborhood structures.

No neighborhood: In this simple case each receiver has only their own past outcome information when deciding which sender to engage with. Each node of the network is only linked with itself. Simplistically this would constitute a hermit. Each receiver is isolated.

Von Neumann neighborhood: Expanding a receiver’s neighborhood locally, the von Neumann neighborhood includes all receivers within a Manhattan distance of radius $v$. For $v = 1$ this means that a receiver’s neighborhood includes the receivers directly to the north, the south, the east and the west of the receiver’s position. Geometrically the neighborhood expands in the shape of a diamond. Each node of the network has a directed link to the 4 nodes which have a Manhattan distance of 1 or lower. This could be interpreted as a neighborhood consisting of friends living nearby.

Kleinberg neighborhood: Following Milgram1967’s work on real world networks and the phenomenon of small world networks, the Kleinberg neighborhood (??) generalises small world properties of networks. In a first step the Kleinberg neighborhood includes a von Neumann neighborhood of radius $v$. Additionally long-range links are formed according to a probability distribution. Consider receiver $r_i$. Then the probability that $r_1$ has a directed link with $r_i$ is

$$Pr(r_1 \rightarrow r_i) = \frac{d(r_1, r_i)^{-\gamma}}{\sum_{r_j \in R} [d(r_1, r_j)^{-\gamma}]}$$

where $d(.)$ represents the Manhattan distance between the two receivers. If $\gamma = 0$ upon engagement. After each receiver has observed their outcome, information about the choice of sender and the observed outcome becomes available in the local neighborhood of a receiver. Note that a receiver’s neighborhood does not change over time, but is determined at the start of the game. To model how different local neighborhoods affect the efficiency of signals, I consider the following concepts of neighborhoods. Figure 1 provides a visual representation of the neighborhood structures.

No neighborhood: In this simple case each receiver has only their own past outcome information when deciding which sender to engage with. Each node of the network is only linked with itself. Simplistically this would constitute a hermit. Each receiver is isolated.

Von Neumann neighborhood: Expanding a receiver’s neighborhood locally, the von Neumann neighborhood includes all receivers within a Manhattan distance of radius $v$. For $v = 1$ this means that a receiver’s neighborhood includes the receivers directly to the north, the south, the east and the west of the receiver’s position. Geometrically the neighborhood expands in the shape of a diamond. Each node of the network has a directed link to the 4 nodes which have a Manhattan distance of 1 or lower. This could be interpreted as a neighborhood consisting of friends living nearby.

Kleinberg neighborhood: Following Milgram1967’s work on real world networks and the phenomenon of small world networks, the Kleinberg neighborhood (??) generalises small world properties of networks. In a first step the Kleinberg neighborhood includes a von Neumann neighborhood of radius $v$. Additionally long-range links are formed according to a probability distribution. Consider receiver $r_i$. Then the probability that $r_1$ has a directed link with $r_i$ is

$$Pr(r_1 \rightarrow r_i) = \frac{d(r_1, r_i)^{-\gamma}}{\sum_{r_j \in R} [d(r_1, r_j)^{-\gamma}]}$$

where $d(.)$ represents the Manhattan distance between the two receivers. If $\gamma = 0$
this models a uniform distribution over long range links. As $\gamma$ increases, the probability of distanced receivers being part of the neighborhood decreases. Network links are therefore distributed according to an inverse $\gamma$-power distribution. By varying the parameter $\gamma$ one can examine the effect of varying degrees of global interrelatedness. This could represent a neighborhood incorporating the effects of modern telecommunication technologies.

A sender $s_i$’s payoff $\pi^i_s$ depends positively on the number of receivers deciding to engage with the sender. It also depends negatively on the cost of sending a signal. A receiver $r_j$’s payoff $\pi^j_r$ depends positively on the observed quality.

Let $c : N \times M \to R$ be $s_i$’s cost of sending a signal which maps the broadcasting radius and the signaling intensity to the payoff space. Suppose further that

$$\frac{\partial c}{\partial d} > 0 \quad \frac{\partial c}{\partial m} > 0 \quad \text{and} \quad \frac{\partial c}{\partial d m} > 0$$

The marginal cost of increasing the broadcasting radius, the signaling intensity or both is positive. Consider the following cost function satisfying these properties. It is increasing in the broadcasting radius and the intensity of the signal sent:

$$c(d, m) = \alpha(2d(d + 1) + 1)m$$

where $\alpha$ is a scalar accounting for the lack of prices in this model. Note that a signal with broadcasting radius $d$ reaches $2d(d + 1) + 1$ receivers. This follows immediately from the fact that when using Manhattan distances, a radius $d$ forms a diamond of receivers around the origin of the signal. Counting the number of receivers reached results in a centred square number which implies the formula above (cf. [7]).

The cost of setting up billboards opposite each consumer’s front lawn is then the cost of setting up billboards opposite each consumer’s front lawn within a radius $d$ around the firm. $\alpha$ takes into account that there are no prices in this model and as a result benefits and costs would not be directly comparable. Assuming that benefits are the numerator, $\alpha$ therefore scales costs accordingly.

Suppose each $s_i$ receives a payoff of one for each $r_j$ that engages with $s_i$. This implies the payoff function

$$\pi^i_s = n^i_s - \alpha(2d_s(d_s + 1) + 1)n^i_s$$

where $n^i_s$ is the number of receivers deciding to engage with the sender and $\alpha \in R$.

Suppose that engaging with sender $s_j$ is costless for the receiver $r_i$ and payoffs are simply given by

$$\pi^i_r = n^i_r$$

where $n^i_r$ is the number of receivers deciding to engage with the sender and $\alpha \in R$.

This models a uniform distribution over long range links. As $\gamma$ increases, the probability of distanced receivers being part of the neighborhood decreases. Network links are therefore distributed according to an inverse $\gamma$-power distribution. By varying the parameter $\gamma$ one can examine the effect of varying degrees of global interrelatedness. This could represent a neighborhood incorporating the effects of modern telecommunication technologies.

Suppose that engaging with sender $s_j$ is costless for the receiver $r_i$ and payoffs are simply given by

$$\pi^i_r = n^i_r$$

where $n^i_r$ is the number of receivers deciding to engage with the sender and $\alpha \in R$.
where $q^u_{ik}$ is the realization of the random variable $q^u_{ik}$ as a result of $r_i$, engaging with $s_k$.

The payoff is therefore either one if $r_i$ observes a good outcome or zero if he observes a bad outcome.

At each stage each $s_k$’s strategy consists of choosing a signaling intensity $v_k$ and a broadcasting radius $d_l$. Each $r_i$’s strategy consists of choosing a sender $s_k$ to engage with. To move from the stage game to the repeated game, I assume that agents have perfect recall. Additionally a sender $s_k$ exits the game if he has a non-positive payoff for $k$ rounds in a row. The exit condition therefore is

$$\pi^u_{st} <= 0 \ \forall \ t = \{u - k, \ldots , u\}$$

where $t$ denotes the time subscript and $u$ is the current period.

5 Approach

To analyze the game I use an agent-based computational instead of a classical game theoretic approach. There are three main benefits of using an agent-based computational approach as opposed to a classical game theoretic approach. Firstly, one can easily study out-of-equilibrium dynamics as one simulates the system’s dynamics. Secondly, it is also straightforward to incorporate myopic behavioral responses as responses have to be implemented in an algorithmic way. Both are features that are in line with the empirical results of how agents behave in experiments. Additionally, one can avoid to explicitly formulate equilibrium refinements, such as Cho (1987) equilibrium dominance, to narrow down the potential equilibrium outcomes of the game. Following Epstein (2012), the generative approach of agent-based models, helps to focus on those equilibria, which are feasibly achievable for agents with bounded rationality over reasonable time frames.

I build a complex adaptive system to solve for the outcome of the game for given set of parameters and realizations of random variables. I follow Tesfatsion’s (2006) simplest definition of a complex adaptive system in an agent-based computational context. A complex adaptive system is a system comprised of agents that are capable of reacting systematically to changes in their environment. For each agent I model behavioral rules and the respective informational constraints explicitly. In particular, I consider two types of senders to construct a counterfactual.

All senders are either quiet or adaptive. A quiet sender does not engage in signaling at all, whereas an adaptive sender changes broadcasting radius and signaling intensity dynamically if another sender engages with more receivers. In particular, senders who do not attract a sufficient number of receivers increase their signaling efforts in order to attract more senders.

Relating the general model to advertisements, firms either do not engage in advertising at all (quiet) or they advertise following a market share maximization strategy

where $q^u_{ik}$ is the realization of the random variable $q^u_{ik}$ as a result of $r_i$, engaging with $s_k$.

The payoff is therefore either one if $r_i$ observes a good outcome or zero if he observes a bad outcome.

At each stage each $s_k$’s strategy consists of choosing a signaling intensity $v_k$ and a broadcasting radius $d_l$. Each $r_i$’s strategy consists of choosing a sender $s_k$ to engage with. To move from the stage game to the repeated game, I assume that agents have perfect recall. Additionally a sender $s_k$ exits the game if he has a non-positive payoff for $k$ rounds in a row. The exit condition therefore is

$$\pi^u_{st} <= 0 \ \forall \ t = \{u - k, \ldots , u\}$$

where $t$ denotes the time subscript and $u$ is the current period.

5 Approach

To analyze the game I use an agent-based computational instead of a classical game theoretic approach. There are three main benefits of using an agent-based computational approach as opposed to a classical game theoretic approach. Firstly, one can easily study out-of-equilibrium dynamics as one simulates the system’s dynamics. Secondly, it is also straightforward to incorporate myopic behavioral responses as responses have to be implemented in an algorithmic way. Both are features that are in line with the empirical results of how agents behave in experiments. Additionally, one can avoid to explicitly formulate equilibrium refinements, such as Cho (1987) equilibrium dominance, to narrow down the potential equilibrium outcomes of the game. Following Epstein (2012), the generative approach of agent-based models, helps to focus on those equilibria, which are feasibly achievable for agents with bounded rationality over reasonable time frames.

I build a complex adaptive system to solve for the outcome of the game for given set of parameters and realizations of random variables. I follow Tesfatsion’s (2006) simplest definition of a complex adaptive system in an agent-based computational context. A complex adaptive system is a system comprised of agents that are capable of reacting systematically to changes in their environment. For each agent I model behavioral rules and the respective informational constraints explicitly. In particular, I consider two types of senders to construct a counterfactual.

All senders are either quiet or adaptive. A quiet sender does not engage in signaling at all, whereas an adaptive sender changes broadcasting radius and signaling intensity dynamically if another sender engages with more receivers. In particular, senders who do not attract a sufficient number of receivers increase their signaling efforts in order to attract more senders.

Relating the general model to advertisements, firms either do not engage in advertising at all (quiet) or they advertise following a market share maximization strategy

1For an introduction to the concepts of agent-based modeling please refer to Macal (2010).
(adaptive). If they do then whenever they do not capture the majority of the market, they increase their advertisement spend by increasing the broadcasting radius, the signaling intensity or both. By comparing the outcome of quiet senders to the outcome of adaptive senders one can assess whether there is any value in setting a signal in this game. If all senders are adaptive and there is value in signaling, higher quality senders should receive more engagements on average compared to a case with only quiet senders.

All receivers choose a sender to engage with in the following way. If a receiver observed a good outcome in the previous period he continues to engage with the same sender. This creates a strong inertia for high quality senders. High quality firms receive many repeat purchases as consumers continue to use the product they believe has a high quality. If a receiver observed a bad outcome in the previous period he searches for a new sender.

Receivers learn from past outcomes and take into account signals. Each receiver holds a belief about the true quality of each sender. This belief is updated using past outcome information from the neighborhood. Each receiver calculates a score for each sender which combines the expected value of a sender’s true quality and the observed signaling intensity. A receiver observes a sender’s current signaling intensity if and only if he is within the sender’s broadcasting radius. Each receiver who observed a bad outcome chooses to engage with the sender with the highest score.

Senders are heterogeneous in their position and inherent quality. There also exists heterogeneity among receivers in the sense that each receiver has his own neighborhood and spatial position. Therefore, the past outcome history and the observed signal will, in general, vary from receiver to receiver.

I use a Monte Carlo approach to generate a distribution of outcomes to draw inferences upon. In particular I simulate 100 different realizations of the system. In what follows I present the simulation results for $N_S = 10$ senders competing for $N_R = 3025$ receivers, i.e. a $55 \times 55$ torus, over 100 time periods. Simulation results might only exhibit meta-stability over these 100 time periods. However, this should not constitute a problem, as for most experience goods it is unlikely that the same agent buys the good more than 100 times over the agent’s lifetime. The radius of the von Neumann neighborhood is $v = 1$, both for the von Neumann neighborhood itself and for the von Neumann neighborhood within the Kleinberg neighborhood. Senders exit the game if they have a non-positive payoff for more than four rounds in a row ($k = 4$). While the exact realizations depend on these parameter values, simulations with different sets of parameters confirm that the main conclusions are robust to changes in the parameters.

Two Further work should increase the number of Monte Carlo draws. The restriction to 100 is due to limitations to computational performance.

Meta-stability refers to a complex system which is stable for prolonged time periods, yet diverges into chaos at a later point in time. (adaptive). If they do then whenever they do not capture the majority of the market, they increase their advertisement spend by increasing the broadcasting radius, the signaling intensity or both. By comparing the outcome of quiet senders to the outcome of adaptive senders one can assess whether there is any value in setting a signal in this game. If all senders are adaptive and there is value in signaling, higher quality senders should receive more engagements on average compared to a case with only quiet senders.

All receivers choose a sender to engage with in the following way. If a receiver observed a good outcome in the previous period he continues to engage with the same sender. This creates a strong inertia for high quality senders. High quality firms receive many repeat purchases as consumers continue to use the product they believe has a high quality. If a receiver observed a bad outcome in the previous period he searches for a new sender.

Receivers learn from past outcomes and take into account signals. Each receiver holds a belief about the true quality of each sender. This belief is updated using past outcome information from the neighborhood. Each receiver calculates a score for each sender which combines the expected value of a sender’s true quality and the observed signaling intensity. A receiver observes a sender’s current signaling intensity if and only if he is within the sender’s broadcasting radius. Each receiver who observed a bad outcome chooses to engage with the sender with the highest score.

Senders are heterogeneous in their position and inherent quality. There also exists heterogeneity among receivers in the sense that each receiver has his own neighborhood and spatial position. Therefore, the past outcome history and the observed signal will, in general, vary from receiver to receiver.

I use a Monte Carlo approach to generate a distribution of outcomes to draw inferences upon. In particular I simulate 100 different realizations of the system. In what follows I present the simulation results for $N_S = 10$ senders competing for $N_R = 3025$ receivers, i.e. a $55 \times 55$ torus, over 100 time periods. Simulation results might only exhibit meta-stability over these 100 time periods. However, this should not constitute a problem, as for most experience goods it is unlikely that the same agent buys the good more than 100 times over the agent’s lifetime. The radius of the von Neumann neighborhood is $v = 1$, both for the von Neumann neighborhood itself and for the von Neumann neighborhood within the Kleinberg neighborhood. Senders exit the game if they have a non-positive payoff for more than four rounds in a row ($k = 4$). While the exact realizations depend on these parameter values, simulations with different sets of parameters confirm that the main conclusions are robust to changes in the parameters.

Further work should increase the number of Monte Carlo draws. The restriction to 100 is due to limitations to computational performance.

Meta-stability refers to a complex system which is stable for prolonged time periods, yet diverges into chaos at a later point in time.
6 Analysis

To assess whether there is any value to spatial signaling, I analyze a counterfactual. How do outcomes change if senders do signal, compared to when they do not signal. To look at how the value of signaling changes as the system becomes more interconnected, I also consider how outcomes change when receivers expand their neighborhood.

Note first that under perfect information, no sender would engage in signaling, as all receivers would choose the sender with the highest quality, regardless of the signal. Therefore, to assess the value of spatial signaling a first starting point is to consider what happens to the market share of the highest quality sender as the system evolves. I then consider what happens to senders' payoffs in general. Finally, I analyze whether signaling improves receivers' payoffs, as it should if senders successfully reveal their private information.

Consider figure 2. This graph plots the mean market share of the highest quality sender over all Monte Carlo draws. Note that the true inherent quality of the highest quality sender varies across each Monte Carlo draw and therefore it is natural to group market share by rank as opposed to quality. The graph shows the market share both for a quiet sender and for an adaptive sender. For each sender type, it distinguishes three sub-cases: receivers have no neighborhood, a von Neumann neighborhood or a Kleinberg neighborhood. In period one the highest quality sender – just as any other sender – has an average market share of 10% (3025 receivers choosing uniformly one of 10 senders). As the system evolves, the average market share of the highest sender increases regardless of how receivers are connected.

Consider first the case when receivers have no neighborhood, i.e. they decide which sender to choose based only on their own outcome information and the signals that reach them. Then engaging in signaling results in a lower average market share (33.4% in round 100) compared to the case in which nobody signals (37.9% in round 100). While this might seem odd at first, breaking down the simulation results unveils that without networking effects signaling is too costly. A sender needs a large broadcasting radius to reach a sufficient number of receivers. This is very costly and results in the
highest quality sender exiting the game for a significant proportion of Monte Carlo draws, as receivers do not learn about the sender’s true quality quickly enough. Thus, the sender’s average market share drops.

When all receivers have a von Neumann neighborhood the world is more interconnected. As a receiver’s neighborhood expands, more outcome information is available and each receiver can better assess a sender’s quality. By making better informed decisions receivers switch to higher quality senders and continue to engage with them due to the inertia of repeat engagements in case a good outcome is observed. Hence average market share for the highest quality sender increases both for quiet (49% in round 100) and for adaptive senders (52.7% in round 100). Note, however, that average market share increases by a lot more for adaptive senders. This is because in addition to more outcome information being available, there is now a word-of-mouth effect through signaling. Suppose a high quality sender engages in signaling. Some receivers will switch to that sender upon observing that signal. As a result, these receivers will observe a good outcome with a high likelihood. Therefore, in the next round the update of beliefs using outcome information will be more favorable. More receivers will engage with the high quality sender. A low quality sender will accrue the cost of signaling, but is less likely to reap the benefit of network effects and will therefore eventually exit the game.

Average market share is even higher for the Kleinberg case (58.2% for quiet and 59.9% for adaptive senders in round 100). Part of that is attributable to the fact that a Kleinberg neighborhood consists of more receivers. But there is also the added benefit that long range links allow receivers to break out of localized patches, where the majority of receivers engage with higher quality senders which are just-good-enough to not switch away. This localized inertia due to specific outcome histories is also the reason that the highest sender’s market share does not converge to 100%. This is consistent with Cooper1997’s empirical finding that convergence towards a unique equilibrium in signaling games depends on the observed history of the game. I will consider localized inertia in more detail when analyzing receivers’ payoffs. Note that the Kleinberg neighborhood becomes more local as \( \alpha \) increases and simulations show that the long range effect indeed diminishes as \( \alpha \) increases.

This suggests that while there might be no value to signaling if receivers are isolated, there is a benefit of engaging in signaling if the world is interconnected.

To better assess whether there is value in signaling if the system is interconnected, consider the payoffs presented in figure 3. It displays the distribution of observed payoffs over the 100 Monte Carlo draws by rank for round four and for round 100. Note that while I only present the case for receivers with a von Neumann neighborhood, results for the Kleinberg neighborhood are qualitatively comparable. Again, as each Monte Carlo draw results in a different inherent quality and, as signaling cost does not directly depend on it, grouping by the senders’ ranking is natural.

In period one (not depicted) the distribution of payoffs is homogeneous across all highest quality sender exiting the game for a significant proportion of Monte Carlo draws, as receivers do not learn about the sender’s true quality quickly enough. Thus, the sender’s average market share drops.

When all receivers have a von Neumann neighborhood the world is more interconnected. As a receiver’s neighborhood expands, more outcome information is available and each receiver can better assess a sender’s quality. By making better informed decisions receivers switch to higher quality senders and continue to engage with them due to the inertia of repeat engagements in case a good outcome is observed. Hence average market share for the highest quality sender increases both for quiet (49% in round 100) and for adaptive senders (52.7% in round 100). Note, however, that average market share increases by a lot more for adaptive senders. This is because in addition to more outcome information being available, there is now a word-of-mouth effect through signaling. Suppose a high quality sender engages in signaling. Some receivers will switch to that sender upon observing that signal. As a result, these receivers will observe a good outcome with a high likelihood. Therefore, in the next round the update of beliefs using outcome information will be more favorable. More receivers will engage with the high quality sender. A low quality sender will accrue the cost of signaling, but is less likely to reap the benefit of network effects and will therefore eventually exit the game.

Average market share is even higher for the Kleinberg case (58.2% for quiet and 59.9% for adaptive senders in round 100). Part of that is attributable to the fact that a Kleinberg neighborhood consists of more receivers. But there is also the added benefit that long range links allow receivers to break out of localized patches, where the majority of receivers engage with higher quality senders which are just-good-enough to not switch away. This localized inertia due to specific outcome histories is also the reason that the highest sender’s market share does not converge to 100%. This is consistent with Cooper1997’s empirical finding that convergence towards a unique equilibrium in signaling games depends on the observed history of the game. I will consider localized inertia in more detail when analyzing receivers’ payoffs. Note that the Kleinberg neighborhood becomes more local as \( \alpha \) increases and simulations show that the long range effect indeed diminishes as \( \alpha \) increases.

This suggests that while there might be no value to signaling if receivers are isolated, there is a benefit of engaging in signaling if the world is interconnected.

To better assess whether there is value in signaling if the system is interconnected, consider the payoffs presented in figure 3. It displays the distribution of observed payoffs over the 100 Monte Carlo draws by rank for round four and for round 100. Note that while I only present the case for receivers with a von Neumann neighborhood, results for the Kleinberg neighborhood are qualitatively comparable. Again, as each Monte Carlo draw results in a different inherent quality and, as signaling cost does not directly depend on it, grouping by the senders’ ranking is natural.

In period one (not depicted) the distribution of payoffs is homogeneous across all

\(^{1}\)Kleinberg (2000) shows that the critical value for moving towards a local (Moore) neighborhood is \( \gamma > 2 \).

\(^{2}\)Kleinberg (2000) shows that the critical value for moving towards a local (Moore) neighborhood is \( \gamma > 2 \).
Figure 3: Sender payoff of by rank for round four (top) and 100 (bottom) without (Quiet) and with (Adaptive) signaling. Receivers have a von Neumann neighborhood, $\alpha = 0.2$, $k = 4$. 1 is the sender with the highest inherent quality, 10 is the sender with the lowest inherent quality.

senders. However, by period four the distribution of payoffs for higher ranked senders is already significantly higher than for lower ranked senders. As time progresses, receivers learn about a sender’s true quality by updating their beliefs using past outcomes. This is driving the shift of the distributions for quiet senders. Payoffs cannot be negative for quiet senders as they do not engage in signaling.

For adaptive senders, there is also a second driver at work. Signaling is costly and higher quality senders have a higher chance of repeat engagement and therefore of recouping their signaling costs. Note that in period four most of the lower ranked senders only observe negative outcomes. As a result, all senders who also had negative payoffs in the previous three periods will exit the game in the next period (as $k = 4$). Note that for some Monte Carlo outcomes even the highest quality sender has a negative payoff because he did not attract a sufficient number of receivers for the amount of signaling he engages in. The median payoff for the highest sender is, however, above the median payoff when there is no signal.

By period 100, lower quality senders have exited the game for both quiet and adaptive cases. While the range of outcomes for the highest quality sender’s payoff is similar for quiet and adaptive senders, the distribution is not. Payoffs are more centered around the median payoff for adaptive senders and the median payoff is higher. Related to the lower median payoff for the highest quiet sender is the fact, that in a significant number of outcomes the second highest sender is able to engage with a substantial number of receivers leading to fewer receivers engaging with the highest receiver. Without signaling it is more difficult for receivers to distinguish between the highest quality sender and the second highest quality sender who might be just slightly worse. This is due to the aforementioned localized inertia. If the second highest sender engages with the majority of receivers due to a favorable outcome history, the adaptive highest sender will increase his signaling level until the majority of receivers engages with him. Then

senders. However, by period four the distribution of payoffs for higher ranked senders is already significantly higher than for lower ranked senders. As time progresses, receivers learn about a sender’s true quality by updating their beliefs using past outcomes. This is driving the shift of the distributions for quiet senders. Payoffs cannot be negative for quiet senders as they do not engage in signaling.

For adaptive senders, there is also a second driver at work. Signaling is costly and higher quality senders have a higher chance of repeat engagement and therefore of recouping their signaling costs. Note that in period four most of the lower ranked senders only observe negative outcomes. As a result, all senders who also had negative payoffs in the previous three periods will exit the game in the next period (as $k = 4$). Note that for some Monte Carlo outcomes even the highest quality sender has a negative payoff because he did not attract a sufficient number of receivers for the amount of signaling he engages in. The median payoff for the highest sender is, however, above the median payoff when there is no signal.

By period 100, lower quality senders have exited the game for both quiet and adaptive cases. While the range of outcomes for the highest quality sender’s payoff is similar for quiet and adaptive senders, the distribution is not. Payoffs are more centered around the median payoff for adaptive senders and the median payoff is higher. Related to the lower median payoff for the highest quiet sender is the fact, that in a significant number of outcomes the second highest sender is able to engage with a substantial number of receivers leading to fewer receivers engaging with the highest receiver. Without signaling it is more difficult for receivers to distinguish between the highest quality sender and the second highest quality sender who might be just slightly worse. This is due to the aforementioned localized inertia. If the second highest sender engages with the majority of receivers due to a favorable outcome history, the adaptive highest sender will increase his signaling level until the majority of receivers engages with him. Then
the second highest receiver will try to fight back by in turn increasing his signaling level. The second highest receiver has, however, less capacity to increase his level of signaling compared to the highest sender as he will produce slightly more bad outcomes and therefore is more likely to exit the game.

This again suggests that there is a value to signaling: localized inertia of good outcomes from a just-good-enough sender can be overcome by signals from the highest quality sender.

Finally consider figure 4 which shows average receiver payoff over time when senders are adaptive for receivers with no neighborhood, a von Neumann neighborhood and a Kleinberg neighborhood. The earlier finding of more receivers engaging with higher quality senders over time implies an increasing average payoff for receivers. This is consistent with the simulation results. In the first round, average receiver payoff is 0.5 as receivers randomize over senders and each sender draws their inherent quality from a uniform distribution with expected value 0.5. As receivers engage with higher quality senders, the expected value of a good outcome increases. The learning process is slowest for receivers without a neighborhood. Thus, the initial increase in receiver payoff is comparatively slow compared to the two cases with social learning. Average payoff is above 0.8 within 5 rounds for receivers with a Kleinberg neighborhood and 6 rounds with a von Neumann neighborhood.

As noted before, localized inertia helps to explain why a Kleinberg neighborhood outperforms a von Neumann neighborhood and why there is value to signaling. Looking at receivers’ payoffs can explain how it arises. For ease of exposition, consider a sender with inherent quality 0.8. Suppose receiver \( r_i \) and the four receivers in his von Neumann neighborhood all engage with the sender. \( r_i \) observes a bad outcome whereas the four others observe a positive outcome. A Kleinberg neighborhood, which adds more receivers who are further away, improves the flow of information through the network and helps ameliorate this problem. This also explains why there is an upper bound to the receivers’ payoff. Once the average payoff has reached a certain level, fewer and fewer receivers observe a bad outcome and actually consider changing the sender.

Finally consider figure 4 which shows average receiver payoff over time when senders are adaptive for receivers with no neighborhood, a von Neumann neighborhood and a Kleinberg neighborhood. The earlier finding of more receivers engaging with higher quality senders over time implies an increasing average payoff for receivers. This is consistent with the simulation results. In the first round, average receiver payoff is 0.5 as receivers randomize over senders and each sender draws their inherent quality from a uniform distribution with expected value 0.5. As receivers engage with higher quality senders, the expected value of a good outcome increases. The learning process is slowest for receivers without a neighborhood. Thus, the initial increase in receiver payoff is comparatively slow compared to the two cases with social learning. Average payoff is above 0.8 within 5 rounds for receivers with a Kleinberg neighborhood and 6 rounds with a von Neumann neighborhood.

As noted before, localized inertia helps to explain why a Kleinberg neighborhood outperforms a von Neumann neighborhood and why there is value to signaling. Looking at receivers’ payoffs can explain how it arises. For ease of exposition, consider a sender with inherent quality 0.8. Suppose receiver \( r_i \) and the four receivers in his von Neumann neighborhood all engage with the sender. \( r_i \) observes a bad outcome whereas the four others observe a positive outcome. A Kleinberg neighborhood, which adds more receivers who are further away, improves the flow of information through the network and helps ameliorate this problem. This also explains why there is an upper bound to the receivers’ payoff. Once the average payoff has reached a certain level, fewer and fewer receivers observe a bad outcome and actually consider changing the sender.
Those who do, find mostly positive outcomes from one sender in their neighborhood. Signaling can help to counteract such a localized positive outcome history with a good-but-not-best sender. Note that the analysis above only applies for certain parameter values. In general, results relating to the value of signaling break down if the maximum number of rounds a sender can have a negative payoff before exiting the game is greater than five ($k > 5$). This is because signals then add too much noise as the punishment mechanism becomes too weak. If the cost parameter $\alpha$ increases over 0.25, signaling becomes too costly even for the highest quality sender. The system shows chaotic behavior as senders exit the game at random.

7 Conclusion

Using an agent-based computational approach I find that the spatial structure is important when analyzing signaling. In particular, simulation results suggest that in this model signaling is only desirable if receivers use information from their neighborhood. By leveraging the spatial structure senders can set signals in a cost-efficient way. They do not have to broadcast their signal to every single receiver to reveal their information. Instead, they can rely on word-of-mouth recommendations. Having a larger, more geographically dispersed neighborhood such as the Kleinberg neighborhood additionally improves the efficiency of the system by counteracting localized inertia. This gives rise to the hypothesis that as the internet became more prevalent, advertisements act more efficiently as signals.

These results support the observation that advertisement campaigns that leverage network effects (e.g. the Evian campaigns) can successfully act as signals, despite being a lot less expensive than global campaigns (e.g. the HTC campaign).

To make the simulation results more robust, further extensions of this model should consider introducing different behavioral rules and agent populations who are heterogeneous in their behavioral rules. Additionally this model currently only analyzes a game with more receivers than senders. Analyzing the converse might yield interesting results. Finally, advertisements often primarily act as a source of information and not as a signal. Initial simulations where receivers only engage with senders whose advertisements they receive, or who have engaged with someone in their neighborhood, show, that the signaling value diminishes in favor of an informational value but a more thorough analysis is needed.

Unlike a game theoretic approach, an agent-based approach has difficulty proofing statements with generality. Yet even without absolute generality, the simulations presented suggest that considering the spatial dimension of signaling, which has been neglected by economic research up to this point, can yield interesting insights.

---

5 To check the robustness of the model against changes in the specification of the game. I have also simulated the model using various different behavioral rules and cost and utility functions. The results suggest that similar overall conclusions hold at least for part of the respective parameter spaces.
8 Bibliography


THE INDIAN INFLATION PUZZLE: IS THE RESERVE BANK OF INDIA USING THE RIGHT PRICE INDEX FOR ITS HEADLINE INFLATION MEASURE?

SAUGATA SEN

London School Of Economics

ABSTRACT

The Reserve Bank of India (RBI) is one of the few Central Banks in the world to use a Wholesale Price Index (WPI), and not a Consumer Price Index (CPI) as its headline inflation rate. This approach has attracted many criticisms from audiences in India and debate continues as to whether this is appropriate for monetary policy purposes. This paper takes a close look at the Indian WPI measure, and seeks to empirically address three key questions concerning its use. First, we consider whether the choice of which price index to use in India has been relevant. To do so, we use a cointegration and vector autoregression approach to investigate the long-run and short-run dynamic relationships between the WPI and major CPI indices. Second, we investigate the representation of consumer-based inflation within the WPI, by using vector autoregressions and impulse response functions to look at a sector-wide breakdown of WPI inflation. Third, we examine whether international price pressures have become more prominent in contributing towards WPI inflation post-economic liberalisation in 1991 through the use of a vector error-correction model. This paper finds that there are neither short-run dynamics nor long-run trends that relate the different price indices in India. The manufacturing sector is shown to be the strongest and most persistent driving force behind WPI inflation. In addition, there is strong evidence that both imported inflation and domestic demand pressures have had greater influence in causing WPI inflation post-economic liberalisation. This paper concludes that the WPI has been a very problematic inflation measure for the RBI to use for its monetary policy. Instead, attention should now switch towards potentially adopting a CPI measure as the new headline inflation rate.
ON THE OTHER SIDE OF THE TRACKS: NEIGHBORHOOD EFFECTS AND SEGREGATION IN MOROCCO

RUSSELL MORTON
Princeton University

ABSTRACT

This paper examines the impact of neighborhood effects using a unique data set from rural Morocco and finds strong evidence that poor households do shift their actions, specifically in the domains of social group membership and investment in education in the next generation, due to living near wealthier households. The estimation of peer effects focuses on comparing the outcomes of two groups shown to be similar across covariates that vary in the distance between their house and that of a third group, specifically those living in the homes of the former European colonists of the town. This identification strategy is rooted in the history of Khouribga, as the segregation of the European and Moroccan residents of Khouribga during the colonial period on separate sides of the railroad track running through town has resulted in the concentration of Khouribga’s wealthiest residents today inside the area where the European colonists used to live. The results suggest that the neighborhood effects contribute to increased participation by adults in work-related professional groups and higher levels of investment in human capital in the next generation of sons. Further support for the presence of neighborhood effects derives from evidence that households where the father participates in a professional group invest even more in the education of their sons, and residents of the former European households are also much more likely to belong to a professional group, facilitating a higher level of interaction between the two groups. Robustness checks focus on two distinct areas. First, alternative hypotheses that may explain this result, such as access to local public goods and other explanations for the limited educational investment by the North are shown to lack support in the data.
Abstract: Previous research has suggested that, in the United States, partisan differences in excess returns are quantitatively large and cannot be explained by partisan differences in risk. In this paper, I argue that partisan differences in excess returns are quantitatively small and can plausibly be explained by partisan differences in risk. I present three pieces of evidence in favor of this argument: First, on the basis of international data, one can strongly reject the hypothesis that partisan differences in excess returns are quantitatively as large as the U.S. time-series data would suggest. Second, average excess returns and the volatility of excess returns have both been higher under social democratic governments than under conservative governments. Third, on the basis of option market movements during the 2004 and 2012 U.S. presidential elections, it appears that financial markets were expecting greater future volatility under a Democratic president than under a Republican.
INEQUALITY AND UNREST: COULD LOCAL INEQUALITY INCREASE MARGINAL PROPENSITY TO STRIKE?

ALYSSA HUBERTS
Georgetown University

ABSTRACT
Does living in an unequal society impact our prospects for social stability? In this paper, I ask if municipalities with greater income inequality experience greater social unrest, considering strikes at the municipal level. I focus my research on Mexico, a country that has experienced both tremendous economic growth and high levels of inequality over the past two decades. I find that local inequality is correlated with more industrial unrest, manifested by significantly more worker strikes in societies with higher levels of inequality. I also find initial confirmation that at the highest levels of inequality; the likelihood of strikes may actually be reduced, due either to limitations in capabilities or the likelihood of comparison between individuals in different income strata.

1 Introduction
Does living in an unequal society impact our prospects for social stability? While poverty or health indicators reflect levels of absolute deprivation, studies of inequality attempt to determine the way that our position in relation to the rest of society impacts our welfare. This comparative character makes it difficult to understand inequality’s direct and indirect consequences. Yet changes in local income distributions can have great consequences on social and industrial stability, leading to important policy and business implications. In this paper, I ask if municipalities with greater income inequality experience greater unrest. I focus my research on Mexico, a country that has experienced both tremendous economic growth and high levels of inequality over the past two decades. I find that local inequality is correlated with more industrial unrest, manifested by significantly more worker strikes in societies with higher levels of inequality. I also find
confirmation that at the highest levels of inequality the likelihood of strikes may actually be reduced, due either to limitations in capabilities or the likelihood of comparison between individuals in different income strata.

2 The Question

Debates about the impact of inequality usually fall along two axes: deep-rooted questions of ethics and the more pragmatic issues of capability. Beyond the ethical questions of an equitable distribution of income, it is often true that structural disparities in the income distribution can hold growing economies back from achieving their full potential. Most commonly, authors suggest that inequality limits the consumption capabilities of an impoverished portion of the population. Here, I consider a second potential consequence of high inequality: industrial unrest.

The issue of inequality has plagued Latin America since the colonial era. Extractive institutions established by colonial powers, race and class-based structures and opportunities, and the disparity between urban and rural development have made Latin America the most unequal region in the world. Export-oriented industrialization policies in recent years have had mixed impacts on this inequality — exacerbating it in some places and mitigating it in others. In recent decades, social policies have attempted to mitigate some of this inequality, be it via sweeping redistributive measures or more moderate changes in social welfare or education policies. In this paper, I focus on Mexico between the years of 2000 and 2005, during which levels of municipal inequality decreased drastically. I test whether these changes in inequality were paralleled with decreases in the industrial unrest manifested by strikes.

3 Background Information: Mexico

Since the 1994-95 Peso Crisis, when the federal government was forced to devalue the national currency in response to massive investor flight, Mexico has undergone a number of important reforms and experienced a period of sustained high economic growth. With the exception of temporary drops in

2001 and 2009, Mexico has grown steadily, with GNP per capita more than doubling since 1990. Much of this growth is a consequence of high levels of foreign direct investment after the signing of the North American Free Trade Agreement (NAFTA), with particular emphasis on the investment in maquilas, factories in free trade zones along the United States border.

Yet despite this significant growth, Mexico remains a profoundly unequal society. In 2010, the World Bank calculated Mexico’s Gini Coefficient at 47.4. In response to high levels of societal inequality, the government of Mexico has undertaken a number of programs for social spending and redistribution. The most famous of these, Oportunidades, was introduced in 1997 as Progresa, a program that required women to attend regular health checkups and send their children to school in order to receive conditional bi-monthly cash transfers. Progresa has since been dramatically expanded and renamed Oportunidades. Explicitly designed to mitigate inequality in human development and social welfare, Oportunidades today serves 6.5 million families across every Mexican municipality.

The data shows that at the municipal level, Mexican inequality is indeed decreasing. Of the 2,454 municipalities in the dataset, 1,873, or 76.3%, experienced a decrease in their Gini coefficient between 2000 and 2005. Considering the common perception that inequality is highly structural and usually requires large periods of time to change, Mexico presents an especially interesting testing ground for theories on the impact of inequality on strike levels.

4 Literature Review

This paper presents the first attempt to my knowledge to incorporate inequality into our understanding of strikes. Here, I outline the common approaches to strike research, and then discuss the ways inequality and relative deprivation literature shape my hypothesis.

Strike literature has traditionally taken two forms. At the micro level, authors attempt to explain why workers strike in a particular given case. The

---

5 Rowe, Nicholas, “Mexico’s Oportunidades: Conditional Cash Transfers as the Solution to Global Poverty?”
6 “Oportunidades, A Human Development Program” http://www.oportunidades.gob.mx/Portal/wb/Web/oportunidades_a_human_development_program

2001 and 2009, Mexico has grown steadily, with GNP per capita more than doubling since 1990. Much of this growth is a consequence of high levels of foreign direct investment after the signing of the North American Free Trade Agreement (NAFTA), with particular emphasis on the investment in maquilas, factories in free trade zones along the United States border.

Yet despite this significant growth, Mexico remains a profoundly unequal society. In 2010, the World Bank calculated Mexico’s Gini Coefficient at 47.4. In response to high levels of societal inequality, the government of Mexico has undertaken a number of programs for social spending and redistribution. The most famous of these, Oportunidades, was introduced in 1997 as Progresa, a program that required women to attend regular health checkups and send their children to school in order to receive conditional bi-monthly cash transfers. Progresa has since been dramatically expanded and renamed Oportunidades. Explicitly designed to mitigate inequality in human development and social welfare, Oportunidades today serves 6.5 million families across every Mexican municipality.

The data shows that at the municipal level, Mexican inequality is indeed decreasing. Of the 2,454 municipalities in the dataset, 1,873, or 76.3%, experienced a decrease in their Gini coefficient between 2000 and 2005. Considering the common perception that inequality is highly structural and usually requires large periods of time to change, Mexico presents an especially interesting testing ground for theories on the impact of inequality on strike levels.

4 Literature Review

This paper presents the first attempt to my knowledge to incorporate inequality into our understanding of strikes. Here, I outline the common approaches to strike research, and then discuss the ways inequality and relative deprivation literature shape my hypothesis.

Strike literature has traditionally taken two forms. At the micro level, authors attempt to explain why workers strike in a particular given case. The
category of strike literature relevant to my analysis takes the macro-level, often empirically-rooted approach, attempting not to explain a single strike scenario but instead the societal factors which make strikes as a social phenomenon more or less likely.

The most seminal paper in this macro strike literature is Haas and Stack (1983), which was the first serious attempt to identify whether late industrial development leads to an increase or decrease in strikes. The authors contrast the “liberal” argument with the “radical” perspective. The liberal perspective suggests that strikes will disappear as industrialization moves in the direction of big business. With management shifting from profit-maximizing owners to technocrats whose aim is predictability in production, workers’ and managers’ interests will begin to overlap more and resources will become available to meet workers’ wage demands. Furthermore, increased wages, job security, and social mobility will lead to decreased unionization and increased heterogeneity of interests.\(^7\)

This liberal perspective is contrasted with the “radical” position that strikes will increase with economic development, which argues that income growth shifts resources to the labor force in a way that makes it easier for them to mobilize. According to authors in this tradition, late industrialization has led to a decrease in working conditions for white collar workers, enabling them to take up common cause with blue collar workers while small scale artisans and farmers are simultaneously being driven out of business and forced to join the industrial labor force. Finally, the expectations argument says that with increased development, workers’ expectations become increasingly unrealistic, leading them to strike even more.\(^7\)

Beyond their neat outline of the two perspectives in the income question for strikes, Haas and Stack conduct econometric, cross-country analysis and find a curvilinear relationship between income and strikes, which they argue supports the liberal perspective. James (1996) weakly confirms their work, but James simultaneously raises the difficulty of attributing causality in strikes to any single variable.\(^8\)

I include the Haas and Stack analysis because the relationship between economic development and strikes lays the foundation for the question of inequality. I believe, however, that the conflict between the liberal and radical perspectives shows the difficulty of attributing causality in strikes to any single variable.\(^8\)

\(^8\) Ibid, 47.
radical perspectives they outline lies in their failure to consider the
distribution of income. By considering only the level of industrialization,
these theories miss the possibility that it is not the society’s wealth as a
whole, but the allocation of this wealth, that creates more or less unrest.

No literature currently links strikes directly to local inequality, but a number
of authors who write about inequality provide useful context, particularly
from a methodological perspective. The most complex problem is the means
by which one determines the grouping or level at which to measure
inequality. In order to identify a meaningful relationship, inequality and
strikes must be measured at the same level. This specification is the reason
that I conduct my analysis using municipal-level information for both my
independent and dependent variables.

Elbers et al (2004) also address the challenges of measuring inequality at the
local level, specifically in poor communities. The paper is an attempt to
identify the ways in which inequality affects program targeting in poor
communities. It provides a useful model on using census and survey data
together, and demonstrates that higher levels of local inequality can be
associated with less effective administration of resources, as local elites
capture the gains meant for society’s poorest. This finding could mean that
the effect of low wages is compounded with an ineffective distribution of
government resources, and frustration over this allocation could prove to be a
mechanism for a relationship between inequality and unrest.

The closest intersection of strike theory and questions of inequality comes
from the theory of relative deprivation, presented by Panning in 1983.11
Drawing on Nagel and the body of literature about inequality and the
likelihood to support revolution, Panning develops a formal model for
relative deprivation conditional on two assumptions:

1. “The more unequal the assets of two individuals, the less happy the poorer will feel if he compares himself with the richer.”

2. The greater the distance in wealth between two individuals, the less likely they are to compare themselves to one another.

Elbers, Chris, “On the Unequal Inequality of Poor Communities,” The World Bank

Panning, William H., “Inequality, Social Comparison and Relative Deprivation,” The

Elbers, Chris, “On the Unequal Inequality of Poor Communities,” The World Bank

Panning, William H., “Inequality, Social Comparison and Relative Deprivation,” The
Thus Panning presents a testable hypothesis – does reducing inequality reduce relative deprivation? This hypothesis can be adapted to my project by asking whether Mexican municipalities with lower inequality experience less industrial unrest. Specific to the second of the above claims, I also test a second hypothesis, which suggests a non-linear relationship. This hypothesis suggests that strikes will be low at extremely low and extremely high levels of inequality, and highest at the intermediate level.

While the question is understudied in the literature, relatively robust arguments emerge to justify each of two conflicting perspectives: 1) that increased worker solidarity will make strikes more likely in societies where the wealth is closely distributed, and 2) that the sentiment of relative deprivation will be sufficiently strong to motivate more strikes in more unequal societies. It is my hypothesis that the relative deprivation argument proves stronger than the solidarity perspective, and that less egalitarian municipalities will exhibit more strikes.

5 Data and Summary Statistics

My data comes from a number of sources, relying heavily on the Mexican Instituto Nacional de Geografía y Estadísticas (INEGI). Due to both historical context and data availability, I choose to focus on the years 2000 and 2005. This period spans the presidency of Vicente Fox and the 2002 expansion of Progresa to its more broadly based form, Oportunidades. In part because of this expansion, this particular five-year span covers significant change in social inequality and makes use of two years for which, due to the Mexican census and surveys, data on both inequality and strikes exists.

My data on strikes, the manufacturing sector and municipal finance are taken from INEGI’s Sistema Estatal y Municipal de Base de Datos (SIMBAD).13 Data on inequality comes from Mexico’s Consejo Nacional de la Evaluación

12 Panning’s model for the difference between curvilinear and linear applications is quite complex, and has to do with the likelihood that particular elements of society compare themselves. Here I believe there is enormous potential for further insight from casework, as a closer look at the culture, class dynamics, and geography of municipalities may help gauge these types of likelihoods. However, for the purpose of this paper I limit this analysis to simply the test of a squared term for the coefficient on inequality.


13 Panning’s model for the difference between curvilinear and linear applications is quite complex, and has to do with the likelihood that particular elements of society compare themselves. Here I believe there is enormous potential for further insight from casework, as a closer look at the culture, class dynamics, and geography of municipalities may help gauge these types of likelihoods. However, for the purpose of this paper I limit this analysis to simply the test of a squared term for the coefficient on inequality.

de la Politica del Desarrollo Social (CONEVAL) for 2000 and 2005\textsuperscript{14}. Finally, my proxy for economic growth – satellite imagery – comes from J. Vernon Henderson’s replication data, originally used for “Measuring Economic Growth from Outer Space.”\textsuperscript{15}

My dependent variable, Total Strikes, comes from SIMBAD and constitutes the number of strikes occurring in the municipality in the given year. The distribution of strikes shows significant zero inflation; while about 600 municipalities exhibited strikes in 2000 and 2005, the remaining 1900-odd municipalities have 0 recorded strikes. When some municipalities have populations of as low as 207 or lack manufacturing sectors this is to be expected; nonetheless, it does run the risk of biasing the model significantly.

In the section on methodology I articulate the variety of approaches I apply in order to address this concern.

For my independent variable of interest, inequality, there are several possible options. I consider the Gini Coefficient, the Kuznets 10% ratio, and the Kuznets 5% ratio in my analysis. All three are available from CONEVAL. Gini coefficients decreased significantly between 2000 and 2005. While a handful of coefficients still remain above .5 in 2005, there is a significant shift in the distribution toward the lower end of the scale, with the mean in 2005 at .4161. Thus, if my hypothesis is correct, we should expect to see significantly fewer strikes in 2005 than in 2000.

The most important of my controls, especially considering the understanding promoted by Haas and Stack, is the control for economic development. At the municipal level, information about local GDP and economic growth is limited at best, and missing for most municipalities. Instead, following an approach introduced by J. Vernon Henderson in his National Bureau of Economic Research paper, “Measuring Growth from Outer Space”, I use satellite imagery data as a proxy for economic growth.\textsuperscript{16} Using geo-located satellite imagery of Mexico merged with a shapefile of Mexican municipalities,\textsuperscript{17} I am able to gauge an average light intensity, scored from 1-

---

\textsuperscript{14} Consejo Nacional de La Evaluación de la Política del Desarrollo Social http://www.coneval.gob.mx/evaluacion/Paginas/Evaluacion.aspx.


\textsuperscript{17} Marco Geostatistica 2000 (Censo de Población y Vivienda) http://www.inegi.org.mx/geo/contenidos/geostadistica/M_Geoestadistico.aspx

---
63, for each municipality. As this score itself is not usefully translated into a measurable dollar figure, I use the change over five-year periods to calculate an approximate growth rate for 1995-2000 and 2000-2005. As Henderson argues, the growth rate of light intensity is less accurate for capturing year-to-year variation but can be a good proxy for increased economic development over time. Thus even at the micro-level of municipalities, I am able to control for growth rates of economic development.

Finally, I include a number of other controls that could be responsible for the variation in strike levels. Manufacturing Sector divides the portion of the population employed in the manufacturing sector by the total municipal population, in order to account for the increased likelihood of strikes in manufacturing-heavy industries. This data comes from the INEGI Economic Census in 1994 and 2004. While this means that the data is lagged by a year for each, I consider it a generally reliable representation of manufacturing levels for 2000 and 2005. Finally, in the sensitivity analysis, I include regional dummies for eight different regions, a dummy for state capitals, and a dummy that reflects whether the municipality is one of the 16 subdivisions of Mexico City.

6 Methodology

My regression methodology is informed mostly by the zero-inflated nature of the dataset. Because about three quarters of municipalities exhibit zeros for Total Strikes, a standard Ordinary least Squares (OLS) regression would inaccurately reflect the distribution, biasing coefficients downward and risks falsely displaying diminished significance levels.

Standard methodology for approaching a zero-inflated set of count data usually assumes that the data takes some form of Poisson distribution, with the mean increasing in accordance with the variance, or the negative binomial distribution, which assumes similar characteristics as a Poisson distribution but allows for a variance far greater than the mean. When the cause of the over dispersion is thought to be “excess zeros,” however, with some outside characteristic motivating “excess zeros” in addition to “true zeros,” the Zero Inflated Poisson (ZIP) method of regression is commonly adopted.18  


Finally, I include a number of other controls that could be responsible for the variation in strike levels. Manufacturing Sector divides the portion of the population employed in the manufacturing sector by the total municipal population, in order to account for the increased likelihood of strikes in manufacturing-heavy industries. This data comes from the INEGI Economic Census in 1994 and 2004. While this means that the data is lagged by a year for each, I consider it a generally reliable representation of manufacturing levels for 2000 and 2005. Finally, in the sensitivity analysis, I include regional dummies for eight different regions, a dummy for state capitals, and a dummy that reflects whether the municipality is one of the 16 subdivisions of Mexico City.

6 Methodology

My regression methodology is informed mostly by the zero-inflated nature of the dataset. Because about three quarters of municipalities exhibit zeros for Total Strikes, a standard Ordinary least Squares (OLS) regression would inaccurately reflect the distribution, biasing coefficients downward and risks falsely displaying diminished significance levels.

Standard methodology for approaching a zero-inflated set of count data usually assumes that the data takes some form of Poisson distribution, with the mean increasing in accordance with the variance, or the negative binomial distribution, which assumes similar characteristics as a Poisson distribution but allows for a variance far greater than the mean. When the cause of the over dispersion is thought to be “excess zeros,” however, with some outside characteristic motivating “excess zeros” in addition to “true zeros,” the Zero Inflated Poisson (ZIP) method of regression is commonly adopted.18  

In order to test a variety of specifications for the model and to take full advantage of the panel dataset, I begin with a basic Difference-in-Difference Poisson regression. I find that my results overwhelmingly confirm my predictions about the positive relationship between strikes and inequality but do not confirm a logarithmic relationship. Noting that the zero-bias of the model may be making incorrect assumptions about the data distribution, however, I consider the Zero Inflated Poisson as my baseline model. Here, manufacturing proves a highly effective predictor for zero inflation, suggesting that the “excess zeros” in the model are a specific consequence of municipalities lacking a manufacturing sector. While the ZIP specification reduces the dataset to a cross-country rather than difference-in-difference approach (thereby eliminating the possibility of fixed effects) I do cluster standard errors by state.

Finally, because the mean is indeed far smaller than the variance for strikes, I adopt a negative binomial model in my robustness checks. The negative binomial model remains in my robustness checks rather than my principal results because the zero-bias in the dataset proves to be specifically and significantly tied to manufacturing, making the Zero Inflated Poisson approach more appropriate. That said, I find the direction of the effects in the negative binomial model to be largely consistent with my findings.

7 Results

7.1 Difference-In-Difference Poisson Model

In the first stage of tests, I assess a difference-in-difference Poisson model comparing the effect of change in inequality 2000 – 2005 on the change in strikes during the same period. Table 1 (Regressions 1 and 2) uses Gini Coefficients as measures for inequality, with Regression 1 testing Gini and Regression 2 incorporating a squared term. Table 2 (Regressions 3 and 4) and Table 3 (Regressions 5 and 6) do the same for the Kuznets Ratio at 10% and 5%, respectively. Here, coefficients are followed in the next column with the Incidence Rate Ratio for the benefit of interpretation.18

The results from the difference-in-difference Poisson model provide preliminary confirmation of the basic hypothesis that the change in inequality has been positively associated with strikes. Both the coefficient on Gini itself and that on the interaction term (Gini×2005) are positive and significant. The coefficient on the Gini×2005 is negative in accordance with In a Poisson distribution, the count data can be perceived as the incident rate, or rate of occurrence in a given period of time, of a given event.

In order to test a variety of specifications for the model and to take full advantage of the panel dataset, I begin with a basic Difference-in-Difference Poisson regression. I find that my results overwhelmingly confirm my predictions about the positive relationship between strikes and inequality but do not confirm a logarithmic relationship. Noting that the zero-bias of the model may be making incorrect assumptions about the data distribution, however, I consider the Zero Inflated Poisson as my baseline model. Here, manufacturing proves a highly effective predictor for zero inflation, suggesting that the “excess zeros” in the model are a specific consequence of municipalities lacking a manufacturing sector. While the ZIP specification reduces the dataset to a cross-country rather than difference-in-difference approach (thereby eliminating the possibility of fixed effects) I do cluster standard errors by state.

Finally, because the mean is indeed far smaller than the variance for strikes, I adopt a negative binomial model in my robustness checks. The negative binomial model remains in my robustness checks rather than my principal results because the zero-bias in the dataset proves to be specifically and significantly tied to manufacturing, making the Zero Inflated Poisson approach more appropriate. That said, I find the direction of the effects in the negative binomial model to be largely consistent with my findings.

7 Results

7.1 Difference-In-Difference Poisson Model

In the first stage of tests, I assess a difference-in-difference Poisson model comparing the effect of change in inequality 2000 – 2005 on the change in strikes during the same period. Table 1 (Regressions 1 and 2) uses Gini Coefficients as measures for inequality, with Regression 1 testing Gini and Regression 2 incorporating a squared term. Table 2 (Regressions 3 and 4) and Table 3 (Regressions 5 and 6) do the same for the Kuznets Ratio at 10% and 5%, respectively. Here, coefficients are followed in the next column with the Incidence Rate Ratio for the benefit of interpretation.19

The results from the difference-in-difference Poisson model provide preliminary confirmation of the basic hypothesis that the change in inequality has been positively associated with strikes. Both the coefficient on Gini itself and that on the interaction term (Gini×2005) are positive and significant. The coefficient on the Gini×2005 is negative in accordance with In a Poisson distribution, the count data can be perceived as the incident rate, or rate of occurrence in a given period of time, of a given event.
the assumptions of Panning’s model; however, it is not significant. The signs and significance of these coefficients are consistent across all three measures of inequality, providing some initial indication of robustness. However, it is important to note that the magnitude is highest in Gini and lowest for the Kuznets 5% ratio. This in itself may provide some weak confirmation of Panning’s assumption. The 5% ratio represents the ratio between the income shares at the 5% and 95% income levels – segments of society which arguably have almost no interaction. By comparison, the Gini coefficient compares at every income level the degree to which the distribution of income is below the level of perfect equality. Thus I have initial confirmation that there is some strong positive relationship between inequality and strikes, with an indication that the direction of this relationship could have to do with the specific level at which inequality is measured.

<table>
<thead>
<tr>
<th>Table 1: Difference-In-Difference Poisson (Gini)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Total Strikes &amp; Total Strikes &amp; Total Strikes</td>
</tr>
<tr>
<td>Gini &amp; 2.411*** &amp; 11.154*** &amp; -15.040** &amp; 2.94x10**</td>
</tr>
<tr>
<td>2005 &amp; -1.293*** &amp; .275*** &amp; -.798 &amp; 0.4501</td>
</tr>
<tr>
<td>Gini<em>2005 &amp; 2.009</em>** &amp; 7.458*** &amp; -.1985 &amp; 0.137</td>
</tr>
<tr>
<td>Gini2 &amp; --- &amp; --- &amp; 17.914** &amp; 550.44**</td>
</tr>
<tr>
<td>Gini2*2005 &amp; --- &amp; --- &amp; -1.985 &amp; 6.03x107</td>
</tr>
<tr>
<td>Population &amp; .0002*** &amp; .0002*** &amp; .0002*** &amp; 1.003***</td>
</tr>
<tr>
<td>Growth (Light) &amp; 0.013 &amp; 1.013 &amp; 0.0113 &amp; 1.011</td>
</tr>
<tr>
<td>Social Spending &amp; -.0001 &amp; .0999 &amp; -.0002 &amp; .9998</td>
</tr>
<tr>
<td>Manufacturing Sector &amp; -3.653 &amp; .026*** &amp; -3.545*** &amp; 0.0289***</td>
</tr>
<tr>
<td>N &amp; 1242 &amp; 1242 &amp; 1242 &amp; 1242</td>
</tr>
<tr>
<td>Fixed Effects &amp; Yes &amp; Yes &amp; Yes &amp; Yes</td>
</tr>
<tr>
<td>IRR &amp; No &amp; Yes &amp; No &amp; Yes</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001

For the scope of this paper, I focus on direction and significance, rather than magnitude of my coefficients.

Table 1: Difference-In-Difference Poisson (Gini)

For the scope of this paper, I focus on direction and significance, rather than magnitude of my coefficients.
Table 2: Difference-in-Difference Poisson (Kuznets 10%)

<table>
<thead>
<tr>
<th></th>
<th>Total Strikes</th>
<th>Total Strikes</th>
<th>Total Strikes</th>
<th>Total Strikes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kuznets 10%</td>
<td>.0123***</td>
<td>1.012***</td>
<td>.0155***</td>
<td>1.016***</td>
</tr>
<tr>
<td>2005</td>
<td>-.913***</td>
<td>0.4013***</td>
<td>-.9682***</td>
<td>.3798***</td>
</tr>
<tr>
<td>Kuznets 10%*2005</td>
<td>.0303***</td>
<td>1.031***</td>
<td>.033***</td>
<td>1.037***</td>
</tr>
<tr>
<td>Kuznets 10%</td>
<td>-.392x10^-9</td>
<td>1</td>
<td>-.14x10^-8</td>
<td>1</td>
</tr>
<tr>
<td>Population</td>
<td>.00026***</td>
<td>1***</td>
<td>.0026***</td>
<td>1.0026***</td>
</tr>
<tr>
<td>Growth (Light)</td>
<td>.0192*</td>
<td>1.019*</td>
<td>.01889*</td>
<td>1.019*</td>
</tr>
<tr>
<td>Social Spending</td>
<td>-.0008</td>
<td>0.9999</td>
<td>-.0004</td>
<td>0.9999</td>
</tr>
<tr>
<td>Manufacturing Sector</td>
<td>-3.581***</td>
<td>0.0279***</td>
<td>-3.5816***</td>
<td>0.0281***</td>
</tr>
<tr>
<td>N</td>
<td>1242</td>
<td>1242</td>
<td>1242</td>
<td>1242</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>IRR</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 3: Difference-In-Difference Poisson (Kuznets 5%)

<table>
<thead>
<tr>
<th></th>
<th>Total Strikes</th>
<th>Total Strikes</th>
<th>Total Strikes</th>
<th>Total Strikes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kuznets 5%</td>
<td>.0046***</td>
<td>1.005***</td>
<td>.0117***</td>
<td>1.912***</td>
</tr>
<tr>
<td>2005</td>
<td>-.8725</td>
<td>.418***</td>
<td>-.807***</td>
<td>0.4463***</td>
</tr>
<tr>
<td>Kuznets 5%*2005</td>
<td>.0156***</td>
<td>1.016*</td>
<td>.013***</td>
<td>1.0134***</td>
</tr>
<tr>
<td>Kuznets 5%</td>
<td>-6.43x10^-9</td>
<td>1</td>
<td>-6.43x10^-9</td>
<td>1</td>
</tr>
<tr>
<td>Kuznets 5%*2005</td>
<td>-3.27x10^9</td>
<td>1</td>
<td>-3.27x10^9</td>
<td>1</td>
</tr>
<tr>
<td>Population</td>
<td>.0003***</td>
<td>1.002***</td>
<td>.00261***</td>
<td>1.00261***</td>
</tr>
<tr>
<td>Growth (Light)</td>
<td>.0235*</td>
<td>1.024*</td>
<td>.0231*</td>
<td>1.0234*</td>
</tr>
<tr>
<td>Social Spending</td>
<td>-.0000257</td>
<td>0.9999</td>
<td>-.00003</td>
<td>0.9999</td>
</tr>
<tr>
<td>Manufacturing Sector</td>
<td>-3.563***</td>
<td>0.02835***</td>
<td>-3.574***</td>
<td>0.0281***</td>
</tr>
<tr>
<td>N</td>
<td>1242</td>
<td>1242</td>
<td>1242</td>
<td>1242</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>IRR</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

7.2 Zero Inflated Poisson Model

Table 4 presents the results of the Zero Inflated Poisson Model, with standard errors in parentheses and clustered standard errors in brackets. The first, and most important conclusion, to draw from this analysis is the finding that manufacturing serves as a strongly significant (p<0.001) predictor for
zero inflation in every iteration of the model. Intuitively, a lack of manufacturing proves to be a strong predictor of zeros.

When the clustered standard errors by state are incorporated, Regression 8, which uses Gini for inequality and includes a squared coefficient, emerges as the most robust. All regressions, however, point to the same conclusion: the relationship between inequality and strikes is positive at moderate levels (unsquared coefficient) and negative at the extremes (squared coefficient). This result is consistent across all three inequality measures, highly significant with unclustered standard errors, and remains significant with clustered standard errors for the Gini Coefficient. Thus with reasonable certainty I argue that there is a statistically significant nonlinear relationship between strikes and inequality.

With respect to the other coefficients included, signs are consistently as expected. The coefficient on population (divided by 100) is positive and significant at all levels, meaning that more populous societies strike more. The coefficient on the year 2005 is negative, a finding which is consistent with the general trend of fewer strikes in 2005. The coefficient on growth (or change in light intensity) is sometimes significant and always negative. This finding confirms Haas and Stack’s liberal perspective, which argues that strikes decrease with increased economic development, but its weakness supports my argument that it is the distribution of income, rather than the level of economic development, which matters for the likelihood of strikes. The coefficient on social spending is positive and significant, but the significance disappears with clustered standard errors. This positive relationship is somewhat counterintuitive under the assumption that social spending would buy off potential strikers, but as it is minimal and disappears with the clustering, I attribute only limited causality to the relationship, questioning whether the correlation is actually motivated in the opposite direction. Finally, the relationship between strikes and manufacturing is positive and highly significant.
<table>
<thead>
<tr>
<th></th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gini</td>
<td>5.154</td>
<td>0.3014***</td>
<td>2.48*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini²</td>
<td>-48.92</td>
<td>-28.77***</td>
<td>-2.35*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuznets 10%</td>
<td>0.00707</td>
<td>0.09</td>
<td>1.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuznets 10%²</td>
<td>-0.0000015</td>
<td>-1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuznets 5%</td>
<td>0.0026</td>
<td>0.0381</td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuznets 5%²</td>
<td>-2.93E-08</td>
<td>-2.93E-08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>-0.16</td>
<td>-0.18</td>
<td>-0.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>0.00000306</td>
<td>0.0000306</td>
<td>1.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth (Light)</td>
<td>-0.0183</td>
<td>-0.0185</td>
<td>-0.0185</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manufacturing Sector</td>
<td>4.283</td>
<td>4.274</td>
<td>2.461***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manufacturing Sector²</td>
<td>-10.9</td>
<td>-10.9</td>
<td>-10.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>4098</td>
<td>4098</td>
<td>4098</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001
For my sensitivity analysis, I first test controls for national and state capitals, and regional fixed effects. The coefficient on the dummy for State Capital is positive, insignificant, and has no effect on the direction of any of the other coefficients. National Capital, a dummy that takes a one for each of Mexico City’s 16 municipios, is dropped due to multicollinearity. None of the eight regional dummies (Northwest, Northeast, North Central, South Central, West, East, Southeast and Southwest) is significant and neither do they change the sign on the coefficient.

Next, I run the same Zero Inflated Poisson regression from Regression 8, this time using Strike Density (strikes/population) instead of Total Strikes. Results are shown below (Table 5), and are fairly robust, although not as significant. The coefficient on Gini is positive, $Gini^2$ negative, and all the others equally consistent with the Total Strikes Model.

Finally, I run a negative binomial model on Total Strikes, with results shown below (Table 6). Once again, the results are confirmed, with the coefficients on Gini and $Gini^2$ significant at the .001 level and all signs correct.
### Table 5: Zero Inflated Poisson (Strike Density as Dependent Variable)

<table>
<thead>
<tr>
<th>Strike Density</th>
<th>Gini</th>
<th>(407.36)</th>
<th>[16.194]*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>39.3879</td>
<td>(407.36)</td>
<td>[16.194]*</td>
</tr>
<tr>
<td>Gini²</td>
<td>-38.433</td>
<td>(436.0734)*</td>
<td>[16.775]*</td>
</tr>
<tr>
<td>2005</td>
<td>-0.13396</td>
<td>(4.45)</td>
<td>[0.2098]</td>
</tr>
<tr>
<td>Population</td>
<td>0.000192</td>
<td>(0.00)</td>
<td>[.000028]***</td>
</tr>
<tr>
<td>Growth (Light)</td>
<td>0.02758</td>
<td>(2.28)</td>
<td>[0.05576]</td>
</tr>
<tr>
<td>Manufacturing Sector</td>
<td>6.397</td>
<td>(20.13)</td>
<td>[.88999]***</td>
</tr>
<tr>
<td>Constant</td>
<td>-19.893</td>
<td>(94.93)</td>
<td>[4.048]***</td>
</tr>
<tr>
<td>Manufacturing Sector</td>
<td>-0.0222</td>
<td>(1.79x10⁸)</td>
<td>[1.32]</td>
</tr>
<tr>
<td>N</td>
<td>4098</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 6: Negative Binomial Regression

<table>
<thead>
<tr>
<th>Total Strikes</th>
<th>Coef.</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gini</td>
<td>57.98844***</td>
<td>8.566237</td>
</tr>
<tr>
<td>Gini²</td>
<td>-57.79339***</td>
<td>9.130669</td>
</tr>
<tr>
<td>2005</td>
<td>-0.0935558</td>
<td>0.1035552</td>
</tr>
<tr>
<td>Population</td>
<td>0.0018959***</td>
<td>0.0001044</td>
</tr>
<tr>
<td>Growth (Light)</td>
<td>0.02758</td>
<td>(2.28)</td>
</tr>
<tr>
<td>Manufacturing Sector</td>
<td>6.397</td>
<td>(20.13)</td>
</tr>
<tr>
<td>Constant</td>
<td>-15.17751***</td>
<td>2.0483</td>
</tr>
</tbody>
</table>

### Table 5: Zero Inflated Poisson (Strike Density as Dependent Variable)

<table>
<thead>
<tr>
<th>Strike Density</th>
<th>Gini</th>
<th>(407.36)</th>
<th>[16.194]*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>39.3879</td>
<td>(407.36)</td>
<td>[16.194]*</td>
</tr>
<tr>
<td>Gini²</td>
<td>-38.433</td>
<td>(436.0734)*</td>
<td>[16.775]*</td>
</tr>
<tr>
<td>2005</td>
<td>-0.13396</td>
<td>(4.45)</td>
<td>[0.2098]</td>
</tr>
<tr>
<td>Population</td>
<td>0.000192</td>
<td>(0.00)</td>
<td>[.000028]***</td>
</tr>
<tr>
<td>Growth (Light)</td>
<td>0.02758</td>
<td>(2.28)</td>
<td>[0.05576]</td>
</tr>
<tr>
<td>Manufacturing Sector</td>
<td>6.397</td>
<td>(20.13)</td>
<td>[.88999]***</td>
</tr>
<tr>
<td>Constant</td>
<td>-19.893</td>
<td>(94.93)</td>
<td>[4.048]***</td>
</tr>
<tr>
<td>Manufacturing Sector</td>
<td>-0.0222</td>
<td>(1.79x10⁸)</td>
<td>[1.32]</td>
</tr>
<tr>
<td>N</td>
<td>4098</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 6: Negative Binomial Regression

<table>
<thead>
<tr>
<th>Total Strikes</th>
<th>Coef.</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gini</td>
<td>57.98844***</td>
<td>8.566237</td>
</tr>
<tr>
<td>Gini²</td>
<td>-57.79339***</td>
<td>9.130669</td>
</tr>
<tr>
<td>2005</td>
<td>-0.0935558</td>
<td>0.1035552</td>
</tr>
<tr>
<td>Population</td>
<td>0.0018959***</td>
<td>0.0001044</td>
</tr>
<tr>
<td>Growth (Light)</td>
<td>0.02758</td>
<td>(2.28)</td>
</tr>
<tr>
<td>Manufacturing Sector</td>
<td>6.397</td>
<td>(20.13)</td>
</tr>
<tr>
<td>Constant</td>
<td>-15.17751***</td>
<td>2.0483</td>
</tr>
</tbody>
</table>

N = 4098
In this paper, I have assessed the relationship between local inequality and strike levels in Mexican municipalities. I find a robust nonlinear relationship between inequality and strikes, suggesting strikes are most common at moderately high inequality and less so at extremes.

I suggest two possible explanations for this nonlinear relationship. The first is grounded in Panning’s theory of relative deprivation, which says that unrest happens most when people of different incomes compare themselves to each other, and that individuals are more likely to compare themselves to each other when their incomes are closer together. By this logic, societies with moderately high levels of inequality experience more strikes because those individuals with low incomes are more exposed to those at the higher end of the income strata, while those societies at the poles lack the exposure that would motivate them to unrest.

Second, it is possible that those living in the most unequal strata of society lack the capabilities necessary to strike, while those in moderately unequal societies can rally the necessary population to strike. While my research presents an empirical analysis of the relationship between strikes and inequality, it does not address those mechanisms that could potentially explain the results. More research is needed here.

The greatest contribution of my research is the incorporation of the concept of inequality in income distributions to the modern understanding of strikes. The Poisson and Zero Inflated Poisson models also more accurately represent the distribution of strike frequency than methods that have been used in the past. Finally, the use of light intensity from satellite imagery to proxy for economic development at the municipal level brings a relatively new approach to the historic problem of measuring GDP at the municipal level.

The most important conclusion of this research for business and policymakers is that within a sphere of moderation, more equal societies exhibit less industrial unrest. This is an important conclusion for supply chain production and for policymakers concerned with civil unrest. At the same time, societies in the process of reducing extreme inequality should know that if they are moving from the extremes of the spectrum toward a more equal society, they might actually see an increase in strikes before they see a decrease due to the nonlinear relationship. However, a continued reduction in inequality should eventually lead to more industrial stability.
10 Bibliography


James, Harvey, “Economic Development and Strikes: An Examination of the Haas and Stack Model” Hartford University Department of Economics, 1996.


“Oportunidades, A Human Development Program” http://www.oportunidades.gob.mx/Portal/wb/Web/oportunidades_a_human_development_program


10 Bibliography


James, Harvey, “Economic Development and Strikes: An Examination of the Haas and Stack Model” Hartford University Department of Economics, 1996.


“Oportunidades, A Human Development Program” http://www.oportunidades.gob.mx/Portal/wb/Web/oportunidades_a_human_development_program


Rowe, Nicholas, “Mexico’s Oportunidades: Conditional Cash Transfers as the Solution to Global Poverty?”
https://www.cmc.edu/keck/student/RoweN%20Fellowship%20REP%202010-11.pdf

Sistema Estatal y Municipal de Base de Datos.
http://sc.inegi.org.mx/sistemas/cobdem/.

“Stata Data Analysis Examples: Negative Binomial Regression,” UCLA: Statistical Consulting Group,

DO U.S. BORDER ENFORCEMENT OPERATIONS INCREASE HUMAN SMUGGLING FEES ALONG THE U.S.-MEXICO BORDER?

KYLE COOMBS

Macalester College

ABSTRACT

Undocumented migrants frequently hire border crossing experts, called “coyotes” to facilitate a successful, safer crossing. U.S. border enforcement actively counters these migrants. U.S. measures of enforcement and coyote fees grew together during the 20th century, suggesting a connection between enforcement and the coyote market. This paper tests the effect of border patrol agents and operations on coyote fees using a dataset compiled from the Mexican Migration Project, U.S. Customs & Border Protection, the Department of Homeland Security, and the United States Sentencing Commission. I do not find a significant connection between coyote fees and border enforcement, but do show that average prison time along the border acted as a shifter of supply prior to 2005.

Undocumented migrants frequently hire border crossing experts, called “coyotes” to facilitate a successful, safer crossing. U.S. border enforcement actively counters these migrants. U.S. measures of enforcement and coyote fees grew together during the 20th century, suggesting a connection between enforcement and the coyote market. This paper tests the effect of border patrol agents and operations on coyote fees using a dataset compiled from the Mexican Migration Project, U.S. Customs & Border Protection, the Department of Homeland Security, and the United States Sentencing Commission. I do not find a significant connection between coyote fees and border enforcement, but do show that average prison time along the border acted as a shifter of supply prior to 2005.
Cointegration is a statistical phenomenon where two time series move together through time. By applying the Granger Representation Theorem, the error correction process that keeps the series cointegrated can be observed. An example can be found in the mortgage and treasury markets. Prior to the Federal Reserve intervening with unconventional monetary policy during the financial crisis of 2008 and 2009, the 7-year Treasury rate and 30-year mortgage rate were not cointegrated. Conversely, in the post-intervention period, the rates became cointegrated again, indicating order had returned to investing. Interestingly, from 1971-2013, the 7-year Treasury is the price leader, but during QE1, price discovery shifted to the 30-year mortgage rate.
error correction process where treasuries are the price leader (1996). The result is very intuitive; since treasuries are considered riskless, it makes sense for mortgage rates to take their initial pricing cues from treasuries. This paper uses the 7-year treasury rate, as opposed to the 30-year rate used by Rahman et al. The use of the 7-year rate will be addressed in the Hypothesis and Model section below. Mehra (1994) concludes that long-term bond rates are positively correlated with the deficit and inflation (1994).

In 2008 and 2009, the United States was in a recession brought about by excessive lending in the subprime mortgage market. The financial crisis caused the Federal Reserve to step in and use unconventional monetary policy in an attempt to bring normalcy back to financial markets. The policies used by the Fed included the usual manipulation of the federal funds rate and several rounds of large-scale asset purchases, also known as quantitative easing (QE). Additionally, the Federal Reserve introduced Operation Twist in which they sold short-term treasuries and bought long-term treasuries in an attempt to further flatten the yield curve.

The three rounds of quantitative easing (also known as QE1, QE2, and QE3), in addition to Operation Twist, expanded the Federal Reserve’s Balance Sheet from around $800 million to roughly $4 trillion. In Figure 2, it can be observed that during QE1 the Fed bought substantially more MBSs (packages of mortgages with roughly the same risk level) than treasuries, which can be seen in Balance Sheet below. As quantitative easing progressed, the Fed bought more treasuries than MBSs, only further increasing the size of their Balance Sheet.

3 Hypothesis

The Engle and Granger two-step procedure requires two time series. In this paper the 30-Year Conventional Mortgage Rate and the 7-Year Treasury Constant Maturity Rate are used. The 7-year treasury was chosen because the duration of a 30-year mortgage is roughly seven years. Additionally, treasury rates have been found to be cointegrated, and since the 10-year and 5-year treasuries are generally seen as benchmark rates, the 7-year is a logical medium.

Both time series have monthly and weekly data available. The weekly data runs from January 8, 2004 to January 9, 2014; however, the data was partitioned to finish with the last observation of December 2013. The monthly data begins at April 1, 1971 and runs through December 2013. The
end dates were chosen so the monthly and weekly data would finish simultaneously.

This part of the paper intends to find the effect of Federal Reserve monetary policy on the relationship between mortgage and treasury rates. This relationship includes whether or not the two series are cointegrated and price leadership (shown through error correction).

### 4 Methodology and Model

If two series are cointegrated, then the two series move together through time (for example, see Figure 1). The method by which the series stay together is interesting. The error correction process displays how two time series can move together without significant separation in the long run. Therefore, the error correction process offers a clear look into the price discovery between two or more series.

![Figure 1](image)

Two time series are cointegrated if the combination of two non-stationary, I(1) series results in a stationary, I(0) series. A non-stationary series is one where a “random walk” or “unit root” exists and spurious results are likely to be found because the errors of the two series may be related. The first step is...
to identify whether each series is I(1). To find out whether or not the series are I(1), an Augmented Dickey-Fuller (ADF) test is used. The test was developed by Dickey and Fuller (1979) and “allows for more dynamics in the DF regression and, consequently, is over-parameterized in the first order case but correctly specified in the higher order cases” (Engle and Granger, 1987). A “first order case” means that to become non-stationary, both sides of the equation must be differenced one time. Only data that is stationary is useful in forecasting time series. Stationary data does not have a random walk and the errors on two time series are not related.

The following basic “autoregressive” AR(1) model is used to derive the ADF test:

\[ y_t = \delta + \varphi y_{t-1} + \epsilon_t, \text{ where } \epsilon_t \sim N(0, \sigma^2) \]  

(1)

Continuing to the ADF test, \( y_{t-1} \) is subtracted from both sides of the above equation. The resulting equation is a first-differenced OLS model, including all statistically significant lags of \( \Delta y_t \).

\[ y_t = \delta + \varphi y_{t-1} + \epsilon_t \]  

(2)

\[ \Delta y_t = \delta + (\varphi - 1) y_{t-1} + \lambda_1 \Delta y_{t-2} + \lambda_2 \Delta y_{t-3} + \epsilon_t \]  

(3)

Also represented as

\[ \Delta y_t = \delta + Y y_{t-1} + \lambda_1 \Delta y_{t-2} + \lambda_2 \Delta y_{t-3} + \epsilon_t, \]  

\[ \text{where } Y = (\varphi - 1) \]  

(4)

The objective is to test whether \( Y \) is equal to 0, which means \( \varphi \) equals 1 and \( y_t \) is a non-stationary I(1) autoregressive variable. The \( t \)-statistic generated by the OLS output is used to test the null hypothesis that \( Y = 0 \). If \( |t| > 2.89 \), then the null hypotheses is rejected at the 95% confidence level and \( y_t \) is I(1). The critical value used in the ADF test is from Dickey and Fuller (1979). This process is then repeated on the second time series and if the second time series is found to be I(1) as well, then the following “cointegrating regression” is estimated:

\[ y_t = \beta_0 + \beta_1 x_t + \epsilon_t \]  

(5)
Calculate the errors \( (e_t) \) from the cointegrating regression and estimate another ADF test using the model below, which is derived in the same fashion as the \( \Delta y_t \) model from above:
\[
\Delta e_t = \gamma_t e_{t-1} + \Delta e_{t-1} + \Delta e_{t-2} + \ldots + \Delta e_{t-n} + u_t
\]  
(6)

Again, the null hypothesis is \( \gamma_t = 0 \). If |\( t_c \)| > |\(-3.17\)|, then reject the hypothesis of non-stationarity and conclude that \( e_t \) is I(0), i.e., stationary and the two time series are cointegrated. The critical value of -3.17 to test for cointegration is from Engle and Granger (1987).

The error correction process explains why the two series stay together and do not separate over long periods of time. To discover the nature of the error correcting processes, let \( z_t = e_t \), from the OLS of \( y_t \), and include all statistically significant lags of \( \Delta y_t \) and \( \Delta x_t \). With one lag of each, the model is:
\[
\Delta y_t = \alpha_0 + \delta_1 z_{t-1} + \alpha_1 \Delta y_{t-1} + \alpha_2 \Delta x_{t-1} + u_t
\]  
(7)
\[
\Delta x_t = \beta_0 + \delta_2 z_{t-1} + \beta_1 \Delta y_{t-1} + \beta_2 \Delta x_{t-1} + u_t
\]  
(8)

The coefficients \( \delta_1 \) and \( \delta_2 \) show which series drives the other, or whether the series may be a two-way error correction process. As long as the coefficients are significant, they have predictive power. Examples of the coefficient results and interpretation are found in Figure 2. The insignificant coefficient is the price leader, and the significant coefficient is the price follower because the error correction adjustments are made by the variable with the significant coefficient on the error term.

**Figure 2**

<table>
<thead>
<tr>
<th>( \delta_1 ) (in ( y_t ))</th>
<th>( \delta_2 ) (in ( x_t ))</th>
<th>Correction Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>= 0</td>
<td>&gt; 0</td>
<td>( y_t ) is the price leader</td>
</tr>
</tbody>
</table>

Calculate the errors \( (e_t) \) from the cointegrating regression and estimate another ADF test using the model below, which is derived in the same fashion as the \( \Delta y_t \) model from above:
\[
\Delta e_t = \gamma_t e_{t-1} + \Delta e_{t-1} + \Delta e_{t-2} + \ldots + \Delta e_{t-n} + u_t
\]  
(6)

Again, the null hypothesis is \( \gamma_t = 0 \). If |\( t_c \)| > |\(-3.17\)|, then reject the hypothesis of non-stationarity and conclude that \( e_t \) is I(0), i.e., stationary and the two time series are cointegrated. The critical value of -3.17 to test for cointegration is from Engle and Granger (1987).

The error correction process explains why the two series stay together and do not separate over long periods of time. To discover the nature of the error correcting processes, let \( z_t = e_t \), from the OLS of \( y_t \), and include all statistically significant lags of \( \Delta y_t \) and \( \Delta x_t \). With one lag of each, the model is:
\[
\Delta y_t = \alpha_0 + \delta_1 z_{t-1} + \alpha_1 \Delta y_{t-1} + \alpha_2 \Delta x_{t-1} + u_t
\]  
(7)
\[
\Delta x_t = \beta_0 + \delta_2 z_{t-1} + \beta_1 \Delta y_{t-1} + \beta_2 \Delta x_{t-1} + u_t
\]  
(8)

The coefficients \( \delta_1 \) and \( \delta_2 \) show which series drives the other, or whether the series may be a two-way error correction process. As long as the coefficients are significant, they have predictive power. Examples of the coefficient results and interpretation are found in Figure 2. The insignificant coefficient is the price leader, and the significant coefficient is the price follower because the error correction adjustments are made by the variable with the significant coefficient on the error term.

**Figure 2**

<table>
<thead>
<tr>
<th>( \delta_1 ) (in ( y_t ))</th>
<th>( \delta_2 ) (in ( x_t ))</th>
<th>Correction Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>= 0</td>
<td>&gt; 0</td>
<td>( y_t ) is the price leader</td>
</tr>
</tbody>
</table>
5 Results

5.1 1971-2013: Monthly data

The two series were found to be I(1) individually, and cointegrated (Table 1) at the .01 level. This conclusion follows from the e_t being I(0). Therefore, the error correction models were estimated. The error correction term was negative and insignificant for the 7-year Treasury and positive and significant at the .01 level for the 30-year mortgage rate (Table 2 and Table 3, respectively). Therefore, there is a one-way error-correction process where 7-year Treasuries are the price leader, as expected. It makes sense that mortgage rates follow Treasury rates, so over a forty-three year time window, it is logical for the mortgage rate to correct back into cointegration equilibrium with the Treasury rate.


The results become more interesting in the years leading up to the financial crisis and during the crisis itself. Using the monthly data, cointegration was tested from January 1998 to the first Federal Reserve intervention in the form of QE1. In this time window, cointegration disappears (Table 1, p-value=0.44).

A potential explanation for the lack of cointegration can be found in the spread between the 30-year mortgage rate and the 7-year Treasury rate as seen in Figure 1. From 2006 to early-mid 2007, the spread between the two rates is relatively flat indicating they were moving together with very little variation. However, beginning in mid-2007, the spread widens as the two rates begin to separate. The peaks can be identified as (1) March, 16 2008 with Bear Stearns merging with JP Morgan Chase, after not having enough short term funds to remain viable, (2) Lehman Brothers filing for Chapter 11 bankruptcy on September 15, 2008, and (3) the Federal Reserve entering the market with QE1 on December 16, 2008. Following the introduction of
QE1, the spread closes and evens out again within six or seven months. A potential explanation for the widening of the spread is the sharp increase in delinquency rates in early to mid-2007. The subprime mortgage bubble becomes apparent over this time period as an increasing number of mortgage owners default on their payments.

As shown in the years 1971-2013, mortgage lenders watched Treasury rates as they attempted to set mortgage rates. However, with increasing delinquency rates, mortgage lenders could no longer rely on Treasury rates as heavily, and a seemingly chaotic period followed in 2007 and 2008 where mortgage rates remained around six percent even as 7-year Treasury rates decreased considerably. This indicates that mortgage lenders and MBS investors were not taking cues from Treasury rates.

5.3 2004 to Federal Reserve Intervention (Dec 2008): Weekly data
The weekly data has the same results as the monthly data. The two series are not cointegrated and have therefore, have no error-correction process (Table 1, p-value=0.60).

5.4 Federal Reserve begins QE1 through Dec 2013: Weekly data
Interestingly, after the Federal Reserve began the first round of quantitative easing, the two series again became cointegrated at the .01 level. As the Fed bought MBSs from Fannie Mae and Freddie Mac, the spread between the two rates decreased. During the original QE1 and the expansion several months later, the Fed bought $1.25 trillion MBSs and $400 billion Treasuries. The initial purchase program and the subsequent programs over the next few years brought the two series back into cointegration.

Because cointegration exists, there must be an error-correction process. Over the time period starting at the beginning of QE1 and finishing at the end of December 2013, the 7-year Treasury rate is the price leader over the 30-year mortgage rate. The coefficient on the error correction term for the Treasury rate is statistically insignificant (Table 2, p-value=0.24). The coefficient on the error correction term in the mortgage equation is significant at the .05 level (Table 3). As stated before, the variable with the stronger t-statistic corrects to the movements of the variable with the weaker t-statistic. Thus, the 7-year Treasury is the price leader.

The conclusion to be drawn from this time window is that the Federal Open Market Committee identified that mortgage and Treasury markets were no longer cointegrated and that the 7-year Treasury rate is the price leader over the 30-year mortgage rate. Because cointegration exists, there must be an error-correction process. Over the time period starting at the beginning of QE1 and finishing at the end of December 2013, the 7-year Treasury rate is the price leader over the 30-year mortgage rate. The coefficient on the error correction term for the Treasury rate is statistically insignificant (Table 2, p-value=0.24). The coefficient on the error correction term in the mortgage equation is significant at the .05 level (Table 3). As stated before, the variable with the stronger t-statistic corrects to the movements of the variable with the weaker t-statistic. Thus, the 7-year Treasury is the price leader.

The conclusion to be drawn from this time window is that the Federal Open Market Committee identified that mortgage and Treasury markets were no longer cointegrated and that the 7-year Treasury rate is the price leader over the 30-year mortgage rate.
longer consistently related and acted to bring the two markets back to cointegration equilibrium.

5.5 Federal Reserve begins QE1 through Dec 2013: Monthly data

The monthly data from the beginning of QE1 to the end of 2013 is extremely similar to the weekly data. The two series are both I(1) and, when combined, are I(0), i.e., cointegrated. Additionally, there is a one way error-correction process where the mortgage rate re-adjusts back to the cointegration equilibrium with the 7-year Treasury rate (Table 3 and Table 2, respectively).

5.6 QE1: Weekly data

Within the Fed intervention-to-present time window, two separate time periods were examined, the first of which was QE1. The first round of quantitative easing took place from December 16, 2008 to the end of March 2010. Additionally, on March 18, 2009 there was an expansion of QE1. During the expansion, the Federal Reserve bought an additional $750 billion of MBSs and $300 billion in longer-term Treasuries.

During QE1, as in the entire Federal Reserve intervention period, the two time series display a cointegration relationship. The series are cointegrated at roughly the .20 level (Table 1). Interestingly, QE1 is the only time period examined when the mortgage rate was the price leader instead of the Treasury rate. The coefficient on the error correction term in the 7-year Treasury equation is statistically significant at the .054 level (Table 2). The coefficient on the error correction term in the 30-year mortgage is insignificant (Table 3).

The change in price leadership may be due to the increase in the volume of MBSs the Federal Reserve was buying compared to Treasuries. Although the Treasury market is very deep, the large influx of funds into the MBS market may have altered the structural relationship between the two rates. Trading volumes for the two series were not available, so no further analysis was conducted. A second potential explanation for the shift in price leadership is that banks were positioning themselves to unload their risky MBSs, thereby, focusing on mortgage debt rather than other rates. The significance of the shift in price leadership is that in all other years, mortgage

---

1 This explanation for mortgages driving Treasuries is owed to Professor Thomas Simpson.
investors responded to changes in Treasury rates. However, during QE1, mortgage investors were not responding to Treasury rates; in fact, Treasury investors were responding to mortgage rates.

5.7 QE3: Weekly data

QE2 and Operation Twist were not analyzed due to insufficient degrees of freedom. Therefore, QE3 is the final time window dissected. The two series are I(1) during this period and when combined are stationary and thus, cointegrated. The series follow QE1 where the series are cointegrated at roughly the .2 level. However, the error correction process is more ambiguous than in QE1. The 7-year Treasury’s coefficient is statistically insignificant at the .274 level (Table 2) and the 30-year mortgage rate coefficient is insignificant (Table 3, p-value=0.25). Thus, nothing can be concluded from these results. The ambiguity can be attributed to either having only 64 observations or that weakly cointegrated time series may produce weak error correction results.

6 Conclusion

The following periods are cointegrated to at least the .2 level:
- 1971-2013: Monthly data
- Fed Intervention-2013: Monthly weekly data
- Fed Intervention-2013: Weekly data
- QE1: Weekly data
- QE3: Weekly data

The following time periods are not cointegrated:
- 1998-Fed Intervention: Monthly data
- 2004-Fed Intervention: Weekly data

In periods of cointegration, the 7-year Treasury rate was the price leader over the 30-year mortgage rate, except during QE1 and QE3. During QE1, the 30-year mortgage rate was the price leader. In QE3, it is not possible to make a price leadership determination. The change in leadership during QE1 may be due to the high ratio of MBSs to Treasuries the Fed was buying during the period, or to banks positioning themselves to unload risky mortgage debt.

In the time windows of non-cointegration, an increase in the delinquency rate seems to have initiated the widening spread between 30-year mortgage rates and 7-year Treasury rates. The increasing delinquency rate and the subprime mortgage crisis led to the falls of Bear Stearns and Lehman Brothers.

investors responded to changes in Treasury rates. However, during QE1, mortgage investors were not responding to Treasury rates; in fact, Treasury investors were responding to mortgage rates.

5.7 QE3: Weekly data

QE2 and Operation Twist were not analyzed due to insufficient degrees of freedom. Therefore, QE3 is the final time window dissected. The two series are I(1) during this period and when combined are stationary and thus, cointegrated. The series follow QE1 where the series are cointegrated at roughly the .2 level. However, the error correction process is more ambiguous than in QE1. The 7-year Treasury’s coefficient is statistically insignificant at the .274 level (Table 2) and the 30-year mortgage rate coefficient is insignificant (Table 3, p-value=0.25). Thus, nothing can be concluded from these results. The ambiguity can be attributed to either having only 64 observations or that weakly cointegrated time series may produce weak error correction results.

6 Conclusion

The following periods are cointegrated to at least the .2 level:
- 1971-2013: Monthly data
- Fed Intervention-2013: Monthly weekly data
- Fed Intervention-2013: Weekly data
- QE1: Weekly data
- QE3: Weekly data

The following time periods are not cointegrated:
- 1998-Fed Intervention: Monthly data
- 2004-Fed Intervention: Weekly data

In periods of cointegration, the 7-year Treasury rate was the price leader over the 30-year mortgage rate, except during QE1 and QE3. During QE1, the 30-year mortgage rate was the price leader. In QE3, it is not possible to make a price leadership determination. The change in leadership during QE1 may be due to the high ratio of MBSs to Treasuries the Fed was buying during the period, or to banks positioning themselves to unload risky mortgage debt.

In the time windows of non-cointegration, an increase in the delinquency rate seems to have initiated the widening spread between 30-year mortgage rates and 7-year Treasury rates. The increasing delinquency rate and the subprime mortgage crisis led to the falls of Bear Stearns and Lehman Brothers.
When the Federal Reserve entered the market with the large-scale asset purchase program, the spread between the two rates decreased and cointegration reappeared. These findings indicate that the Federal Reserve brought the mortgage and Treasury markets back from being unrelated to a more orderly structure that allowed for rational investing. Also, because the mortgage rate became the price leader instead of the 7-year Treasury during QE1; the Federal Reserve demonstrated its ability to alter the structure of financial markets.
Bibliography


### Table 1: ADF on Errors

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Variable</th>
<th>Coeff.</th>
<th>SE</th>
<th>T</th>
<th>df</th>
<th>obs</th>
<th>R-sq</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971-2013-M</td>
<td>e(t-1)</td>
<td>-0.105</td>
<td>0.022</td>
<td>-4.868***</td>
<td>507</td>
<td>507</td>
<td>0.167</td>
<td>0.248</td>
</tr>
<tr>
<td></td>
<td>de(t-1)</td>
<td>0.164</td>
<td>0.042</td>
<td>3.928***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>de(t-2)</td>
<td>-0.277</td>
<td>0.044</td>
<td>-6.486***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>de(t-5)</td>
<td>0.083</td>
<td>0.044</td>
<td>2.841**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998-Fed-M</td>
<td>e(t-1)</td>
<td>-0.026</td>
<td>0.034</td>
<td>-0.771</td>
<td>126</td>
<td>126</td>
<td>0.062</td>
<td>0.127</td>
</tr>
<tr>
<td></td>
<td>de(t-1)</td>
<td>0.179</td>
<td>0.091</td>
<td>1.966**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>de(t-2)</td>
<td>-0.183</td>
<td>0.093</td>
<td>-1.979**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004-Fed-W</td>
<td>e(t-1)</td>
<td>0.006</td>
<td>0.011</td>
<td>0.533</td>
<td>252</td>
<td>252</td>
<td>0.072</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td>de(t-1)</td>
<td>-0.133</td>
<td>0.063</td>
<td>-2.090**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>de(t-2)</td>
<td>0.141</td>
<td>0.063</td>
<td>2.224**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>de(t-5)</td>
<td>-0.168</td>
<td>0.063</td>
<td>-2.669***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fed-2013-M</td>
<td>e(t-1)</td>
<td>-0.223</td>
<td>0.073</td>
<td>-3.032**</td>
<td>54</td>
<td>54</td>
<td>0.198</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>de(t-1)</td>
<td>0.243</td>
<td>0.132</td>
<td>1.845*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>de(t-3)</td>
<td>0.290</td>
<td>0.116</td>
<td>2.494**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fed-2013-W</td>
<td>e(t-1)</td>
<td>-0.082</td>
<td>0.018</td>
<td>-4.482***</td>
<td>257</td>
<td>257</td>
<td>0.141</td>
<td>0.067</td>
</tr>
<tr>
<td></td>
<td>de(t-1)</td>
<td>-0.144</td>
<td>0.060</td>
<td>-2.416**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>de(t-3)</td>
<td>0.146</td>
<td>0.056</td>
<td>2.598**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>de(t-4)</td>
<td>0.157</td>
<td>0.057</td>
<td>2.764***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QE1-W</td>
<td>e(t-1)</td>
<td>-0.088</td>
<td>0.033</td>
<td>-2.680</td>
<td>61</td>
<td>61</td>
<td>0.107</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td>de(t-1)</td>
<td>-0.222</td>
<td>0.084</td>
<td>-2.634</td>
<td>64</td>
<td>64</td>
<td>0.099</td>
<td>0.055</td>
</tr>
<tr>
<td>QE3-W</td>
<td>e(t-1)</td>
<td>-0.088</td>
<td>0.033</td>
<td>-2.680</td>
<td>61</td>
<td>61</td>
<td>0.107</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td>de(t-1)</td>
<td>-0.222</td>
<td>0.084</td>
<td>-2.634</td>
<td>64</td>
<td>64</td>
<td>0.099</td>
<td>0.055</td>
</tr>
<tr>
<td>Time Period</td>
<td>Variable</td>
<td>Coef</td>
<td>SE</td>
<td>T</td>
<td>df</td>
<td>obs</td>
<td>R-sq</td>
<td>SE</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>1971-2013-M</td>
<td>z(t-1)</td>
<td>-0.007</td>
<td>0.029</td>
<td>-0.248</td>
<td>507</td>
<td>508</td>
<td>0.192</td>
<td>0.316</td>
</tr>
<tr>
<td></td>
<td>dy(t-1)</td>
<td>0.606</td>
<td>0.061</td>
<td>9.859***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dy(t-3)</td>
<td>0.133</td>
<td>0.050</td>
<td>2.634***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dy(t-4)</td>
<td>-0.120</td>
<td>0.051</td>
<td>-2.363**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dy(t-5)</td>
<td>0.192</td>
<td>0.061</td>
<td>3.153***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dx(t-1)</td>
<td>-0.479</td>
<td>0.067</td>
<td>-7.121***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dx(t-5)</td>
<td>-0.166</td>
<td>0.065</td>
<td>-2.558**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998-Fed-M</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2004-Fed-W</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Fed-2013-M</td>
<td>z(t-1)</td>
<td>-0.053</td>
<td>0.172</td>
<td>-0.310</td>
<td>53</td>
<td>54</td>
<td>0.036</td>
<td>0.214</td>
</tr>
<tr>
<td></td>
<td>dy(t-1)</td>
<td>0.186</td>
<td>0.136</td>
<td>1.373</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fed-2013-W</td>
<td>z(t-1)</td>
<td>-0.035</td>
<td>0.029</td>
<td>-1.181</td>
<td>256</td>
<td>257</td>
<td>0.036</td>
<td>0.107</td>
</tr>
<tr>
<td></td>
<td>dy(t-1)</td>
<td>0.177</td>
<td>0.062</td>
<td>2.875***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QE1-W</td>
<td>z(t-1)</td>
<td>-0.093</td>
<td>0.047</td>
<td>-1.964**</td>
<td>60</td>
<td>61</td>
<td>0.061</td>
<td>0.125</td>
</tr>
<tr>
<td>QE3-W</td>
<td>z(t-1)</td>
<td>-0.037</td>
<td>0.136</td>
<td>-0.274</td>
<td>63</td>
<td>64</td>
<td>0.004</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td>dy(t-1)</td>
<td>0.048</td>
<td>0.128</td>
<td>0.377</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Table 3: Error Correction for 30-Year Mortgage Rates**

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Coef</th>
<th>SE</th>
<th>T</th>
<th>df</th>
<th>obs</th>
<th>R-sq</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971-2013-M</td>
<td>z(t-1)</td>
<td>0.094</td>
<td>0.026</td>
<td>3.632***</td>
<td>507</td>
<td>508</td>
<td>0.320</td>
</tr>
<tr>
<td></td>
<td>dx(t-2)</td>
<td>-0.107</td>
<td>0.044</td>
<td>-2.403**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dy(t-1)</td>
<td>0.508</td>
<td>0.043</td>
<td>11.921***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998-Fed-M</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2004-Fed-W</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Fed-2013-M</td>
<td>z(t-1)</td>
<td>0.214</td>
<td>0.123</td>
<td>1.742*</td>
<td>53</td>
<td>54</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>dy(t-4)</td>
<td>-0.173</td>
<td>0.096</td>
<td>-1.804*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fed-2013-W</td>
<td>z(t-1)</td>
<td>0.048</td>
<td>0.023</td>
<td>2.068**</td>
<td>256</td>
<td>257</td>
<td>0.102</td>
</tr>
<tr>
<td></td>
<td>dx(t-1)</td>
<td>-0.360</td>
<td>0.094</td>
<td>-3.843***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dy(t-1)</td>
<td>0.374</td>
<td>0.071</td>
<td>4.875***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QE1-W</td>
<td>z(t-1)</td>
<td>0.022</td>
<td>0.038</td>
<td>0.590</td>
<td>60</td>
<td>61</td>
<td>0.006</td>
</tr>
<tr>
<td>QE3-W</td>
<td>z(t-1)</td>
<td>0.192</td>
<td>0.165</td>
<td>1.160</td>
<td>63</td>
<td>64</td>
<td>0.021</td>
</tr>
</tbody>
</table>

**Key**
- **Fed** = Federal Reserve Intervention
- **Monthly** = Monthly data
- **Weekly** = Weekly data
- e(t-1) = errors lagged 1 period for ADF on errors
- e(t-n) = errors lagged n periods
- dx(t-n) = change (delta) errors lagged n periods
- z(t-1) = errors lagged 1 period for error correction model
- dx(t-n) = change in mortgage rate lagged n periods
- dy(t-n) = change in Treasury rate lagged n periods

**Table 3: Error Correction for 30-Year Mortgage Rates (x)**

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Coef</th>
<th>SE</th>
<th>T</th>
<th>df</th>
<th>obs</th>
<th>R-sq</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971-2013-M</td>
<td>z(t-1)</td>
<td>0.094</td>
<td>0.026</td>
<td>3.632***</td>
<td>507</td>
<td>508</td>
<td>0.320</td>
</tr>
<tr>
<td></td>
<td>dx(t-2)</td>
<td>-0.107</td>
<td>0.044</td>
<td>-2.403**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dy(t-1)</td>
<td>0.508</td>
<td>0.043</td>
<td>11.921***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998-Fed-M</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>2004-Fed-W</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fed-2013-M</td>
<td>z(t-1)</td>
<td>0.214</td>
<td>0.123</td>
<td>1.742*</td>
<td>53</td>
<td>54</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>dy(t-4)</td>
<td>-0.173</td>
<td>0.096</td>
<td>-1.804*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fed-2013-W</td>
<td>z(t-1)</td>
<td>0.048</td>
<td>0.023</td>
<td>2.068**</td>
<td>256</td>
<td>257</td>
<td>0.102</td>
</tr>
<tr>
<td></td>
<td>dx(t-1)</td>
<td>-0.360</td>
<td>0.094</td>
<td>-3.843***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dy(t-1)</td>
<td>0.374</td>
<td>0.071</td>
<td>4.875***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QE1-W</td>
<td>z(t-1)</td>
<td>0.022</td>
<td>0.038</td>
<td>0.590</td>
<td>60</td>
<td>61</td>
<td>0.006</td>
</tr>
<tr>
<td>QE3-W</td>
<td>z(t-1)</td>
<td>0.192</td>
<td>0.165</td>
<td>1.160</td>
<td>63</td>
<td>64</td>
<td>0.021</td>
</tr>
</tbody>
</table>

**Key**
- **Fed** = Federal Reserve Intervention
- **Monthly** = Monthly data
- **Weekly** = Weekly data
- e(t-1) = errors lagged 1 period for ADF on errors
- e(t-n) = errors lagged n periods
- dx(t-n) = change (delta) errors lagged n periods
- z(t-1) = errors lagged 1 period for error correction model
- dx(t-n) = change in mortgage rate lagged n periods
- dy(t-n) = change in Treasury rate lagged n periods
HETEROGENEITY IN CONDITIONAL CASH TRANSFER EFFECTS: EVIDENCE FROM A RANDOMIZED TRIAL IN MALAWI

Johnny Huynh
Pomona College

“A growing body of literature documents the effectiveness of conditional cash transfers (CCTs) on increasing investments in children's human capital. While CCTs increase mean rates of school enrollment and test scores, how are these gains distributed across treated households? This paper constructs a labor supply model that predicts the impact of a CCT is concentrated among a subset of participants. To test our theory, we utilize data from a randomized controlled trial in Malawi from 2007 to 2009. We use a set of observable test scores to identify latent academic ability, and estimate quantile treatment effects for the CCT program. Our evidence shows that the impact of the CCT on test scores is only significant at the median of the latent academic ability distribution. Whereas the CCT increases test scores, on average, by 0.13 standard deviation, the median treatment effect is largest with 0.23 standard deviation. Unconditional transfers do not show a similar distributional heterogeneity. Nonparametric estimates of school enrollment confirm these results. Our findings suggest the primary beneficiaries of CCTs are higher-ability children, while having little to no effect on lower-ability, often poorer, children.”
MOTIVATION AND WORK PERFORMANCE: THE EFFECTS OF PROSOCIAL AND INTRINSIC MOTIVATION ON TEACHER PERFORMANCE IN RURAL CHINA

KATHERINE LOOSLEY

The London School of Economics

ABSTRACT

This study examines the effects of motivation on work performance. Two contrasting types of motivation are evaluated, prosocial motivation, that is, motivation that comes from the desire to help others, and intrinsic motivation, motivation from enjoyment in the job itself. Work performance is studied in the context of education in rural China using data from a 2012 randomized controlled trial. Findings indicate that prosocial motivation of teachers has a significant but relatively small positive effect on academic performance, while intrinsic motivation, conditional on prosocial motivation, may actually have a negative effect on student scores. The positive effect of prosocial motivation occurs primarily among students with high ability, indicating that prosocially motivated teachers may focus their efforts on top-achieving students who have a chance to continue their education. Finally, by studying the interaction of motivation and teacher incentives, it appears that monetary incentives may strengthen the effect of prosocial motivation on student performance. The findings illustrate that the type of motivation may matter for work performance and that contrary to the claim that monetary incentives crowd out motivation, incentives may serve to increase motivation and performance.
INDUSTRY AND FINANCE IN INTERWAR INDIA

RUSSELL BLACK

Oxford University

ABSTRACT

I use a panel dataset to test the effect of exogenous commodity price shocks on agriculture, banks and industry in India in the interwar period. The Great Depression limits India’s ability to export primary products. However, the Indian financial system reacts to the loss of export earnings by redirecting capital to industrial investment, now relatively more profitable.

1 Introduction

At the beginning of the 1920s, India was a peripheral economy supplying primary goods to European markets, and in return importing finished products. However, this economic structure was fragile, and vulnerable to adverse commodity price shocks starting from the mid-1920s and intensifying with the Great Depression. How did India adapt to the changing global economy during the interwar period? Did the Great Depression shock India out of its underdevelopment trap? To answer these questions, I create a panel dataset on district and city level agricultural capacity, bank balance sheets and industrial plants. I use this data to show the effects of falling commodity prices on the structure of the Indian economy. I find that a lack of demand for exports of primary products induces capital owners to invest in industry for domestic consumption.

For the period 1921-1938, I have collected the balance sheets of the 595 largest banks in India and Burma, and the spread of each bank across 900 cities and 250 administrative districts. In addition, I have collected data on the number of industrial firms employing more than 100 employees, and the total acreage of 30 crops, for each of the 250 districts. There is significant variance in specialisation between regions in India, which allows for the

2 Tomlinson, op Cit, p44
effects of commodity price shocks on districts in my data to be identified, using a difference-in-difference approach. I calculate price shocks by using world prices to value a district’s potential output of primary products. Using panel regressions, I first test the effects of a change in the price shock on the financial sector of each district. Second, I test whether an increase in available credit causes an increase in the number of firms in each district. There are more large firms in districts with more available funds to loan, even when controlling for other financial coverage, such as deposits.

The Great Depression imposed import substitution on India, whether India willed it or not. In the course of the 1920s and 30s, merchants and financiers that relied on the export trade in primary products became less profitable. There was less incentive to engage in economic activity that ‘trapped’ the Indian economy into being a peripheral exporter of raw materials. My results show that, faced with a hostile world economy, these financier groups changed their business practices and redeployed their capital into industry, starting India on a path out of underdevelopment.

2 Background

The idea that the terms of trade between a peripheral and a developed economy might hinder the development of the peripheral economy has been greatly influential within Indian economic thought. It was a motivating force in the emergence of Indian nationalism in the nineteenth century, most famously in the writings of Romesh Dutt. The first nationalist campaign was the _swadeshi_ movement, the ‘homemade’ movement, which called for the boycott of foreign companies and foreign-made manufactures. Following on from an era of famine in the late nineteenth century, economic imperialism – in the form of adverse terms of trade – was felt by the Indian nationalist movement as keenly as political imperialism. Underdevelopment was an important political topic, discussed extensively by both Indian and British intellectuals and policymakers.

According to the ‘Lewis Thesis’, developed economies specialise in relatively high productivity manufacturing, forcing peripheral economies to specialise in agriculture and reduce the size of their less competitive manufacturing sector. Underdeveloped economies remain underdeveloped because they are trapped exporting agricultural products to colonial

---

3 R. Dutt, _India in the Victorian Age: an Economic History of the People_, 1904
4 M. Misra, _Vishnu’s Crowded Temple: India since the Great Rebellion_, 2007, p76
5 A. Bagchi, _Private Investment in India, 1900-1939_, pp. 49-53

60
metropoles, and denied investment to extend into more productive industry. Besides these factors, colonies suffer from purely extractive imperial institutions. This model is a popular explanation for underdevelopment, with various refinements.

The Lewis thesis might explain India’s industrial underdevelopment. If metropolitan investment funded improvements in infrastructure only to facilitate the export of primary products, it would support the Lewis thesis. Tomlinson proposes that the Great Depression shocked India out of its underdevelopment trap – capital was redeployed from primary production into industry. This study, therefore, addresses the relationship between the financing of primary production and underinvestment in industry, in order to test Tomlinson’s argument.

The growth of India’s industry was below average for Asia between 1890 and 1913, even compared with East Asia, or Russia, a primary exporting economy attempting import substitution policies. A pressing concern in development economics is the ‘missing middle’: developing countries have large numbers of very small firms, which face some constraint preventing them from expanding, and achieving the particular benefits of larger scale. What these constraints and benefits are has not been definitively answered.

The ‘international’ view of the Great Depression, expressed for example by Temin and Eichengreen, explains the crisis through the Gold Standard. The money supply of the advanced economies was linked to an asset of fixed quantity. The underlying cause was that two countries (France and the US) between them owned the majority of the world’s gold, restricting the money supply and aggregate demand available to all others. For several years real output collapsed, as gold stocks pinged back and forth between countries, until fiat currencies were introduced. After a century as the world financial capital, Britain was in the centre of the crisis. India in turn was a vital prop in Britain’s balance of payments, and its connection to the rest of Asia.

---

7 Tomlinson, op Cit, p42
8 Bénétrix, O’Rourke and Williamson, forthcoming
10 Temin, “The Transmission of the Great Depression”, JEP, 7 (2), 1993, pp.87-102
11 Eichengreen, op Cit, p192
12 Ibid, pp. 279-285
Despite India’s importance, the quantitative literature has paid little attention to India’s interaction with the rest of the world during the interwar period.

Conquered later, and distinct both ethnically and religiously, Burma was never as fully integrated into the administration of the British Raj. It is often left out of the history of colonial India. However, the era saw considerable migration from India to Burma, and strong trade and financial links between the two. With data that covers Burma, this study describes the experience of the Indian and Burmese economies, and their reaction to the interwar world economy.

3 Conceptual Framework and Identification Strategy

3.1 Conceptual Framework

I consider the following simplified model for the Indian economy in the interwar period. There are two business models in the financial system. In the first, financiers lend money fractionally, against a reserve of mainly physical stocks of gold and other commodities. In the second, financiers lend money fractionally, against a reserve of mainly financial assets (bonds, deposits at other banks, etc.).

There are two sectors in the Indian economy. The first is primary product production and export. The second is industrial capital accumulation and production mainly for the domestic market. Both sectors require capital, and both sectors must mortgage future profits as collateral (i.e. forfeit farmland or plant). Due to ingrained historical practice, the first financial business model prefers serving the first sector over the second sector, given rates of return are equal.

I assume there is a fixed level of capital and an infinite supply of labour. Each time period, capital holders allocate between the two sectors to maximise return, depending on their relative returns in that time period. They also minimise the opportunity cost of their reserve and collateral, depending on the external price of gold, farmland and financial assets.

The rate of return to capital in each sector is set exogenously of the level of capital, either in the current or previous time period. If these are the same, the initial equilibrium favours capital held in the primary product sector, to the detriment of the secondary sector.

I consider the following simplified model for the Indian economy in the interwar period. There are two business models in the financial system. In the first, financiers lend money fractionally, against a reserve of mainly physical stocks of gold and other commodities. In the second, financiers lend money fractionally, against a reserve of mainly financial assets (bonds, deposits at other banks, etc.).

There are two sectors in the Indian economy. The first is primary product production and export. The second is industrial capital accumulation and production mainly for the domestic market. Both sectors require capital, and both sectors must mortgage future profits as collateral (i.e. forfeit farmland or plant). Due to ingrained historical practice, the first financial business model prefers serving the first sector over the second sector, given rates of return are equal.

I assume there is a fixed level of capital and an infinite supply of labour. Each time period, capital holders allocate between the two sectors to maximise return, depending on their relative returns in that time period. They also minimise the opportunity cost of their reserve and collateral, depending on the external price of gold, farmland and financial assets.

The rate of return to capital in each sector is set exogenously of the level of capital, either in the current or previous time period. If these are the same, the initial equilibrium favours capital held in the primary product sector, to the detriment of the secondary sector.

13 Tomlinson advances this model. Ibid, pp. 159-161.
3.2 Identification Strategy

There are three tests. The first examines the link between the formal financial sector and primary product exports. I use ordinary least squares to estimate (standard errors are clustered by district):

\[
\text{credit availability}_{it} = \alpha + \beta \text{commodity prices}_{jt} \times \text{crop suitability}_{ij} + \gamma t + \delta i + \epsilon_{jt} + \eta_i \times t
\]

(1)

Year and district fixed effects, \( \gamma t + \delta i \), and district time trends, \( \eta_i \times t \), absorb any variation due to fixed effects (this district has more railways, that year’s rainfall was bad across all India). ‘Credit availability’ is the size of the formal financial sector in a particular district in a particular year – how much is deposited, and how much is loaned out. ‘Commodity prices’ is a vector of world prices for \( n \) crops, \( j \), in each year. ‘Crop suitability’ is a vector of measures of the capacity of a district to grow each \( j \) primary crop. The dot product of these two vectors is the price shock: the potential export value of a district’s primary produce in each year.

\[
\text{priceshock}_{it} = \sum_{j=1}^{n} \text{commodity price}_{jt} \times \text{crop suitability}_{ij}
\]

The model predicts a negative coefficient for \( \beta \). The higher the commodity prices, the more profit to be made lending to villagers against gold. If farmland is worth less because the industrialised world is in Depression, investors sell gold and invest the proceeds where they will produce the highest dividends.

The second regression tests if it is true that financial assets are invested in more industry.

\[
\text{industrial investment}_{it} = \alpha + \beta \text{credit availability}_{it} + \zeta \text{controls} + \gamma t + \delta i + \epsilon_{jt} + \eta_i \times t
\]

(2)

The model predicts a positive value for \( \beta \), because Indian firms are capital constrained. The controls are other measures of financial sector size, to test if the variation is purely in credit availability or if some other factor, such as the paid-up capital of a bank, causes an effect on industry. In addition, there will be many other factors to firm location that will be absorbed into the fixed effects.
Thirdly, I use an instrumental variable approach to ensure that credit availability is robust to endogeneity concerns. The first stage equation is (1). For the second stage I estimate:

\[
\text{industrial_investment}_t = \alpha + \beta \hat{C}_t + \gamma_t + \delta_i + \epsilon_{jt} + \eta_i \times t
\] (3)

The second stage is a regression of industry on predicted credit availability, \( \hat{C} \), where the price shock is used as an instrument for credit availability. As the variation in the price shock is caused by exogenous circumstances in faraway markets, this tests the overall hypothesis, that price shocks cause industrialisation. Additionally, I estimate using OLS:

\[
\text{industrial_investment}_t = \alpha + \beta \text{credit_availability}_t + \zeta \text{commodity_prices}_t \times \text{crop_suitability}_i + \gamma_t + \delta_i + \epsilon_{jt} + \eta_i \times t
\] (4)

which tests if price shocks have an effect on industry independently of finance. The coefficient \( \zeta \) shows how much the price shock causes industry directly. The model predicts an insignificant value for \( \zeta \).

4 Data

I create a dataset of the Indian economy drawn from government statistics published in the 1920s and 30s.

The data was entered by hand from 31 annually published statistics books across the three series: Statistical Tables relating to Banks in India, Large Industrial Establishments in India and Agricultural Statistics of India. The agricultural statistics were published since the 1880s by the Indian government; the s banking and industrial censuses that were introduced in the 1910s. My dataset comprises 5,913 annual entries over 595 banks, including balance sheet and branch locations, 14,476 factories by location and category from 9 censuses and 684 entries of acreage for 30 crops by district.

The survey of banks is published annually. Every bank with over Rs. 10,000 (£1,083) in capital is given an entry. There are four categories of banks: presidency banks, exchange banks, joint stock banks and co-operative banks. In terms of deposits, the first three sectors are roughly the same size. The co-operative bank sector was smaller, but with a faster rate of growth (it increased from 15% of the size of the joint stock bank sector to 35% from
1922 to 1938). Table 1 provides statistics for formal the financial sector between 1914 and 1938.

Table 2 gives the summary statistics per district. The negative value for the assets of the Imperial Bank of India in 1922 are carried over – one district has more branches of the Imperial Bank of India than other banks with more assets in that year.

The agricultural statistics series gives the area under cultivation of each major crop in each year and district. I have taken data from the beginning of the period to use a measure of crop suitability: each district can potentially produce this much of each crop, even if it produces less revenue in future years. For each district and princely state I have taken the average area sown of 28 crops over the 1919 and 1920 agricultural cycles, as well as the total area sown and area sown multiple times. Some, such as cinchona or Indian hemp, are very rare. Tobacco is very common. Particular regions specialise in producing crops such as tea (Assam), jute (east Bengal), wheat (the Punjab) or rice (Burma was a major exporter). Food-grains including beans and pulses comprise 77% of total area sown, cotton 5%. About 10% of agricultural land produces two harvests a year. The statistics for, say, rice in Bengal thus give the area sown of winter rice plus the area sown of summer rice.

The final measure is industrial capacity. The biennial industrial censuses list factories that employ over 20 people, by district and number of employees and according to 71 categories. The biggest factories are the cotton and jute mills in Bombay, Ahmedabad and Calcutta, as well as the iron and steel rolling mills at Jamshedpur, Bhadravathi and Kulti. The Tata mill in Jamshedpur employed over 27,000 people at its height in 1925. As a measure of the level of Industry, I have counted the number of factories employing over 100 people in each district and industrial sector. Some firms list several plants, for which, I have divided their employees evenly between the plants, and counted each of the plants if they employ more than 100. A further variable is calculated only counting plants in ‘heavy’ industry categories, such as textile mills or engineering works.

---

14 Bombay is two districts (Bombay Island and Bombay Suburban). These have been combined. Calcutta spills over into its neighbouring districts, Howrah, Hooghly and 24 Parganas, which are otherwise standard for rural Bengal. Calcutta’s banks are all located within the city’s district, but most of the industry is spread outside. I have combined all four into one large district. Following the Government of India Act 1935, Burma was split into a separate colony in 1937. For 1937 and 1938 Burma is absent from the statistics series. The tests are all robust to this omission.
5 Results

Figure 1 shows the growth of the formal financial sector over the period. Cooperative banks were only included in the surveys in 1922, causing the steep rise that year. There is no such administrative explanation for the spurt in the early 1930s, and the increase in the gradient. The increase is due to more bank formations, or existing banks growing over the capital threshold for inclusion in the survey. The size of the total Indian financial sector before the 1930s was not small – most were not reported. More banks were switching to become joint stock banks, and more capital was flowing to the recorded sector away from lending to cash crop cultivators. More banks enter the survey than exit. Of the 595 banks covered in total over the period, 479 are extant in 1938.

Figure 1

Crore acres of crop sown are weighted by an index of the world price. Potential acreage is kept constant, so the variation over time is only due to world prices. The collapse of agricultural prices began at the beginning of the decade, after the spike of postwar inflation. The subsequent rallies in 1923 and 1924 and 1928 were brief, and the downward trend only began to recover by the end of the period for food crops. Cash crops never reached their former peak.

The financial aggregates are regressed on combinations of price shocks and fixed effects for district, time and time-district trends. As expected, there is a
significant negative relationship along a variety of specifications of credit and debt. The significance level is 1% for assets on the aggregate price shock. One standard deviation downwards of the price shock is worth roughly half a standard deviation of assets upwards. When the price shock is disaggregated to cash crops and food staples, most of the variation is captured by the cash crops for export. However, cash crops alone do not have the same magnitude as the total shock.

Bombay, Rangoon and Madras districts are major financial sectors, but have no agricultural land and do not experience the price shocks. If uncaptured variation of these districts is correlated with finance, it will distort the results. Minus these and Calcutta, the rest of India experiences lower variation in financial assets due to food price shocks, while primary products grown for export are very significant.

When the shock is weighted by the proportion of multiple harvests, the food price shock reverses direction. With shorter agricultural cycles, there is less demand for loans for food cultivation. However, for cash crops the direction and magnitude of the effect is unchanged, after accounting for the weighting. This may reflect greater control over more valuable cash crops by moneylenders, e.g. forcing cultivators to borrow seeds.

Total deposits, and cash and bullion behave in mostly the same way as assets. Shorter agricultural cycles increase demand for credit, at least for staple food crops. However, even accounting for loans between co-operative banks, the financial centres are large enough to swamp this result. Weighted food price shocks are significant, with a positive correlation, once the financial sectors are removed. Otherwise, the measure for consumer loans follows the measure for all assets.

The regressions for industry and finance are shown in table 4. The coefficients are small – a district must have an extra Rs. 550,000 (£60,000 at the then fixed exchange rate) in loans to gain one large factory, almost twice the mean loans due in a district in the period.

The assets variable gives the portion of the bank’s balance sheet not held as cash. If it is excluded from the regression there is no significant correlation – deposits or other variables for the financial sector do not matter for the growth of industry, only the level of loans given by banks. Once isolated, the effect from assets on industry is significant and positive, as the model predicts.
The result is not robust to dropping the presidency capitals. However, for this regression table, dropping the presidency cities is distortionary. The test is for whether the total financial assets in a district are correlated with industry, and Bombay and Calcutta dominate both sectors (so do Madras and Rangoon in their respective provinces) – the question is whether there is some other unobservable that is causing industrialisation, rather than access to finance.

6 Conclusion

I have assembled a panel dataset on the economy of India between the two World Wars. My data covers agriculture, banks and industry for 250 districts. I calculated variables that show the amount of loans owed to banks in a district and the potential export value of the agricultural produce of a district. The export value was calculated using world prices, to produce an exogenous price shock on a district. Panel regressions show that there is a significant negative relationship between the price shock and the level of financial assets. There were other sources of finance apart from banks, indigenous networks that specialised in serving the agricultural sector. The first result suggests that as agricultural exports were profitable, the indigenous networks dominated. When primary product exports were unprofitable, the indigenous networks moved into formal banking.

At the same time as agricultural export prices fell, the world price of gold rose. India could sell the stock of gold amassed through decades of an overvalued exchange rate, releasing in exchange a slew of financial assets not tied down to the primary product export.

Secondly, I show that there is a significant positive relationship between bank assets and large firms. The increase in the number of firms with more than 100 employees is evidence that small and medium firms were constrained by capital availability, although there were also other factors, such as a government more willing to protect industry with tariffs. Combining the two results in an IV regression shows that an increase in capital, from a decrease in primary product export value, causes an increase in industry.

There are limitations in the data, which future research might overcome. The proxy for industry is the number of firms with over 100 employees, but this does not accurately equate to the extent of fixed investment in plant and machinery, that was the main constraints on firms. Several fixed effects might vary over the course of the period, for instance new railways, the increase in sugar cultivation or the introduction of new tariffs on certain primary products.
Donaldson (2010) uses annual rainfall as an exogenous shock to district level agricultural output. A similar measure might refine the price shocks used in this study.

Tomlinson argues that the trade in primary products was the basis for British rule in India: 

"By the eve of the Second World War the Indian economy was a great deal less "colonial" than it had been twenty years before." My study confirms this thesis, that India became more autarkic due to the Great Depression. A financial sector that primarily funded agricultural exports with a capital base of bullion was transformed into a network of banks ready to invest in neglected industry.
7 Bibliography

Official Publications
Government of India Press, Calcutta:
Agricultural Statistics of India, vols I and II (annual)
Large Industrial Establishments in India (biennial)
Statistical Tables relating to Banks in India (annual)
Travancore Administrative Report, 1926/7

Published Secondary Sources


Balachandaran, G., 1996, John Bullion’s Empire: Britain’s Gold Problem and India between the Wars, Richmond: Curzon


Hardiman, D., 1996, Feeding the Baniya: Peasants and Userers in Western India, Delhi: Oxford University Press


Table 1: Bank Statistics 1914-38

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-operative Capital</td>
<td>3.24</td>
<td>176</td>
<td>230</td>
<td>-</td>
<td>3,342</td>
</tr>
<tr>
<td>Reserve</td>
<td>3.23</td>
<td>139</td>
<td>214</td>
<td>-</td>
<td>2,773</td>
</tr>
<tr>
<td>Deposits</td>
<td>2.24</td>
<td>1,329</td>
<td>2,528</td>
<td>-</td>
<td>45,712</td>
</tr>
<tr>
<td>Consumer Loans</td>
<td>4.24</td>
<td>363</td>
<td>908</td>
<td>-</td>
<td>10,799</td>
</tr>
<tr>
<td>Loans to other Banks</td>
<td>2.67</td>
<td>989</td>
<td>1,567</td>
<td>-</td>
<td>326,23</td>
</tr>
<tr>
<td>Cash and Bullion</td>
<td>2.32</td>
<td>86</td>
<td>332</td>
<td>-</td>
<td>2,471,99</td>
</tr>
<tr>
<td>Assets</td>
<td>3.32</td>
<td>1,518</td>
<td>2,681</td>
<td>454</td>
<td>45,923</td>
</tr>
<tr>
<td>Balsheet</td>
<td>7</td>
<td>1,602</td>
<td>2,824</td>
<td>-</td>
<td>46,201</td>
</tr>
<tr>
<td>Exchange Capital</td>
<td>183</td>
<td>18,623</td>
<td>27,872</td>
<td>659</td>
<td>115,096</td>
</tr>
<tr>
<td>Reserve</td>
<td>183</td>
<td>23,166</td>
<td>34,604</td>
<td>818</td>
<td>138,921</td>
</tr>
<tr>
<td>Deposits</td>
<td>183</td>
<td>3,345</td>
<td>425,289</td>
<td>623</td>
<td>1,884,20</td>
</tr>
<tr>
<td>Cash and Bullion</td>
<td>183</td>
<td>5,370</td>
<td>10,788</td>
<td>97</td>
<td>41,501</td>
</tr>
<tr>
<td>Assets</td>
<td>183</td>
<td>9,895</td>
<td>575,203</td>
<td>2</td>
<td>322,339</td>
</tr>
</tbody>
</table>

Table 1: Bank Statistics 1914-38

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-operative</td>
<td>2.00</td>
<td>926</td>
<td>2,287</td>
<td>-</td>
<td>22,669</td>
</tr>
<tr>
<td>Capital</td>
<td>2.00</td>
<td>489</td>
<td>1,482</td>
<td>-</td>
<td>13,221</td>
</tr>
<tr>
<td>Reserve</td>
<td>2.00</td>
<td>8,432</td>
<td>28,814</td>
<td>-</td>
<td>334,662</td>
</tr>
<tr>
<td>Deposits</td>
<td>2.00</td>
<td>1,383</td>
<td>5,171</td>
<td>-</td>
<td>88,145</td>
</tr>
<tr>
<td>Cash and Bullion</td>
<td>2.00</td>
<td>8,451</td>
<td>27,184</td>
<td>-</td>
<td>322,339</td>
</tr>
<tr>
<td>JSB</td>
<td>2.00</td>
<td>926</td>
<td>2,287</td>
<td>-</td>
<td>22,669</td>
</tr>
<tr>
<td>Capital</td>
<td>2.00</td>
<td>489</td>
<td>1,482</td>
<td>-</td>
<td>13,221</td>
</tr>
<tr>
<td>Reserve</td>
<td>2.00</td>
<td>8,432</td>
<td>28,814</td>
<td>-</td>
<td>334,662</td>
</tr>
<tr>
<td>Deposits</td>
<td>2.00</td>
<td>1,383</td>
<td>5,171</td>
<td>-</td>
<td>88,145</td>
</tr>
<tr>
<td>Cash and Bullion</td>
<td>2.00</td>
<td>8,451</td>
<td>27,184</td>
<td>-</td>
<td>322,339</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>2,00</td>
<td>9,834</td>
<td>31,882</td>
<td>-</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>---</td>
</tr>
<tr>
<td>Balsheet</td>
<td>2,00</td>
<td>7</td>
<td>9,834</td>
<td>31,882</td>
<td>-</td>
</tr>
<tr>
<td>Branches</td>
<td>2,00</td>
<td>7</td>
<td>9</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Presidency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital</td>
<td>39</td>
<td>3,270</td>
<td>2,246</td>
<td>750</td>
<td>5,630</td>
</tr>
<tr>
<td>Reserve</td>
<td>39</td>
<td>3,079</td>
<td>2,111</td>
<td>480</td>
<td>5,850</td>
</tr>
<tr>
<td>Deposits</td>
<td>39</td>
<td>43,015</td>
<td>27,742</td>
<td>7,620</td>
<td>77,833</td>
</tr>
<tr>
<td>Cash and Bullion</td>
<td>39</td>
<td>10,759</td>
<td>5,352</td>
<td>2,560</td>
<td>20,970</td>
</tr>
<tr>
<td>Assets</td>
<td>39</td>
<td>11,212</td>
<td>14,242</td>
<td>-4,367</td>
<td>45,499</td>
</tr>
<tr>
<td>Balsheet (Thousand Rupees)</td>
<td>39</td>
<td>49,364</td>
<td>32,008</td>
<td>9,160</td>
<td>88,386</td>
</tr>
</tbody>
</table>
Table 2: Summary Statistics 1914-38

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital</td>
<td>4,334</td>
<td>1,224</td>
<td>3,886</td>
<td>-</td>
<td>51,677</td>
</tr>
<tr>
<td>Reserve Capital</td>
<td>4,334</td>
<td>1,241</td>
<td>4,108</td>
<td>-</td>
<td>51,481</td>
</tr>
<tr>
<td>Deposits and other Liabilities</td>
<td>4,334</td>
<td>12,330</td>
<td>38,158</td>
<td>-</td>
<td>382,759</td>
</tr>
<tr>
<td>Cash and Bullion</td>
<td>4,334</td>
<td>1,903</td>
<td>6,458</td>
<td>-</td>
<td>97,615</td>
</tr>
<tr>
<td>Loans to other Banks</td>
<td>4,334</td>
<td>621</td>
<td>1,564</td>
<td>-</td>
<td>29,483</td>
</tr>
<tr>
<td>Assets</td>
<td>4,334</td>
<td>4,422</td>
<td>9,526</td>
<td>-</td>
<td>51,580</td>
</tr>
<tr>
<td>BalSheet</td>
<td>4,334</td>
<td>16,961</td>
<td>54,959</td>
<td>-</td>
<td>659,835</td>
</tr>
<tr>
<td>Consumer Loans (Hundred Rupees)</td>
<td>4,334</td>
<td>3,801</td>
<td>8,505</td>
<td>-</td>
<td>115,036</td>
</tr>
<tr>
<td>capratio</td>
<td>4,452</td>
<td>18.72%</td>
<td>13.21%</td>
<td>0.00%</td>
<td>100.00</td>
</tr>
<tr>
<td>cashratio</td>
<td>4,452</td>
<td>12.39%</td>
<td>8.09%</td>
<td>0.00%</td>
<td>70.19%</td>
</tr>
<tr>
<td>Proportional area of cash crops</td>
<td>5,292</td>
<td>8.10%</td>
<td>10.18%</td>
<td>0.00%</td>
<td>64.23%</td>
</tr>
<tr>
<td>Proportional area multiple harvests</td>
<td>5,292</td>
<td>10.56%</td>
<td>9.12%</td>
<td>0.00%</td>
<td>41.91%</td>
</tr>
<tr>
<td>priceshock</td>
<td>5,292</td>
<td>45.53</td>
<td>53.44</td>
<td>-</td>
<td>891.16</td>
</tr>
<tr>
<td>priceshock food</td>
<td>5,292</td>
<td>38.60</td>
<td>42.89</td>
<td>-</td>
<td>577.78</td>
</tr>
<tr>
<td>priceshock cash</td>
<td>5,292</td>
<td>6.92</td>
<td>18.17</td>
<td>-</td>
<td>417.37</td>
</tr>
<tr>
<td>xpriceshock</td>
<td>5,274</td>
<td>5.26</td>
<td>7.76</td>
<td>-</td>
<td>64.94</td>
</tr>
<tr>
<td>xpriceshock food</td>
<td>5,274</td>
<td>4.75</td>
<td>7.19</td>
<td>-</td>
<td>61.84</td>
</tr>
<tr>
<td>xpriceshock cash</td>
<td>5,274</td>
<td>0.51</td>
<td>1.01</td>
<td>-</td>
<td>15.51</td>
</tr>
<tr>
<td>biennial priceshock</td>
<td>5,292</td>
<td>47.31</td>
<td>57.52</td>
<td>-</td>
<td>1,186</td>
</tr>
<tr>
<td>biennial priceshock food</td>
<td>5,292</td>
<td>40.06</td>
<td>46.53</td>
<td>-</td>
<td>761.06</td>
</tr>
<tr>
<td>biennial priceshock cash</td>
<td>5,292</td>
<td>7.25</td>
<td>19.08</td>
<td>-</td>
<td>425.72</td>
</tr>
<tr>
<td>Industry</td>
<td>2,553</td>
<td>3.77</td>
<td>18.77</td>
<td>0</td>
<td>308.00</td>
</tr>
<tr>
<td>Heavy Industry</td>
<td>2,553</td>
<td>3.45</td>
<td>17.24</td>
<td>0</td>
<td>277.00</td>
</tr>
</tbody>
</table>

Table 2: Summary Statistics 1914-38

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital</td>
<td>4,334</td>
<td>1,224</td>
<td>3,886</td>
<td>-</td>
<td>51,677</td>
</tr>
<tr>
<td>Reserve Capital</td>
<td>4,334</td>
<td>1,241</td>
<td>4,108</td>
<td>-</td>
<td>51,481</td>
</tr>
<tr>
<td>Deposits and other Liabilities</td>
<td>4,334</td>
<td>12,330</td>
<td>38,158</td>
<td>-</td>
<td>382,759</td>
</tr>
<tr>
<td>Cash and Bullion</td>
<td>4,334</td>
<td>1,903</td>
<td>6,458</td>
<td>-</td>
<td>97,615</td>
</tr>
<tr>
<td>Loans to other Banks</td>
<td>4,334</td>
<td>621</td>
<td>1,564</td>
<td>-</td>
<td>29,483</td>
</tr>
<tr>
<td>Assets</td>
<td>4,334</td>
<td>4,422</td>
<td>9,526</td>
<td>-</td>
<td>51,580</td>
</tr>
<tr>
<td>BalSheet</td>
<td>4,334</td>
<td>16,961</td>
<td>54,959</td>
<td>-</td>
<td>659,835</td>
</tr>
<tr>
<td>Consumer Loans (Hundred Rupees)</td>
<td>4,334</td>
<td>3,801</td>
<td>8,505</td>
<td>-</td>
<td>115,036</td>
</tr>
<tr>
<td>capratio</td>
<td>4,452</td>
<td>18.72%</td>
<td>13.21%</td>
<td>0.00%</td>
<td>100.00</td>
</tr>
<tr>
<td>cashratio</td>
<td>4,452</td>
<td>12.39%</td>
<td>8.09%</td>
<td>0.00%</td>
<td>70.19%</td>
</tr>
<tr>
<td>Proportional area of cash crops</td>
<td>5,292</td>
<td>8.10%</td>
<td>10.18%</td>
<td>0.00%</td>
<td>64.23%</td>
</tr>
<tr>
<td>Proportional area multiple harvests</td>
<td>5,292</td>
<td>10.56%</td>
<td>9.12%</td>
<td>0.00%</td>
<td>41.91%</td>
</tr>
<tr>
<td>priceshock</td>
<td>5,292</td>
<td>45.53</td>
<td>53.44</td>
<td>-</td>
<td>891.16</td>
</tr>
<tr>
<td>priceshock food</td>
<td>5,292</td>
<td>38.60</td>
<td>42.89</td>
<td>-</td>
<td>577.78</td>
</tr>
<tr>
<td>priceshock cash</td>
<td>5,292</td>
<td>6.92</td>
<td>18.17</td>
<td>-</td>
<td>417.37</td>
</tr>
<tr>
<td>xpriceshock</td>
<td>5,274</td>
<td>5.26</td>
<td>7.76</td>
<td>-</td>
<td>64.94</td>
</tr>
<tr>
<td>xpriceshock food</td>
<td>5,274</td>
<td>4.75</td>
<td>7.19</td>
<td>-</td>
<td>61.84</td>
</tr>
<tr>
<td>xpriceshock cash</td>
<td>5,274</td>
<td>0.51</td>
<td>1.01</td>
<td>-</td>
<td>15.51</td>
</tr>
<tr>
<td>biennial priceshock</td>
<td>5,292</td>
<td>47.31</td>
<td>57.52</td>
<td>-</td>
<td>1,186</td>
</tr>
<tr>
<td>biennial priceshock food</td>
<td>5,292</td>
<td>40.06</td>
<td>46.53</td>
<td>-</td>
<td>761.06</td>
</tr>
<tr>
<td>biennial priceshock cash</td>
<td>5,292</td>
<td>7.25</td>
<td>19.08</td>
<td>-</td>
<td>425.72</td>
</tr>
<tr>
<td>Industry</td>
<td>2,553</td>
<td>3.77</td>
<td>18.77</td>
<td>0</td>
<td>308.00</td>
</tr>
<tr>
<td>Heavy Industry</td>
<td>2,553</td>
<td>3.45</td>
<td>17.24</td>
<td>0</td>
<td>277.00</td>
</tr>
</tbody>
</table>
### Table 3: Agriculture and Finance 1921-38

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>(\text{Con. Loans})</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
</tr>
<tr>
<td>Assets</td>
<td>Assets</td>
<td>Assets</td>
<td>Assets</td>
<td>Assets</td>
<td>Assets</td>
<td>Loans</td>
</tr>
</tbody>
</table>

- **priceshock**
  - \(-55.04^*\)  
  - **(19.86)**
- **priceshock food**
  - \(-49.13\)
  - **(36.09)**
  - \(-10.47\)
  - **(8.84)**
  - \(-5.806\)
  - **(7.925)**
- **priceshock cash**
  - **65.63^*\)
  - **77.55^*\)
  - **60.87^*\)
  - **(11.57)**
- **xpriceshock**
  - **-48.11\)
  - **(53.89)**
- **xfoodshock**
  - **127.9^*\)
  - **(42.66)**
- **xcashshock**
  - **1,454^*\)
  - **(393.6)**

- **Non Presidency**
  - No
  - Yes
  - Yes
  - Yes
  - Yes
  - Yes

- **Fixed Effects**
  - Yes
  - Yes
  - Yes
  - Yes
  - Yes
  - Yes

- **Observations**
  - 3,839
  - 3,839
  - 3,788
  - 3,788
  - 3,771
  - 3,771

- **Number of Districts**
  - 226
  - 226
  - 222
  - 222
  - 222

- **R-squared**
  - 0.265
  - 0.265
  - 0.327
  - 0.281
  - 0.324
  - 0.301
Table 3 (cont.): Agriculture and Finance 1921-38

<table>
<thead>
<tr>
<th></th>
<th>G OLS</th>
<th>H OLS</th>
<th>I OLS</th>
<th>J OLS</th>
<th>K OLS</th>
<th>L OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Con. Loans</td>
<td>Deposits</td>
<td>Deposits</td>
<td>Cash</td>
<td>Cash</td>
<td></td>
</tr>
<tr>
<td>priceshock</td>
<td>** 127.2*</td>
<td>32.36*</td>
<td>** 155.5*</td>
<td>** 136.5*</td>
<td>** 1172*</td>
<td>1119*</td>
</tr>
<tr>
<td>priceshock food</td>
<td>-3.893 (16.03)</td>
<td>0.784 (3.066)</td>
<td>- -</td>
<td>- -</td>
<td>75.84* (14.92)</td>
<td>10.44* (2.253)</td>
</tr>
<tr>
<td>priceshock cash</td>
<td>- -</td>
<td>- -</td>
<td>5.88* (523.8)</td>
<td>355.6 (51.8)</td>
<td>1172* (344.7)</td>
<td>1119* (344.7)</td>
</tr>
<tr>
<td>xpriceshock</td>
<td>- -</td>
<td>- -</td>
<td>5.88* (523.8)</td>
<td>355.6 (51.8)</td>
<td>1172* (344.7)</td>
<td>1119* (344.7)</td>
</tr>
<tr>
<td>Non Presidency</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of Districts</td>
<td>226</td>
<td>222</td>
<td>222</td>
<td>222</td>
<td>222</td>
<td>222</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.239</td>
<td>0.297</td>
<td>0.173</td>
<td>0.173</td>
<td>0.075</td>
<td>0.075</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered by district. Non Presidency excludes Bombay, Madras, Calcutta and Rangoon. Fixed Effects are district and time fixed effects, and district fixed effects multiplied by time.
Table 4: Industry and Finance 1921-38

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>OLS</th>
<th>OLS</th>
<th>OLS</th>
<th>OLS</th>
<th>OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Total</td>
<td>Total</td>
<td>Heavy</td>
<td>Heavy</td>
<td>Heavy</td>
</tr>
<tr>
<td>Industry</td>
<td>Industry</td>
<td>Industry</td>
<td>Industry</td>
<td>Industry</td>
<td>Industry</td>
<td>Industry</td>
</tr>
<tr>
<td>Capital</td>
<td>4.17</td>
<td>5.79</td>
<td>-5.79</td>
<td>9.54</td>
<td>4.26</td>
<td>-5.66</td>
</tr>
<tr>
<td></td>
<td>(10.3)</td>
<td>(10.3)</td>
<td>(10.0)</td>
<td>(7.86)</td>
<td>(3.95)</td>
<td>(10.3)</td>
</tr>
<tr>
<td>Deposits</td>
<td>-0.597</td>
<td>-0.943</td>
<td>0.257</td>
<td>0.042</td>
<td>-0.709</td>
<td>0.297</td>
</tr>
<tr>
<td></td>
<td>(0.874)</td>
<td>(0.874)</td>
<td>(0.375)</td>
<td>(0.652)</td>
<td>(0.619)</td>
<td>(0.337)</td>
</tr>
<tr>
<td>Cash</td>
<td>3.09</td>
<td>3.13*</td>
<td>0.664</td>
<td>1.55</td>
<td>2.49</td>
<td>0.463</td>
</tr>
<tr>
<td></td>
<td>(19.5)</td>
<td>(19.5)</td>
<td>(0.269)</td>
<td>(1.03)</td>
<td>(1.58)</td>
<td>(0.358)</td>
</tr>
<tr>
<td>Assets</td>
<td>1.77**</td>
<td>0.379</td>
<td>1.51***</td>
<td>0.392</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.573)</td>
<td>(0.491)</td>
<td>(0.475)</td>
<td>(0.252)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non Presidency</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>2,011</td>
<td>2,011</td>
<td>1,960</td>
<td>1,784</td>
<td>2,011</td>
<td>1,960</td>
</tr>
<tr>
<td>Number of Districts</td>
<td>222</td>
<td>222</td>
<td>218</td>
<td>222</td>
<td>222</td>
<td>218</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.612</td>
<td>0.597</td>
<td>0.470</td>
<td>0.646</td>
<td>0.563</td>
<td>0.454</td>
</tr>
</tbody>
</table>

Table 4: Industry and Finance 1921-38

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>OLS</th>
<th>OLS</th>
<th>OLS</th>
<th>OLS</th>
<th>OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Total</td>
<td>Total</td>
<td>Heavy</td>
<td>Heavy</td>
<td>Heavy</td>
</tr>
<tr>
<td>Industry</td>
<td>Industry</td>
<td>Industry</td>
<td>Industry</td>
<td>Industry</td>
<td>Industry</td>
<td>Industry</td>
</tr>
<tr>
<td>Capital</td>
<td>4.17</td>
<td>5.79</td>
<td>-5.79</td>
<td>9.54</td>
<td>4.26</td>
<td>-5.66</td>
</tr>
<tr>
<td></td>
<td>(10.3)</td>
<td>(10.3)</td>
<td>(10.0)</td>
<td>(7.86)</td>
<td>(3.95)</td>
<td>(10.3)</td>
</tr>
<tr>
<td>Deposits</td>
<td>-0.597</td>
<td>-0.943</td>
<td>0.257</td>
<td>0.042</td>
<td>-0.709</td>
<td>0.297</td>
</tr>
<tr>
<td></td>
<td>(0.874)</td>
<td>(0.874)</td>
<td>(0.375)</td>
<td>(0.652)</td>
<td>(0.619)</td>
<td>(0.337)</td>
</tr>
<tr>
<td>Cash</td>
<td>3.09</td>
<td>3.13*</td>
<td>0.664</td>
<td>1.55</td>
<td>2.49</td>
<td>0.463</td>
</tr>
<tr>
<td></td>
<td>(19.5)</td>
<td>(19.5)</td>
<td>(0.269)</td>
<td>(1.03)</td>
<td>(1.58)</td>
<td>(0.358)</td>
</tr>
<tr>
<td>Assets</td>
<td>1.77**</td>
<td>0.379</td>
<td>1.51***</td>
<td>0.392</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.573)</td>
<td>(0.491)</td>
<td>(0.475)</td>
<td>(0.252)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non Presidency</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>2,011</td>
<td>2,011</td>
<td>1,960</td>
<td>1,784</td>
<td>2,011</td>
<td>1,960</td>
</tr>
<tr>
<td>Number of Districts</td>
<td>222</td>
<td>222</td>
<td>218</td>
<td>222</td>
<td>222</td>
<td>218</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.612</td>
<td>0.597</td>
<td>0.470</td>
<td>0.646</td>
<td>0.563</td>
<td>0.454</td>
</tr>
</tbody>
</table>
A MACRO-FINANCIAL MODEL FOR CREDIT RISK STRESS TESTING: THE CASE OF LATVIA

VIKTORYIA PILINKO
ANDREI ROMANCENCO

Stockholm School of Economics

ABSTRACT

This paper investigates the sources of credit risk in Latvian banking system. Using a three-step rigorous methodology we estimate resilience of Latvian banking system to shocks in chosen macro-financial variables. We build reduced form VAR model to uncover the relationship between macroeconomic and financial variables and forecast their development over a period of one year. Next, to identify the influence of macro-financial variables on banks' credit risk, measured as a share of non-performing loans to gross loans, we use panel data regression including chosen macro-financial variables as independent regressors. After we establish the relationship between NPLs and those variables, we introduce a plausible shock to the system based on a historical scenario of GDP development. Shocked NPL values are later used to calculate new capital adequacy ratios (CAR) for banks in order to check if they are still above the threshold of regulatory requirement of CAR equal to 8%. Results of the analysis indicate that not all the banks in the Latvian banking system would withstand a crisis that would entail a drop of GDP equal to that during the recent financial crisis. However, the Latvian banking system overall is quite resilient to macro-financial shocks due to its sound capitalization.

1 Introduction

Stress testing, whose popularity had been growing even in pre-crisis years, became one of the most widely advocated tools for banks following the 2008 financial crisis. Authorities such as the European Banking Authority, the European Commission, the Basel Committee introduce regulations that
specify requirements for the use and quality of stress tests by financial institutions. A methodological approach for examining the “potential vulnerability to exceptional but plausible events”, stress testing has proved to be an integral part of banks’ risk management practices.

The scope of this paper is to assess the resilience of the Latvian banking system to external shocks affecting its credit risk parameters. Credit risk is, by far, the most relevant risk factor in contemporary banking, and Latvia is no exception. Usually regulatory capital provisions for credit risk considerably surpass capital charges for other categories of risk (Bonti, Kalkbrener, Lotz, & Stahl, 2005). With loans amounting to nearly two thirds of the Latvian banks total assets (Bank of Latvia, 2012), the counterparty risk appears to be the major individual source of risk.

In this paper we scrupulously devise and discuss an approach for stress testing the banking system with respect to credit risk, given the Latvian context. In line with the literature, we measure credit risk as the proportion of non-performing loans to the total loans. Additionally, we formulate our own specification for the regression equation for the case of Latvia, examining whether shocks to macro-financial variables result in escalation of non-performing loans and subsequent capital losses.

We formulate the research problem at hand into the following research question:

**How resilient is the Latvian banking system to adverse macro-financial shocks simulated through the credit risk model?**

### 2 Overview of Latvian Banking System

The year 2014 started off with 3 credit institutions halting their operations – GE Money bank, UniCredit Bank and Latvijas Hipotēku un zemes banka - and 26 institutions present on Latvian market, 9 of them being branches of foreign banks operating in Latvia. Quality of the loans extended to households kept improving but at a pace slower than expected. Since its spike of 19.4% in the third quarter of 2010, the share of loans past due 90 days NPLs has been contracting, but at the end of 2013 it was still almost 6 times higher than before the crisis and equaled 8.3%. Besides, the ratio of provisions to the principal amount of the loans with payments of more than 90 days overdue remains high (around 74%). With regards to credit risk, the

### 2 Overview of Latvian Banking System

The year 2014 started off with 3 credit institutions halting their operations – GE Money bank, UniCredit Bank and Latvijas Hipotēku un zemes banka - and 26 institutions present on Latvian market, 9 of them being branches of foreign banks operating in Latvia. Quality of the loans extended to households kept improving but at a pace slower than expected. Since its spike of 19.4% in the third quarter of 2010, the share of loans past due 90 days NPLs has been contracting, but at the end of 2013 it was still almost 6 times higher than before the crisis and equaled 8.3%. Besides, the ratio of provisions to the principal amount of the loans with payments of more than 90 days overdue remains high (around 74%). With regards to credit risk, the

---

1 Credit risk is usually referred to as counterparty risk in the context of commercial banking.
banking sector remains well-capitalized: the capital adequacy ratio reached an all-time high of 18.94% at the end of 2013 versus regulatory minimum of 8%. Currently, it is well above the Euro Area average of 15%. Also, the lending pace remains slow - the loan portfolio continued to decline and contracted by 6.5% in 2013. At the same time, due to the introduction of euro, there has been a substantial increase in resident deposits (13.3%), which slightly improved the liquidity situation (Financial and Capital Market Commission, 2014).

In spite of gradually improving banking indicators, external risks coming from the sluggish growth and sovereign debt crisis in the euro area, can have a negative impact on the pace of Latvian financial system’s recovery. The beginning of the year 2014 brought concerns over Latvian economic situation with regards to the crisis between Russia and Ukraine. If the European Union imposes economic sanctions against Russia, and Russia does so in return, Latvian economy will be the one affected most.

Domestically, risks Latvian economy is facing are connected with several factors. The most alarming issue is labor market overheating: real gross wages have seen a slightly faster growth than productivity in the past year, which might pose a risk for competitiveness if the trend strengthens. Also, exports and investments have been quite sluggish; however, they are expected to pick up in 2014-2015 (Swedbank, 2014). Given these and other developments, there is a critical need for constant, quality monitoring of financial stability and banks’ ability to withstand crises (Bank of Latvia, 2012).

Given the trends and expectations outlined above, in particular strong capitalization of Latvian banks, we put forward the following hypothesis that should allow us to answer the research question in a more quantitative fashion:

**The capital adequacy ratio of Latvian banks will not fall below the minimum regulatory requirement of 8% in response to adverse macro-financial shocks.**

2 Literature Review

2.1 Banking Sector

Within the financial system, which is a complex set of closely interlinked markets and institutions used for financial contracting and the exchange of assets and risks, banks play a crucial role. These financial institutions take on the role of intermediation between borrowers and lenders: they channel the
funds from investors with an excess of them to those who experience a shortage of funds. Banks, along with other financial institutions, start operating in response to the existing market frictions, such as transaction costs and information asymmetry, and by their activities facilitate resource allocation. The ability of institutions to acquire and process large amounts of relevant information – something that is impossible to do on an individual level - as well as the ability to agglomerate and distribute capital while assuming risks faced by individuals lowers individuals’ transaction costs and moral hazards, thus, facilitating exchange (Levine, 1997).

Due to their very nature, banks are accompanied by multiple risks in their activity. As any organization, a bank is exposed to operational risk, the risk of loss due to the failure of internal systems and processes through which the company operates, as well as reputational risk – that of a loss due to damages in a company’s reputation. The inherent funding liquidity risk stems from the maturity mismatch of loans and deposits, with the former having longer maturities and the latter having short maturities. The focus of the literature, especially recently, has been largely on transitions from funding liquidity risk to market liquidity risk.

Banks are highly interlinked among themselves, thus, the risk of a contagious failure of financial system due to the rise in any of the mentioned risks has been termed as systemic risk. It is in particular the systemic nature of the risks that endangers the stability of the whole financial system. However, exposure to credit risk, arguably, remains to be the most pervasive and important source of problems for the banks worldwide, as discussed, among others, by Ali and Daly (2010), Mileris (2012), Castro (2013). Credit risk is the risk of loss of principal or financial reward due to a borrower's failure to repay a loan or otherwise meet a contractual obligation (Jorion, 2001). It stems from the fact that a borrower expects to cover the current funds from investors with an excess of them to those who experience a shortage of funds. Banks, along with other financial institutions, start operating in response to the existing market frictions, such as transaction costs and information asymmetry, and by their activities facilitate resource allocation. The ability of institutions to acquire and process large amounts of relevant information – something that is impossible to do on an individual level - as well as the ability to agglomerate and distribute capital while assuming risks faced by individuals lowers individuals’ transaction costs and moral hazards, thus, facilitating exchange (Levine, 1997).

Due to their very nature, banks are accompanied by multiple risks in their activity. As any organization, a bank is exposed to operational risk, the risk of loss due to the failure of internal systems and processes through which the company operates, as well as reputational risk – that of a loss due to damages in a company’s reputation. The movements in market prices may cause losses for a bank’s investment or trading portfolios, which is defined as market risk, with interest rate risk and currency risk being the most relevant for banks. Especially in the light of the recent events in the global financial markets, regulatory changes present a substantial risk for banks.

As illustrated by the global financial crisis of 2008-2009, risks connected with liquidity are among the ones that have the most crucial effects on the whole banking system. In the relevant literature, the notion of liquidity refers to the unrestrained ability of economic agents to exchange their wealth for other assets (Williamson, 2008). Funding liquidity is the ability of banks to meet their liabilities as they are due and fund more assets, as in BIS (2008). The inherent funding liquidity risk stems from the maturity mismatch of loans and deposits, with the former having longer maturities and the latter having short maturities. The focus of the literature, especially recently, has been largely on transitions from funding liquidity risk to market liquidity risk.

Banks are highly interlinked among themselves, thus, the risk of a contagious failure of financial system due to the rise in any of the mentioned risks has been termed as systemic risk. It is in particular the systemic nature of the risks that endangers the stability of the whole financial system. However, exposure to credit risk, arguably, remains to be the most pervasive and important source of problems for the banks worldwide, as discussed, among others, by Ali and Daly (2010), Mileris (2012), Castro (2013). Credit risk is the risk of loss of principal or financial reward due to a borrower's failure to repay a loan or otherwise meet a contractual obligation (Jorion, 2001). It stems from the fact that a borrower expects to cover the current
obligations with the future cash flows that are uncertain, and thus, he may be unable or unwilling to perform the given obligations. As banks assume this risk, they require a compensation for lending their capital. During the underwriting process, a credit institution has to assess the credit risk of lending to a particular potential borrower; naturally, the required return will increase if the perceived credit risk increases.

Another crucial part of analysis of borrowers’ creditworthiness involves measurement and control of credit risk determinants that affect the ability of borrowers to repay at the aggregated level. For the context of our study, important findings were made by Festi, Kavklera, and Repina (2011), who support the previous research by suggesting that the rapid credit growth and the rise in the amount of available finance for the newly admitted EU member states, among which Latvia can be found, can have a deteriorating effect on the performance of the countries’ banking sectors reflected in the non-performing loans dynamics. They state that the risk accumulated in a boom while consumption, investment, capital accumulation grows, materializes in the downturn, which was clearly demonstrated during the last crisis. In many cases a sharp rise in crediting was accompanied by lax credit standards and irresponsible assessment of credit risks. In Latvia, the rapid growth of GDP was possible due to the substantial growth in loan portfolio, which, paired with insufficient analysis, led to instability and growth in non-performing loans (Fainstein &Novikov, 2011). Similar results were achieved in a study on Croatia, in which the authors concluded that rapid growth in loans increased the probability of a rise in credit risk (Kraft & Jankov, 2005).

2.2 Stress Testing

In the last decade, especially in the context of the global financial crisis that highlighted the necessity of developing improved methodological and practical approaches to risk identification and monitoring in banks, stress testing has gained huge popularity among academicians, supervisory authorities, international organizations such as IMF as well as in banks internally. Defined as a methodological approach for examining the “potential vulnerability to exceptional but plausible events” (Bank for International Settlements, 2000); (Virolainen, 2004)), stress testing proved to be an integral and most comprehensive tool for an on-going assessment of banks’ resilience to various shocks.

In essence, a stress test represents an ample form of sensitivity analysis to adverse events such as a dramatic fall in exports, deterioration in total factor productivity, hyperinflation and other macroeconomic shocks as well as obligations with the future cash flows that are uncertain, and thus, he may be unable or unwilling to perform the given obligations. As banks assume this risk, they require a compensation for lending their capital. During the underwriting process, a credit institution has to assess the credit risk of lending to a particular potential borrower; naturally, the required return will increase if the perceived credit risk increases.

Another crucial part of analysis of borrowers’ creditworthiness involves measurement and control of credit risk determinants that affect the ability of borrowers to repay at the aggregated level. For the context of our study, important findings were made by Festi, Kavklera, and Repina (2011), who support the previous research by suggesting that the rapid credit growth and the rise in the amount of available finance for the newly admitted EU member states, among which Latvia can be found, can have a deteriorating effect on the performance of the countries’ banking sectors reflected in the non-performing loans dynamics. They state that the risk accumulated in a boom while consumption, investment, capital accumulation grows, materializes in the downturn, which was clearly demonstrated during the last crisis. In many cases a sharp rise in crediting was accompanied by lax credit standards and irresponsible assessment of credit risks. In Latvia, the rapid growth of GDP was possible due to the substantial growth in loan portfolio, which, paired with insufficient analysis, led to instability and growth in non-performing loans (Fainstein &Novikov, 2011). Similar results were achieved in a study on Croatia, in which the authors concluded that rapid growth in loans increased the probability of a rise in credit risk (Kraft & Jankov, 2005).
The rationale for using stress tests is three-fold. Firstly, they complement the value-at-risk (hereafter VaR) approach by showing the direction of exposure, by computing the magnitude of exposure and by looking at the long-term effects of a shock. Secondly, it is highly recommended that stress tests are applied by financial regulators. Stress testing unveils the systemic risks incorporated in the banking sector, and makes these risks more transparent via calculating potential losses exacerbated by a given shock. Thirdly, on a microeconomic level, the results obtained in stress tests are insightful for projecting capital allocations and checking the effectiveness of organizational systems (Vazquez, Tabak, & Souto, 2012).

3 Empirics

3.1 Dependent Variable

One of the most commonly used proxies for the probability of default is the ratio of non-performing loans to gross loans. By non-performing loans regulators and practitioners usually mean impaired loans and loans whose repayments are 90 or more days overdue. The definition may vary from country to country; however, the conventional notion used by the European Central Bank is 90+ days overdue. Originally, this metric characterizes the loan quality of a lender: the higher the ratio, the worse the quality of his/her loan portfolio. The on-going deterioration of the EU-wide loan portfolio quality provides additional incentives to study their dynamics. Invalid source specified. If the amount of non-performing loans increases, a bank has to increase its loan loss provisions (LLPs), allowances for doubtful and bad loans. However, given that LLPs do not magnify equally with NPLs, they [LLPs] would capture less information about the overall loan quality. Thus, we find it reasonable to employ the share of non-performing loans to the gross amount of loans as a proxy for the probability of default.

In case of Latvia, data on non-performing loans is quite limited. Quarterly data on NPLs classified as loans >90 days overdue is available from the fourth quarter of 2004 till the fourth quarter of 2013 on aggregated level (total for all banks) from the Financial and Capital Market Commission (FCMC). Data on individual banks’ NPLs is available at yearly frequency from year 2006 to year 2012 from Bankscope database, however, data is missing for several banks (e.g. Danske Bank, Expobank), and for others it’s not available for the whole period. In Bankscope, NPLs are classified as shocks in banking, financial systems, etc. In fact, they can be designed for testing any of the risks described in the previous section: liquidity, market, credit, etc.

The rationale for using stress tests is three-fold. Firstly, they complement the value-at-risk (hereafter VaR) approach by showing the direction of exposure, by computing the magnitude of exposure and by looking at the long-term effects of a shock. Secondly, it is highly recommended that stress tests are applied by financial regulators. Stress testing unveils the systemic risks incorporated in the banking sector, and makes these risks more transparent via calculating potential losses exacerbated by a given shock. Thirdly, on a microeconomic level, the results obtained in stress tests are insightful for projecting capital allocations and checking the effectiveness of organizational systems (Vazquez, Tabak, & Souto, 2012).

3 Empirics

3.1 Dependent Variable

One of the most commonly used proxies for the probability of default is the ratio of non-performing loans to gross loans. By non-performing loans regulators and practitioners usually mean impaired loans and loans whose repayments are 90 or more days overdue. The definition may vary from country to country; however, the conventional notion used by the European Central Bank is 90+ days overdue. Originally, this metric characterizes the loan quality of a lender: the higher the ratio, the worse the quality of his/her loan portfolio. The on-going deterioration of the EU-wide loan portfolio quality provides additional incentives to study their dynamics. Invalid source specified. If the amount of non-performing loans increases, a bank has to increase its loan loss provisions (LLPs), allowances for doubtful and bad loans. However, given that LLPs do not magnify equally with NPLs, they [LLPs] would capture less information about the overall loan quality. Thus, we find it reasonable to employ the share of non-performing loans to the gross amount of loans as a proxy for the probability of default.

In case of Latvia, data on non-performing loans is quite limited. Quarterly data on NPLs classified as loans >90 days overdue is available from the fourth quarter of 2004 till the fourth quarter of 2013 on aggregated level (total for all banks) from the Financial and Capital Market Commission (FCMC). Data on individual banks’ NPLs is available at yearly frequency from year 2006 to year 2012 from Bankscope database, however, data is missing for several banks (e.g. Danske Bank, Expobank), and for others it’s not available for the whole period. In Bankscope, NPLs are classified as
impaired loans. In order to include data for the longest period possible and get the maximum number of observations, we pick 15 banks that have balanced data on NPLs from year 2008 to year 2012, leaving us with 75 observations. The list of the banks in the sample can be found in Appendix 3.

3.2 Explanatory Variables

There is consensus among economists that the probability of default is associated with changes in macroeconomic and financial indicators (Hamerle, Dartsch, Jobst, & Plank, 2011); (Festic, Kavkler, & Repina, 2011). As suggested by Figlewski, Frydman, Liang (2012), there are three groups of macro-factors that influence a firm’s (a system’s) creditworthiness:

1. Factors related to general macroeconomic conditions (e.g. unemployment, inflation, etc.);
2. Factors characterizing real economy (e.g. growth of real GDP, terms of trade, etc.);
3. Factors reflecting financial market conditions (e.g. interest rates, equity returns, etc.).

Now, we have to come up with a set of exogenous macro-financial variables that would best explain the variation in the probability of default. We first consider different macro-financial variables that are most common in the literature, and then justify the choice of variables for our regression that are applicable specifically for Latvia. Below we present the ratio of NPLs to gross loans as a function of those variables and provide the expected signs of their coefficients:

\[ y_t = \frac{NPL_t}{GL_t} = f(GDP \text{ growth}, CPI, REER, Interest \text{ Rate}, \text{Stock Market Index}, \text{U, Loans-to-assets}) \]

Let us briefly provide the economic motivation behind each of the coefficients:

**GDP growth** (−) – the higher growth in total output would increase real salaries, lessen unemployment, thus, leading to greater loan repayment rates.

**Consumer Price Index** (+) – an increase in consumer prices would erode a part of the individual/corporate budget, thus, limiting the loan repayment ability.
Real Effective Exchange Rate (–) – an increase in REER means an increase in a country’s competitiveness, implying higher loan quality of the banking system.

Interest rate (+) – higher interest rates mean higher interest payments on loans for borrowers, therefore, they should be positively related with NPLs.

Stock Market Index (+) – an increase in financial development leads to an increase in the overall economic growth, thus benefiting the loan quality. However, because of the marginal development of the Baltic stock market, this measure is inapplicable for Latvia.

Unemployment (+) – as mentioned above, the higher unemployment rates implies the lower loan quality because of the less repayment ability.

Loans-to-assets (+) – share of a bank’s total loans in assets is considered as a proxy of risk taken by a bank. Rising loan-assets ratio is leads to increases in NPL ratio.

The data have been retrieved from the Central Statistical Bureau of Latvia and Eurostat. All data are seasonally adjusted and are collected at both quarterly and annual frequency.

4 Methodological Approach

The first stage of our methodological approach is estimation of the relationship between macroeconomic and financial variables that supposedly affect the banks’ loan portfolio quality. To forecast their development, we employ the reduced form VAR model that uses 5 variables (namely, change in GDP growth, change in Inflation, change in unemployment, change in 3-month EURIBOR, and growth in real effective exchange rate (REER)). Our regression specification contains only one lag of every explanatory variable. The rationale behind this approach consists in the attempt to preserve degrees of freedom in our small sample.

Lastly the system-wide shocks are introduced to the model and are reiterated to check the resilience of the Latvian banking sector.
At the next step, we use panel data regression to estimate the relationship between probability of default and the set of macro-financial indicators:

\[
\Delta NPL_{it} = \alpha_0 + \alpha_1 GDP_{it} + \alpha_2 \Delta M_{it} + \alpha_3 \Delta \text{Unemployment}_{it} + \alpha_4 \Delta \text{Loantoasset}_{it} + \alpha_5 \Delta \text{Euribor}_{it} + \alpha_6 \Delta \text{REERgr}_{it} + \epsilon_{it}
\]

Where:

- \(\Delta NPL_{it}\) is the change in the ratio of NPLs to gross loans of bank \(i\) at time \(t\)
- \(GDP_{it}\) is GDP growth at time \(t\)
- \(\Delta M_{it}\) is change in inflation (based on harmonized consumer price index) at time \(t\)
- \(\Delta \text{Unemployment}_{it}\) is change in unemployment rate at time \(t\)
- \(\Delta \text{Loantoasset}_{it}\) is change in ratio of a bank \(i\)'s total loans to total assets at time \(t\)
- \(\Delta \text{Euribor}_{it}\) is change in European interbank offer rate at time \(t\)
- \(\Delta \text{REERgr}_{it}\) is growth of real effective exchange rate at time \(t\)
- \(\epsilon_{it}\) is the white noise process, which is assumed to be independently and identically distributed with a mean of 0 and a variance of 1.

With this approach, we aim to estimate the impact of various macro determinants on the loan quality.

Next, we calculate the expected change in NPL ratio over one year using the coefficients from panel regression and values of macro-financial variables forecasted for the case of a system-wide shock in the first part of the analysis.

At the last step, we make calculations that answer the research question in a quantitative fashion. Using the predicted change in NPLs and additional assumptions, we estimate banks' new capital adequacy ratios and compare
them to a regulatory minimum of 8%. The following equation is used for this calculation:

\[ \frac{ CAR^* - OwnCapital - \Delta P }{ RWA - \Delta P } \]

Here, OwnCapital is bank’s own funds, RWA is risk weighted assets, and \( \Delta P \) is the expected increase in loan loss provisions in response to an increase in non-performing loans. It is calculated as \( \Delta P = \Delta NPLs \times \text{Loss given default} \), meaning that the amount of the additional loan loss provisions that the bank is supposed to make will equal to the amount of the loans that it doesn’t expect to recover. In line with the recovery rate typically assumed in Financial Stability Reports by the Bank of Latvia, we assume it to be equal to 40%, thus LGD is 60% (Bank of Latvia, 2012). When performing this calculation we make a conservative assumption that a bank’s profit for the forecasted year equals zero, thus, own capital is not increased by the amount of profit. The data on own funds and RWA are retrieved from the FCMC.

5 Review of Empirical Findings
5.1 After-shock Macro-financial Variables Development

We choose the historical scenario analysis in order to test the resilience of the Latvian banking system. In our set-up, we design the dramatic fall in GDP equal to 19% over the year of 2014, with assumed equal drops of 5.13% in each quarter. Below one can study the dynamics of forecasted values for variables over the next 4 quarters (created by authors):

---

2 Risk-weighted assets (RWA) – a bank’s assets weighted according to risk of different asset classes, used in order to calculate minimum capital needs (Jakubík, 2007).
3 A 19% drop in GDP corresponds to the all-time maximum drop of GDP in the Latvian contemporaneous history. GDP soared to this record low in 2009 (Bank of Latvia, 2012).
5.2 Results of the Panel Data Regression

We estimate the equation with changes in NPL ratio as dependent variable using the data on 15 Latvian banks for 5 years, starting with 2008 till 2012. The chosen 15 banks are the ones that have balanced NPL data during the period of 2008-2012, and they comprise more than 70% of total Latvian banking sector’s assets as of 2012. We run the regression using several different approaches. First, we use ordinary OLS regression to estimate the coefficients of the independent variables with heteroscedasticity-robust standard errors. Secondly, taking an advantage of panel data properties, we run the fixed effects regression with individual fixed effects using least square dummy variables (LSDV) regression with heteroskedasticity and autocorrelation consistent standard errors. As OLS regression may suffer from omitted variable bias, we choose fixed effects approach which controls for individual fixed variables that are constant over time. Euribor variable was omitted due to collinearity, so the final equation we used for regression is:

\[ \Delta \text{NPL}_{it} = \alpha_0 + \alpha_1 \Delta \text{GDPGr}_{it} + \alpha_2 \Delta \text{HICP}_{it} + \alpha_3 \Delta \text{Unemployment}_{it} + \alpha_4 \Delta \text{LoanToAsset}_{it}, \]

\[ + \alpha_5 \text{REERGr}_{it} + \nu_{it}, \]

We report the coefficients of the regressors in Appendix 1. The results of the regression are broadly in line with our expectations. All coefficients have expected signs. Loan-to-asset ratio has a significant coefficient with a positive sign: increasing ratio proxies for a rise in a bank’s risk, thus causing NPLs to increase as well. GDP growth has also proven to have a statistically significant negative effect on NPL growth: as expected, as the economy’s growth accelerates and both household and corporate incomes rise, there is a
positive effect on banks’ credit risk as proxied by NPL decreases. Besides, the coefficient before GDP growth is the largest in magnitude out of all variables – a one percentage point increase in GDP growth results in 0.7 percentage point decrease in NPL share in total loan portfolio. This confirms our expectations of a pervasive impact GDP growth has on Latvian borrowers’ probability of default.

5.3 Capital Adequacy

At the final stage of the analysis we assess the resilience of the Latvian banking system to the specified shock using capital adequacy ratio (CAR) as a metrics. According to the Basel Agreement, which applies to all Latvian banks, the minimum requirement for CAR is 8%, meaning that the banks’ own funds be should at least 8% of its risk-weighted assets (RWA). Using the coefficients acquired in the panel data regression of NPL ratio on macro-financial variable, we calculate the estimated change in NPL ratio that will happen in the year 2014. Materialisation of the baseline scenario, according to which the macro-financial environment is expected to develop in a favourable way, would lead to a decrease in NPL ratio of 3.07%. If, however, the shock in a form of a 19% drop of GDP materialises, NPL ratio is expected to rise by 12%. During the recent financial crisis after an actual drop of GDP of 18.74% in year 2009, NPL development was akin to the one simulated through our model. Then, NPL ratio for the whole banking sector increased from 3.6% in the fourth quarter of 2008 to 16.4% in the fourth quarter of 2009.

Next, we study the effect of a 12% rise in NPL ratio on banks’ capital adequacy ratio (Table 2). Using the formulas described in the methodology part and banks’ existing own capital, CAR, RWA and total loans data as of the end of 2013, we calculate the changed CAR for all banks in response to increased loan loss provisions. This forecast is not made for the three banks that seized their operations in 2013 and for Reverta Bank, whose financial results for the year 2013 have not been published yet. The results indicate that while for some banks the fall in CAR is not critically large, for two of them (DNB Bank and Norvik Bank) the ratio actually falls below the regulatory minimum of 8%, 5.44% for DNB and 7.70% for Norvik. For SEB Bank and SMP Bank the ratio decreases to quite a low level as well, being just slightly above 8%. Under this scenario the required capital injection in order to absorb the rise in credit risk and bring the CAR up to 8% again is 27 829 thousand lat (19 506 thousand euro), or 1.71% of total assets for DNB Bank and 1 159 thousand lat (812 thousand euro), or 0.21% of total assets for Norvik Bank. The estimation of this amount was made under the
The following results show that the Latvian banking system is overall resilient to a shock in macroeconomic environment in the country due to the high capitalisation of its largest banks. However, deteriorating CAR ratios, which fall below the regulatory requirement for 2 out of 11 analysed banks, indicate that concerns over credit risk’s impact on stability of Latvian banks remain highly relevant in today’s environment.

Table 2: Forecasted 2014 CAR by banks
(Figures are given in thousand lat)
Source: simulation; FCMC

<table>
<thead>
<tr>
<th>Bank</th>
<th>Own Capital</th>
<th>CAR</th>
<th>RWA</th>
<th>Loans</th>
<th>Change in NPLs</th>
<th>Provison Change</th>
<th>CAR*</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABLV Bank</td>
<td>187 318</td>
<td>17.53%</td>
<td>1 068 556</td>
<td>562 371</td>
<td>67 485</td>
<td>40 491</td>
<td>14.28%</td>
</tr>
<tr>
<td>Baltikums Bank</td>
<td>33 891</td>
<td>19.12%</td>
<td>177 254</td>
<td>31 190</td>
<td>3 743</td>
<td>2 246</td>
<td>18.08%</td>
</tr>
<tr>
<td>DNB Bank</td>
<td>157 241</td>
<td>13.27%</td>
<td>1 184 935</td>
<td>1 362 851</td>
<td>163 542</td>
<td>98 125</td>
<td>5.44%</td>
</tr>
<tr>
<td>Norvik Bank</td>
<td>39 617</td>
<td>10.11%</td>
<td>391 859</td>
<td>142 320</td>
<td>17 078</td>
<td>10 247</td>
<td>7.70%</td>
</tr>
<tr>
<td>PrivatBank</td>
<td>35 017</td>
<td>17.01%</td>
<td>205 861</td>
<td>86 015</td>
<td>10 321</td>
<td>6 193</td>
<td>14.44%</td>
</tr>
<tr>
<td>Regional Investment Bank</td>
<td>32 949</td>
<td>21.21%</td>
<td>155 346</td>
<td>49 090</td>
<td>5 890</td>
<td>3 535</td>
<td>19.38%</td>
</tr>
<tr>
<td>Rietumu Bank</td>
<td>208 798</td>
<td>18.46%</td>
<td>1 131 083</td>
<td>826 460</td>
<td>99 175</td>
<td>59 505</td>
<td>13.93%</td>
</tr>
<tr>
<td>SEB Bank</td>
<td>273 675</td>
<td>16.19%</td>
<td>1 690 395</td>
<td>1 874 385</td>
<td>224 926</td>
<td>134 95</td>
<td>8.92%</td>
</tr>
<tr>
<td>SMP Bank</td>
<td>15 734</td>
<td>12.86%</td>
<td>122 348</td>
<td>86 311</td>
<td>10 357</td>
<td>6 214</td>
<td>8.20%</td>
</tr>
<tr>
<td>Swedbank</td>
<td>625 049</td>
<td>27.95%</td>
<td>2 236 311</td>
<td>2 304 123</td>
<td>276 495</td>
<td>165 897</td>
<td>22.18%</td>
</tr>
<tr>
<td>Trust Commercia l Bank</td>
<td>28 766</td>
<td>16.65%</td>
<td>172 768</td>
<td>79 723</td>
<td>9 567</td>
<td>5 740</td>
<td>13.79%</td>
</tr>
</tbody>
</table>

The following results show that the Latvian banking system is overall resilient to a shock in macroeconomic environment in the country due to the high capitalisation of its largest banks. However, deteriorating CAR ratios, which fall below the regulatory requirement for 2 out of 11 analysed banks, indicate that concerns over credit risk’s impact on stability of Latvian banks remain highly relevant in today’s environment.

Table 2: Forecasted 2014 CAR by banks
(Figures are given in thousand lat)
Source: simulation; FCMC

<table>
<thead>
<tr>
<th>Bank</th>
<th>Own Capital</th>
<th>CAR</th>
<th>RWA</th>
<th>Loans</th>
<th>Change in NPLs</th>
<th>Provison Change</th>
<th>CAR*</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABLV Bank</td>
<td>187 318</td>
<td>17.53%</td>
<td>1 068 556</td>
<td>562 371</td>
<td>67 485</td>
<td>40 491</td>
<td>14.28%</td>
</tr>
<tr>
<td>Baltikums Bank</td>
<td>33 891</td>
<td>19.12%</td>
<td>177 254</td>
<td>31 190</td>
<td>3 743</td>
<td>2 246</td>
<td>18.08%</td>
</tr>
<tr>
<td>DNB Bank</td>
<td>157 241</td>
<td>13.27%</td>
<td>1 184 935</td>
<td>1 362 851</td>
<td>163 542</td>
<td>98 125</td>
<td>5.44%</td>
</tr>
<tr>
<td>Norvik Bank</td>
<td>39 617</td>
<td>10.11%</td>
<td>391 859</td>
<td>142 320</td>
<td>17 078</td>
<td>10 247</td>
<td>7.70%</td>
</tr>
<tr>
<td>PrivatBank</td>
<td>35 017</td>
<td>17.01%</td>
<td>205 861</td>
<td>86 015</td>
<td>10 321</td>
<td>6 193</td>
<td>14.44%</td>
</tr>
<tr>
<td>Regional Investment Bank</td>
<td>32 949</td>
<td>21.21%</td>
<td>155 346</td>
<td>49 090</td>
<td>5 890</td>
<td>3 535</td>
<td>19.38%</td>
</tr>
<tr>
<td>Rietumu Bank</td>
<td>208 798</td>
<td>18.46%</td>
<td>1 131 083</td>
<td>826 460</td>
<td>99 175</td>
<td>59 505</td>
<td>13.93%</td>
</tr>
<tr>
<td>SEB Bank</td>
<td>273 675</td>
<td>16.19%</td>
<td>1 690 395</td>
<td>1 874 385</td>
<td>224 926</td>
<td>134 95</td>
<td>8.92%</td>
</tr>
<tr>
<td>SMP Bank</td>
<td>15 734</td>
<td>12.86%</td>
<td>122 348</td>
<td>86 311</td>
<td>10 357</td>
<td>6 214</td>
<td>8.20%</td>
</tr>
<tr>
<td>Swedbank</td>
<td>625 049</td>
<td>27.95%</td>
<td>2 236 311</td>
<td>2 304 123</td>
<td>276 495</td>
<td>165 897</td>
<td>22.18%</td>
</tr>
<tr>
<td>Trust Commercia l Bank</td>
<td>28 766</td>
<td>16.65%</td>
<td>172 768</td>
<td>79 723</td>
<td>9 567</td>
<td>5 740</td>
<td>13.79%</td>
</tr>
</tbody>
</table>
In this paper we have presented a broad review of credit risk determinants as studied in academic literature and institutional empirical papers, as well as developed the methodology to stress test Latvian banking system with regards to its credit risk. While we specified a complex, multi-step approach to study the relationships between banking system’s credit risk and macro-financial variables and examine its resilience to adverse shocks, further exploration of relevant variables could improve the fit of the model; more bank-specific variables such as deposit-to-loan ratio or loan-to-GDP ratio could be added.

Moreover, significant contribution to existing literature could be made by developing a model that accounts for feedback loops between banking system and real economy. Leitner (2005) defines financial contagion as a shock initiated by one institution that is widely translated to the other parts of the system. The main motivation for analysing financial contagion is that the risks may multiply and materialize into a market-wide spiral. In case of Latvia, no similar research has been done yet.

In this paper, we pursue several objectives while performing our analysis. First of all, we perform an extensive review of the existing literature on credit risk determinants and justify the ones that would best suit Latvian context. Secondly, we build macro-financial VAR model that allows to examine relations between macroeconomic and financial variables, and make their forecasts for future periods. Thirdly, we use panel data to identify how responsive ratio of NPLs to gross loans is to changes in macro-financial variables. Lastly, we introduce a plausible shock into the system and check how banks’ CARs react to it.

By performing the last part of the analysis we are able to get the answer to the research question we posed at the outset of our research: How resilient is the Latvian banking system to adverse macro-financial shocks simulated through the credit risk model? Our analysis shows that we can reject our hypothesis: the capital adequacy ratio of Latvian banks will not fall below the minimum regulatory requirement of 8% in response to adverse macro-financial shocks. Two banks, DNB Bank and Norvik Bank, fail to comply with the requirement in response to a rising credit risk. Despite this, the results of our analysis demonstrate the general health of Latvian banking system and its overall resilience: even in the case of such an adverse macroeconomic development as a 19% fall in GDP, the majority of Latvian banks, including the four largest ones by assets, remain well-

Moreover, significant contribution to existing literature could be made by developing a model that accounts for feedback loops between banking system and real economy. Leitner (2005) defines financial contagion as a shock initiated by one institution that is widely translated to the other parts of the system. The main motivation for analysing financial contagion is that the risks may multiply and materialize into a market-wide spiral. In case of Latvia, no similar research has been done yet.

In this paper, we pursue several objectives while performing our analysis. First of all, we perform an extensive review of the existing literature on credit risk determinants and justify the ones that would best suit Latvian context. Secondly, we build macro-financial VAR model that allows to examine relations between macroeconomic and financial variables, and make their forecasts for future periods. Thirdly, we use panel data to identify how responsive ratio of NPLs to gross loans is to changes in macro-financial variables. Lastly, we introduce a plausible shock into the system and check how banks’ CARs react to it.

By performing the last part of the analysis we are able to get the answer to the research question we posed at the outset of our research: How resilient is the Latvian banking system to adverse macro-financial shocks simulated through the credit risk model? Our analysis shows that we can reject our hypothesis: the capital adequacy ratio of Latvian banks will not fall below the minimum regulatory requirement of 8% in response to adverse macro-financial shocks. Two banks, DNB Bank and Norvik Bank, fail to comply with the requirement in response to a rising credit risk. Despite this, the results of our analysis demonstrate the general health of Latvian banking system and its overall resilience: even in the case of such an adverse macroeconomic development as a 19% fall in GDP, the majority of Latvian banks, including the four largest ones by assets, remain well-
capitalized with the capital adequacy ratios above 8%. However, the failure of the two banks to withstand the shock shows that credit risk remains a highly relevant source of vulnerability for Latvian banks. Even though the overall macroeconomic outlook for Latvian economy is positive, prudential credit risk monitoring should stay among top priorities in banks’ risk assessment procedures.
Bibliography


Appendix 1

Output of panel regression on NPLs

Source: Created by the authors

<table>
<thead>
<tr>
<th>Dependent variable: change in ratio of NPLs to Gross loans (dNPL)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanatory variables</td>
<td>Regression coefficients</td>
</tr>
<tr>
<td>GDP growth</td>
<td>-0.701893* (0.3835481)</td>
</tr>
<tr>
<td>REER growth</td>
<td>-0.0188962 (0.7755508)</td>
</tr>
<tr>
<td>Δ Unemployment</td>
<td>0.1174022 (0.0798654)</td>
</tr>
<tr>
<td>Δ HICP</td>
<td>0.6227669 (0.5628968)</td>
</tr>
<tr>
<td>Δ Loans-to-Assets</td>
<td>0.2194955** (0.0949454)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.729479 (4.469601)</td>
</tr>
</tbody>
</table>

Summary statistics:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>R²</td>
<td>0.2958</td>
</tr>
<tr>
<td>N</td>
<td>75</td>
</tr>
</tbody>
</table>

Note: The table exhibits results of the basic regression of changes in ratio of NPLs to Gross loans. Regression is run on the data from 15 companies over 5 years, from 2008 to 2012. Heteroskedasticity and autocorrelation consistent standard errors are given in parentheses. Coefficients of dummy variables resulting from LSDV regression are not reported.

* - significant at the 10% significance level
** - significant at the 5% significance level
*** - significant at the 1% significance level
Bilateral Investment Treaties (BITs) have proliferated throughout the international system. While ostensibly commercial in purpose, do BITs have domestic political ramifications? I argue that BITs affect a leader’s tenure through their effect on the property rights environment in developing countries. BITs, by segmenting a country’s property rights environment for foreign and domestic firms, reduce the incentive for foreign firms to lobby for property rights protections in the host country thus reducing the quality of the domestic property rights environment. In autocracies, a stagnating domestic property rights regime benefits domestic business elites who can continue to stymie small and medium enterprises (SMEs). For democracies, a stagnating property rights regime harms SMEs. Using a dataset of developing country leaders over the period 1960-2001, I find support for my hypothesis that BITs increase the tenure of autocrats and decrease the tenure of democrats. The results highlight the consequences of the legalization of global investment on the domestic political economy.
It is commonly assumed in economics that the self-employed save more relative to employees due to greater precautionary demand for saving. However, this line of reasoning disregards self-selection into self-employment whereby unobservable characteristics weaken the desire to save. This paper utilises a combination of standard theoretical savings models and literature on the determinants of self-employment in order to evaluate the impact of these unobservable characteristics on the saving patterns of the self-employed. Using a combination of two-stage least squares regression and the Heckman selection model it is found that these unobservable characteristics have a quantitatively large negative effect on the stock of wealth held by the self-employed, although the effect is only statistically significant in the two-stage least squares model. These results call doubt upon previous research which has not accounted for selection bias, challenge the conventional notion the self-employed save more, and may have policy implications.

1 Introduction

A common assertion in economics is that the self-employed have a greater precautionary demand motive for saving. Their incomes are typically more volatile and, therefore, greater precautionary saving is necessary to prevent consumption falling below a minimum threshold level. As rational economic agents, they would accumulate a greater stock of wealth. However, insufficient attention has been paid to potential self-selection into self-employment. There may be a number of unobserved characteristics, such as risk levels or time preferences, which affect both an individual’s decision to save more, and may have policy implications.
enter self-employment and their wealth level. These characteristics could dampen the precautionary motive and actually lead the self-employed to save less than employees. Furthermore, not accounting for this in regression models could lead to endogeneity, creating an omitted variable bias in the estimated coefficient of being self-employed. Seeing as this self-selection could have important implications, the motivation of this paper will be to address the neglect of self-selection in economic literature and try and measure the impact of these unobservable characteristics.

This paper will take the original approach of synthesising theories on savings patterns and the lifecycle model with evidence regarding the determinants of entry into self-employment in order to address self-selection. Using the self-employment status of the father as an instrument, it will be shown that in a two-stage least squares model, self-selection based on these unobservable characteristics seem to lead the self-employed to save significantly, in both a statistical and quantitative sense, less than employees. In a Heckman model, we find these unobservable characteristics are large and negative but not statistically significant, which could imply a heterogeneous effects framework is required. That these unobserved characteristics appear important has a number of implications. Firstly, it calls doubt on previous estimations of saving patterns which do not take account of this endogeneity. Secondly, it challenges conventional thinking that the self-employed save more due to the precautionary motive, imploring future research to further explore and account for the effects of self-selection into self-employment. Finally, if it is felt this leads to the self-employed making sub-optimal saving decisions it could have policy implications.

2 Literature Review

The common assertion that the self-employed have a greater precautionary demand motive for saving was first made in Leland’s (1968) pioneering paper on a theoretical precautionary savings model. Using the two period model of consumption, Leland showed that:

\[ C = f(Y_t, Y_{t+1}, A, r, \sigma^2) \]

\[ \frac{\partial f}{\partial \sigma^2} < 0 \]

The consumption function (C) is a function of current income (Yt), discounted expected future income (Y_{t+1}), assets (A), interest rate (r) and, most importantly, variability of Y (\sigma^2). The negative partial differentiation of \sigma^2 shows the positive relationship between income variability and savings. A prediction of this theoretical model is that "we would expect the self-
employed to save more, as their incomes are more variable” (Leland, 1968, p.471).

Whilst there is evidence that the self-employed have greater volatility of incomes (Jensen & Shore, 2008), there is less consensus over whether this translates into greater saving. Skinner (1988) finds that the self-employed save less than those with other employment statuses, whilst Kazarosian’s (1997) panel study finds a positive relationship between saving and self-employment consistent with the precautionary motive narrative. Whilst both acknowledge potential endogeneity, neither satisfactorily account for this; therefore, the disparity in results could result from selection bias. This selection bias has not been adequately addressed in economic literature and as such this paper hopes to fill a gap in the literature and offer methods that could be used to improve the reliability of the results attained in future research.

In order to create an unbiased estimate of the effect of self-employment on wealth due to unobservable characteristics a suitable instrument needs to be found. Rees and Shah (1986) undergo a comprehensive study of the determinants of self-employment. However, many particularly significant determinants, such as education and age, are not suitable instruments as they are almost certain to have a causal effect on wealth.

It may therefore be useful to think beyond socio-economic factors. Environmental, psychological or social capital factors may influence these unobservable factors, as discussed by Sørensen (2007). Furthermore, genetic factors may be important, as discussed by Nicolaou et al. (2008). A number of studies, including Henley (2004) and Hout and Rosen (2000) find that the employment status of parents is significant, and this may work through the factors previously outlined. Parents may pass on genes which influence personality traits and dispositions to certain employment types. Psychologically, having a self-employed parent may make that occupational choice seem more viable and prestigious. Socially, there may be a transfer of social capital as the parent exposes an individual to networks and connections. Expected earnings are typically considered to be a function of human capital and, also for the self-employed specifically, endowments of managerial ability (Boden, 1996). As self-employed parents may pass down skills or managerial ability then the expected earnings of becoming self-employed would increase. Finally, there may be an inter-generational transfer of a business. The employment status of parents may influence an individual’s employment status but not wealth beyond that and hence be a suitable instrument; this will be returned to in the analysis section.

employed to save more, as their incomes are more variable” (Leland, 1968, p.471).

Whilst there is evidence that the self-employed have greater volatility of incomes (Jensen & Shore, 2008), there is less consensus over whether this translates into greater saving. Skinner (1988) finds that the self-employed save less than those with other employment statuses, whilst Kazarosian’s (1997) panel study finds a positive relationship between saving and self-employment consistent with the precautionary motive narrative. Whilst both acknowledge potential endogeneity, neither satisfactorily account for this; therefore, the disparity in results could result from selection bias. This selection bias has not been adequately addressed in economic literature and as such this paper hopes to fill a gap in the literature and offer methods that could be used to improve the reliability of the results attained in future research.

In order to create an unbiased estimate of the effect of self-employment on wealth due to unobservable characteristics a suitable instrument needs to be found. Rees and Shah (1986) undergo a comprehensive study of the determinants of self-employment. However, many particularly significant determinants, such as education and age, are not suitable instruments as they are almost certain to have a causal effect on wealth.

It may therefore be useful to think beyond socio-economic factors. Environmental, psychological or social capital factors may influence these unobservable factors, as discussed by Sørensen (2007). Furthermore, genetic factors may be important, as discussed by Nicolaou et al. (2008). A number of studies, including Henley (2004) and Hout and Rosen (2000) find that the employment status of parents is significant, and this may work through the factors previously outlined. Parents may pass on genes which influence personality traits and dispositions to certain employment types. Psychologically, having a self-employed parent may make that occupational choice seem more viable and prestigious. Socially, there may be a transfer of social capital as the parent exposes an individual to networks and connections. Expected earnings are typically considered to be a function of human capital and, also for the self-employed specifically, endowments of managerial ability (Boden, 1996). As self-employed parents may pass down skills or managerial ability then the expected earnings of becoming self-employed would increase. Finally, there may be an inter-generational transfer of a business. The employment status of parents may influence an individual’s employment status but not wealth beyond that and hence be a suitable instrument; this will be returned to in the analysis section.
The Wealth and Assets Survey (WAS) is a representative longitudinal data set covering Great Britain that at the time of writing had two waves. In this paper, only wave 2 data, which was collected between July 2008 and June 2010, will be considered, making this research cross-sectional. WAS was selected for this analysis due to its exhaustive collection of household and individual wealth data. The dependent variable in the model therefore will be the log of gross wealth, with wealth defined as an individual’s aggregate pension and financial wealth. Most studies which try to measure precautionary saving use some measure of saving as the dependent variable, such as the difference between income and consumption or the change in wealth over time periods. Considering the self-employed may have a good year followed by a bad one, or good/bad runs of years, using such flow variables may not be appropriate within the context of focusing on the self-employed. This provides justification for using the stock variable wealth. An individual is deemed to be self-employed in WAS if they earn “income directly from their own business, trade or profession” (ONS, 2010a).

Table 1 presents summary statistics from WAS wave 2, broken down by employment status. There are 19,074 individuals with either of these employment statuses and 2,599 of these are self-employed meaning 13.6% of the sample is self-employed, a similar figure found in the much larger Labour Force Survey between 2008 and 2010 (ONS, 2010b) - suggesting the sample may be representative of the population as a whole. It can also be seen that significantly more men than women are self-employed and those with higher educations are more likely to be self-employed compared to low education types, matching the results of Rees and Shah (1986).
Table 1. Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Employee</th>
<th>Self-employed</th>
<th>Total</th>
<th>Employee</th>
<th>Self-employed</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>No.</td>
<td>No.</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Total</td>
<td>16475</td>
<td>2599</td>
<td>19074</td>
<td>86.4</td>
<td>13.6</td>
<td>100</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>8021</td>
<td>1762</td>
<td>9783</td>
<td>48.7</td>
<td>67.8</td>
<td>51.3</td>
</tr>
<tr>
<td>Female</td>
<td>8454</td>
<td>837</td>
<td>9291</td>
<td>51.3</td>
<td>32.2</td>
<td>48.7</td>
</tr>
<tr>
<td>Age bands</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-29</td>
<td>1364</td>
<td>98</td>
<td>1462</td>
<td>8.3</td>
<td>3.8</td>
<td>7.7</td>
</tr>
<tr>
<td>30-34</td>
<td>1597</td>
<td>165</td>
<td>1762</td>
<td>9.7</td>
<td>6.3</td>
<td>9.2</td>
</tr>
<tr>
<td>35-39</td>
<td>2009</td>
<td>237</td>
<td>2246</td>
<td>12.2</td>
<td>9.1</td>
<td>11.8</td>
</tr>
<tr>
<td>40-44</td>
<td>2374</td>
<td>360</td>
<td>2734</td>
<td>14.4</td>
<td>13.9</td>
<td>14.3</td>
</tr>
<tr>
<td>45-49</td>
<td>2262</td>
<td>385</td>
<td>2647</td>
<td>13.7</td>
<td>14.8</td>
<td>13.9</td>
</tr>
<tr>
<td>50-54</td>
<td>2025</td>
<td>318</td>
<td>2343</td>
<td>12.3</td>
<td>12.2</td>
<td>12.3</td>
</tr>
<tr>
<td>55-59</td>
<td>1805</td>
<td>357</td>
<td>2162</td>
<td>11</td>
<td>13.7</td>
<td>11.3</td>
</tr>
<tr>
<td>60-64</td>
<td>1185</td>
<td>336</td>
<td>1521</td>
<td>7.2</td>
<td>12.9</td>
<td>8</td>
</tr>
<tr>
<td>Education Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No qualifications</td>
<td>1504</td>
<td>313</td>
<td>1817</td>
<td>9.1</td>
<td>12</td>
<td>9.5</td>
</tr>
<tr>
<td>Other qualifications</td>
<td>10008</td>
<td>1485</td>
<td>11493</td>
<td>60.7</td>
<td>57.1</td>
<td>60.3</td>
</tr>
<tr>
<td>Degree or above</td>
<td>4953</td>
<td>799</td>
<td>5752</td>
<td>30.1</td>
<td>30.7</td>
<td>30.2</td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>4628</td>
<td>440</td>
<td>5068</td>
<td>28.1</td>
<td>16.9</td>
<td>26.6</td>
</tr>
<tr>
<td>Married or in civil partnership</td>
<td>9636</td>
<td>1799</td>
<td>11435</td>
<td>58.5</td>
<td>69.2</td>
<td>60</td>
</tr>
<tr>
<td>Separated or</td>
<td>1945</td>
<td>305</td>
<td>2250</td>
<td>11.8</td>
<td>11.7</td>
<td>11.8</td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>4628</td>
<td>440</td>
<td>5068</td>
<td>28.1</td>
<td>16.9</td>
<td>26.6</td>
</tr>
<tr>
<td>Married or in civil partnership</td>
<td>9636</td>
<td>1799</td>
<td>11435</td>
<td>58.5</td>
<td>69.2</td>
<td>60</td>
</tr>
<tr>
<td>Separated or</td>
<td>1945</td>
<td>305</td>
<td>2250</td>
<td>11.8</td>
<td>11.7</td>
<td>11.8</td>
</tr>
</tbody>
</table>
A particular novelty of the WAS data is that it attempts to provide a measure of risk attitudes. This comes from a question asking respondents whether they would prefer a guaranteed £1000 or a 1 in 5 chance of winning £10,000. It is possible the responses to this survey question are related to income, as per Pratt’s (1964) measure of risk aversion which suggests risky investment is not an inferior good. However, it may still serve as a useful measure of risk aversion, with those who selected the 1 in 5 chance considered less risk averse and those the guaranteed payment, more risk averse. A number of studies have noted that self-employment is a more risky occupational choice (such as Sheshinski & Dréze, 1976) as well as that the self-employed have lower levels of risk aversion (see Friedman, 1957 and Skinner, 1988). It is encouraging therefore to find in the WAS data that the self-employed are over 5 percentage points more likely to choose the riskier 1 in 5 chance. This variable could allow for the typically unobservable characteristic of risk aversion to be directly measured in regression modelling.

4 Methodology

A standard OLS regression would look something like this:

\[ \ln \text{Wealth}_i = \alpha X_i + b \delta E_i + \mu_i \]
Where $X_i$ is a vector of exogenous variables and $SE_i$ is a dummy equalling one if the individual is self-employed. However, self-selection in self-employment would mean that:

$$\text{Prob}(SE = 1) = cA_i + v_i$$

Where $A_i$ is a vector of unobservable characteristics that, if observed, we would be happy to include in the regression. Not including it places it in the error term, so that:

$$\text{Cov}(SE_i, \mu_i) \neq 0$$

And hence:

$$b = \frac{dy}{dx} = \beta + \frac{du}{dx}$$

Where $\beta$ is the true effect of being self-employed. Unobservable characteristics are therefore creating an omitted variable bias and violating an OLS assumption. What is required is a variable, termed an instrument ($Z_i$), picking up these unobservable effects. It should correlated with the casual variable of interest, SE, but not any other determinant of the dependent variable, such that:

$$\text{Cov}(Z_i, \mu_i) = 0$$

$$\text{Cov}(Z_i, SE_i) \neq 0$$

Econometrically therefore it is vital to find a variable correlated with self-employment but which has no other influencing factor on an individual’s stock of wealth. It should also have a theoretical underpinning which provides a reason why it is an effective instrument.

A two-stage least squares equation uses the instrument, as well as other exogenous variables, in the first stage to provide estimates of SE. These predicted values are then used in the second-stage to provide unbiased coefficients:

$$SE_i = dZ_i + eX_i + \epsilon_i$$

$$\ln Wealth_i = fX_i + gSE_i + \mu_i$$

The Heckman selection model (Heckman, 1979) also deals with selection bias as a form of omitted variable bias. Here, the dependent variable in the outcome equation ($\ln Wealth$) is observed if and only if a latent
variable, $SE^*$, which measures the propensity to be self-employed, exceeds a particular threshold:

$$SE_i = \begin{cases} 1 & \text{if } SE > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$Pr(SE_i = 1) = \Phi(a'w_i) = \text{selection equation}$$

Where $\alpha$ and $w$ are the vectors of coefficients and covariates (including both $X_i$ and $Z_i$) for the selection equation respectively. For Heckman, an estimate of the omitted variable would overcome the problem of selection bias, and this estimate comes from computing the inverse Mill’s ratio ($\lambda$): the ratio of the probability density function ($\Phi$) over the cumulative distribution function of a distribution. A strong assumption that needs to be made for Heckman to be applicable is the error terms of the selection and outcomes equation must be normally distributed.

5 Analysis

When a set of OLS regressions are run in order to check the coefficient of the self-employed dummy, allowing it to be compared with the results when self-selection is accounted for, it is found that the self-employment dummy varies between -0.84 and -0.93 (not shown). The introduction of additional variables improves the model’s adjusted $R^2$ and increases the negativity of the self-employment dummy.

In the literature review section, it was discussed that there may be a relation between employment status of the parent and the individuals. Table 2 confirms a correlation exists:

<table>
<thead>
<tr>
<th>Table 2. Cross-tabulation between employment status and father’s employment status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment status</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Employee</td>
</tr>
<tr>
<td>Self-employed</td>
</tr>
</tbody>
</table>

Note: other possible employment statuses of the father show no evidence of a correlation with the individual’s employment status and have not been reported.

variable, $SE^*$, which measures the propensity to be self-employed, exceeds a particular threshold:

$$SE_i = \begin{cases} 1 & \text{if } SE > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$Pr(SE_i = 1) = \Phi(a'w_i) = \text{selection equation}$$

Where $\alpha$ and $w$ are the vectors of coefficients and covariates (including both $X_i$ and $Z_i$) for the selection equation respectively. For Heckman, an estimate of the omitted variable would overcome the problem of selection bias, and this estimate comes from computing the inverse Mill’s ratio ($\lambda$): the ratio of the probability density function ($\Phi$) over the cumulative distribution function of a distribution. A strong assumption that needs to be made for Heckman to be applicable is the error terms of the selection and outcomes equation must be normally distributed.

5 Analysis

When a set of OLS regressions are run in order to check the coefficient of the self-employed dummy, allowing it to be compared with the results when self-selection is accounted for, it is found that the self-employment dummy varies between -0.84 and -0.93 (not shown). The introduction of additional variables improves the model’s adjusted $R^2$ and increases the negativity of the self-employment dummy.

In the literature review section, it was discussed that there may be a relation between employment status of the parent and the individuals. Table 2 confirms a correlation exists:

<table>
<thead>
<tr>
<th>Table 2. Cross-tabulation between employment status and father’s employment status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment status</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Employee</td>
</tr>
<tr>
<td>Self-employed</td>
</tr>
</tbody>
</table>

Note: other possible employment statuses of the father show no evidence of a correlation with the individual’s employment status and have not been reported.
The self-employed are 8.4 percentage points more likely to have a self-employed father than employees. Interestingly, the effect isn’t stronger for fathers and sons than fathers and daughters (not shown). A number of other potential instruments have been considered, including mother’s employment status, father’s education and parental household type. When adding these to the OLS regressions conducted earlier it is found that all except father’s employment status is significant, and when an F-test is run on that sub-set of coefficients the F-value is 1.02, meaning the null cannot be rejected. Resultantly, it is concluded that the most suitable instrument is the employment status of the father, and this will be the sole instrument used in the two-stage least squares process and Heckman modelling.

In the below regressions, the discarded ‘instruments’ of mother’s employment status, parent’s household tenancy and father’s education are included in the specification due to their apparent significance as noted previously.

Table 3. Truncated two-stage least squares specifications. Dependent variable: lnWealth

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-employment dummy</td>
<td>-1.78*</td>
<td>-1.87**</td>
<td>-1.74*</td>
<td>-1.85**</td>
</tr>
<tr>
<td>Age bands</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-34</td>
<td>0.90***</td>
<td>0.82***</td>
<td>0.82***</td>
<td>0.78***</td>
</tr>
<tr>
<td>35-39</td>
<td>1.55***</td>
<td>1.45***</td>
<td>1.47***</td>
<td>1.42***</td>
</tr>
<tr>
<td>40-44</td>
<td>2.35***</td>
<td>2.25***</td>
<td>2.29***</td>
<td>2.25***</td>
</tr>
<tr>
<td>45-49</td>
<td>2.84***</td>
<td>2.73***</td>
<td>2.78***</td>
<td>2.73***</td>
</tr>
<tr>
<td>50-54</td>
<td>3.20***</td>
<td>3.08***</td>
<td>3.14***</td>
<td>3.10***</td>
</tr>
<tr>
<td>55-59</td>
<td>3.67***</td>
<td>3.56***</td>
<td>3.65***</td>
<td>3.61***</td>
</tr>
<tr>
<td>60-64</td>
<td>4.24***</td>
<td>4.11***</td>
<td>4.20***</td>
<td>4.15***</td>
</tr>
<tr>
<td>Female gender dummy</td>
<td>-0.50***</td>
<td>-0.51***</td>
<td>-0.51***</td>
<td>-0.50***</td>
</tr>
</tbody>
</table>

2 However, in the interest of brevity they have not been reported. The two-stage least squares regressions were also run without including these additional variables with little change to the self-employed dummy coefficient.
The results from Table 3 are stark. The coefficient for being self-employed is quantitatively very large, statistically significant between the 0.05 and 0.1 levels and fairly robust to specification changes. This suggests unobservable characteristics that influence selection into self-employment and have a major influence on the wealth holdings of the self-employed. Solely according to these results, it seems apparent that instrumenting for it paternal employment is critically important in the context of the saving habits of the self-employed and should compel future research to take account of it.

When the Heckman model is ran the following estimates are obtained:

### Educational level

<table>
<thead>
<tr>
<th>Other qualifications</th>
<th>1.20***</th>
<th>1.17***</th>
<th>1.15***</th>
<th>1.15***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree or above</td>
<td>2.30***</td>
<td>2.27***</td>
<td>2.21***</td>
<td>2.20***</td>
</tr>
</tbody>
</table>

### Marital status

<table>
<thead>
<tr>
<th>Married or in civil partnership</th>
<th>0.31***</th>
<th>0.29***</th>
<th>0.30***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separated or divorced</td>
<td>-0.00072</td>
<td>0.0042</td>
<td>0.011</td>
</tr>
<tr>
<td>Widowed</td>
<td>-0.0045</td>
<td>-0.001</td>
<td>-0.015</td>
</tr>
</tbody>
</table>

### Health status

<table>
<thead>
<tr>
<th>Good</th>
<th>-0.18***</th>
<th>-0.18***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fair</td>
<td>-0.56***</td>
<td>-0.56***</td>
</tr>
<tr>
<td>Bad</td>
<td>-0.80***</td>
<td>-0.81***</td>
</tr>
<tr>
<td>Less risk averse dummy</td>
<td>0.31***</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Constant</th>
<th>6.92***</th>
<th>6.85***</th>
<th>7.00***</th>
<th>7.01***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>11143</td>
<td>11143</td>
<td>11143</td>
<td>10943</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.242</td>
<td>0.242</td>
<td>0.251</td>
<td>0.246</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1
Table 4. Truncated Heckman selection model specifications. Dependent variable: lnWealth

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome Equation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age bands</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-34</td>
<td>0.87***</td>
<td>0.88***</td>
<td>0.90**</td>
<td>0.78*</td>
</tr>
<tr>
<td>35-39</td>
<td>1.43***</td>
<td>1.45***</td>
<td>1.63***</td>
<td>1.60***</td>
</tr>
<tr>
<td>40-44</td>
<td>2.27***</td>
<td>2.29***</td>
<td>2.39***</td>
<td>2.38***</td>
</tr>
<tr>
<td>45-49</td>
<td>3.05***</td>
<td>3.07***</td>
<td>3.30***</td>
<td>3.28***</td>
</tr>
<tr>
<td>50-54</td>
<td>3.49***</td>
<td>3.51***</td>
<td>3.68***</td>
<td>3.67***</td>
</tr>
<tr>
<td>55-59</td>
<td>3.86***</td>
<td>3.88***</td>
<td>4.13***</td>
<td>4.13***</td>
</tr>
<tr>
<td>60-64</td>
<td>4.51***</td>
<td>4.51***</td>
<td>4.91***</td>
<td>4.86***</td>
</tr>
<tr>
<td>Female gender dummy</td>
<td>-0.50***</td>
<td>-0.51***</td>
<td>-0.56***</td>
<td>-0.55***</td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other qualifications</td>
<td>0.71***</td>
<td>0.74***</td>
<td>0.78***</td>
<td>0.78***</td>
</tr>
<tr>
<td>Degree or above</td>
<td>1.59***</td>
<td>1.60***</td>
<td>1.62***</td>
<td>1.61***</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married or in civil partnership</td>
<td>-0.0005</td>
<td>-0.079</td>
<td>-0.08</td>
<td></td>
</tr>
<tr>
<td>Separated or divorced</td>
<td>-0.16</td>
<td>-0.22</td>
<td>-0.24</td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>0.71</td>
<td>0.64</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Health status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>-0.039</td>
<td>-0.047</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair</td>
<td>-0.27</td>
<td>-0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bad</td>
<td>-0.55</td>
<td>-0.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less risk averse dummy</td>
<td>0.36***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Selection Equation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Truncated Heckman selection model specifications. Dependent variable: lnWealth

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome Equation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age bands</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-34</td>
<td>0.87***</td>
<td>0.88***</td>
<td>0.90**</td>
<td>0.78*</td>
</tr>
<tr>
<td>35-39</td>
<td>1.43***</td>
<td>1.45***</td>
<td>1.63***</td>
<td>1.60***</td>
</tr>
<tr>
<td>40-44</td>
<td>2.27***</td>
<td>2.29***</td>
<td>2.39***</td>
<td>2.38***</td>
</tr>
<tr>
<td>45-49</td>
<td>3.05***</td>
<td>3.07***</td>
<td>3.30***</td>
<td>3.28***</td>
</tr>
<tr>
<td>50-54</td>
<td>3.49***</td>
<td>3.51***</td>
<td>3.68***</td>
<td>3.67***</td>
</tr>
<tr>
<td>55-59</td>
<td>3.86***</td>
<td>3.88***</td>
<td>4.13***</td>
<td>4.13***</td>
</tr>
<tr>
<td>60-64</td>
<td>4.51***</td>
<td>4.51***</td>
<td>4.91***</td>
<td>4.86***</td>
</tr>
<tr>
<td>Female gender dummy</td>
<td>-0.50***</td>
<td>-0.51***</td>
<td>-0.56***</td>
<td>-0.55***</td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other qualifications</td>
<td>0.71***</td>
<td>0.74***</td>
<td>0.78***</td>
<td>0.78***</td>
</tr>
<tr>
<td>Degree or above</td>
<td>1.59***</td>
<td>1.60***</td>
<td>1.62***</td>
<td>1.61***</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married or in civil partnership</td>
<td>-0.0005</td>
<td>-0.079</td>
<td>-0.08</td>
<td></td>
</tr>
<tr>
<td>Separated or divorced</td>
<td>-0.16</td>
<td>-0.22</td>
<td>-0.24</td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>0.71</td>
<td>0.64</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Health status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>-0.039</td>
<td>-0.047</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair</td>
<td>-0.27</td>
<td>-0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bad</td>
<td>-0.55</td>
<td>-0.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less risk averse dummy</td>
<td>0.36***</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Father self-employed
Father unemployed
Father other

\[ \lambda \] which provides an estimate of the unobserved characteristics influencing self-selection into self-employment, is negative and quantitatively large. However, it is not statistically significant, and tends towards zero as the number of variables increase.

It should be noted at this point that there may be a number of other reasons also depressing wealth holdings of the self-employed. For example, the self-employed may be expecting to continue work after reaching retirement age and hence save less for retirement. Furthermore, it has been noted that self-employment rates have increased in Great Britain since the beginning of the 2008-2009 financial crisis (Goodall, 2014). It is possible therefore that some of the self-employed in the 2008-2010 wave are newly self-employed who, potentially for a number of reasons, have lower stocks of wealth. Finally, the self-employed may have invested some of their wealth in a business, which is not fully reflected in the WAS data set.

Furthermore, that \( \lambda \) is negative and quantitatively large but not statistically significant could suggest the impact of these unobserved characteristics may be heterogeneous – affecting some self-employed individuals but not others. A heterogeneous effects framework may therefore be necessary. For example, if having a self-employed father has an effect mostly on the first sibling then the estimates would be relevant to that group. Related to this, it may be useful to interact self-employment with occupational sector – self-employed farmers, for example, may have very different characteristics and

### Observations

<table>
<thead>
<tr>
<th>Father self-employed</th>
<th>0.33***</th>
<th>0.33***</th>
<th>0.32***</th>
<th>0.32***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Father unemployed</td>
<td>0.048</td>
<td>0.048</td>
<td>0.0019</td>
<td>0.011</td>
</tr>
<tr>
<td>Father other</td>
<td>-0.11</td>
<td>-0.11</td>
<td>-0.12</td>
<td>-0.11</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>-0.77</td>
<td>-0.73</td>
<td>-0.62</td>
<td>-0.43</td>
</tr>
<tr>
<td>Constant (outcome)</td>
<td>7.66***</td>
<td>7.57***</td>
<td>7.31***</td>
<td>6.91***</td>
</tr>
<tr>
<td>Constant (selection)</td>
<td>-1.17***</td>
<td>-1.17***</td>
<td>-1.18***</td>
<td>-1.19***</td>
</tr>
<tr>
<td>Observations</td>
<td>13258</td>
<td>13258</td>
<td>11275</td>
<td>11252</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1.
saving patterns from self-employed lawyers. These above points show that there are clearly a number of areas for future research.

Nevertheless, despite the limitations listed above, considering there is some evidence that unobservable characteristics have a negative effect on the wealth holdings of the self-employed, we should return to why this may be the case. As has been already discussed, the self-employed are typically thought of as having lower levels of risk aversion. However, in Table 3 the less risk averse dummy is both positive and does not dampen the coefficient of the self-employed dummy. Whilst it is positive again in the Heckman model, the inclusion of the dummy in Table 4 Column 4 reduces the value of $\lambda$ from -0.62 to -0.43.

An alternative reason is that the self-employed may have different time preferences to employees. It has been suggested the self-employed are especially optimistic by Arabsheibani et al. (2000) Cooper et al. (1988). It seems possible that this over-optimism among the self-employed may in turn lead to a higher discount rate among the self-employed as they have a greater sense that ‘tomorrow will take care of itself’. These traits could be related to the instrument of father’s employment status through genetic and environmental factors. As far as known, that the self-employed have a higher discount rate in the mean has not been empirically tested, but such a narrative would fit with the results obtained here and research into this is encouraged.

Overall, the results of this paper have a number of implications. The apparent importance of unobservable characteristics shows they should be taken seriously and accounted for in any future modelling. It also goes against the conventional wisdom that the self-employed save more because of the precautionary demand motive for saving – their unobservable characteristics seem to result in lower levels of saving, perhaps due to a higher discount rate. Economists, therefore, may need to reconsider assumptions about the saving habits of the self-employed.

There may also be a number of policy implications. One of the motives for social security is a paternalistic concern that individuals may not act rationally or with full information when making inter-temporal decisions, leading to sub-optimality. Whether that is strictly true or not is beyond the scope of this paper, but if it was felt this was the case there could be a case for government intervention to ‘nudge’ the self-employed into saving more. It is also calls into the question the current state pension system in the UK. The self-employed, through National Insurance contributions can earn entitlements to the Basic State Pension but, unlike employees, cannot thereby save less due to the precautionary motive.
accumulate a State Second Pension on top of this. If the self-employed are
saving less anyway, the government deciding to provide them with a lower
state pension on average may appear unwise. Furthermore, auto-enrolment
does not apply to the self-employed.

6 Conclusion

There has been a tendency in economics literature to emphasise the
importance of precautionary saving for the self-employed, and whilst this
may be an important factor, the role of unobservable characteristics and self-
selection into self-employment has often been neglected in modelling,
leading to potential bias. This paper makes an initial attempt to account for
the size of these unobservable characteristics. In both models estimated, their
effects are found to be very large, but they are not statistically significant in
the Heckman selection model. It may be the case that a heterogeneous effects
framework is more appropriate therefore. As such, research which builds on
this paper is encouraged – it is fair to say this paper is valuable in terms of
originally combining two sets of economic literature, on savings and
determinants of self-employment, in order to try and measure the impact of
unobservable characteristics affecting selection into self-employment and
overcome endogeneity, but should be viewed as an initial starting point for
research into this area. Yet based on the results obtained here unobservable
characteristics seem important, possibly due to their impact on risk aversion
levels or discount rates, suggesting endogeneity bias may be plaguing
research which does no account for this. As such, the results obtained here
challenge the conventional story of greater savings due to precautionary motives and encourages future research to have this at the forefront of its
mind. Furthermore, it raises questions about the optimality of saving
decisions among the self-employed and hence could influence policy.

3 It should be noted that from April 2016 a single-tier pension will be introduced in the
UK which will equalise the amount of state pension employees and self-employed can
accrue.

4 Auto-enrolment is currently being introduced incrementally into the UK and means
employees are automatically enrolled in an occupational pension scheme unless they
actively decide to opt-out. The aim is to increase the number of people with a non-state pension.
Bibliography


A MECHANISM DESIGN APPROACH TO INCREASING
THE EFFICIENCY OF AID DISTRIBUTION

MATTHEW KLEIN
University of Chicago

ABSTRACT

This paper brings to light evidence that development aid is, and historically has been, distributed inefficiently. The critical friction, discussed at length by anthropologists, is an incentive misalignment: multilateral aid agencies face a trade-off between efficiently distributing aid to the poor and allocating resources to themselves. Incentives are aligned so that development agencies, by maximizing utility, necessarily decrease the efficiency with which aid is distributed. By analyzing the work of two eminent anthropologists, James Ferguson and David Mosse, I demonstrate that this friction exists and that a simple mechanism can realign incentives to realize an improvement in efficiency. I then adopt Holmstrom and Tirole's (1997) model of outside financing capacity to frame my mechanism design problem. After developing the Groves scheme, I revisit Ferguson's work to explain how this incentive misalignment negatively impacts the efficiency of aid distribution and how much the developed mechanism could have increased the welfare of those receiving aid in one particular case. The goals of my project, then, are twofold: to produce a long sought after mechanism that may positively impact the efficiency of aid distribution, and to bridge two literatures in order to increase the rate of theoretical innovation. There are outstanding questions that merit further research, two of which are as follows: To what extent is aid distribution inefficiency caused by this misalignment and what does implementation look like today?
THE CHALLENGE OF NATIONAL TREATMENT IN INTERNATIONAL TRADE: EVIDENCE FROM WTO DISPUTE SETTLEMENT

HEATHER HEDGES

Georgetown University

ABSTRACT

The principle of national treatment, enshrined in GATT Article III, is a central pillar of the multilateral trading system established by the World Trade Organization (WTO). National treatment has also become a basic principle of bilateral and multilateral preferential trade agreements (PTAs), underscoring its importance at the core of international trade law. Interpretation of national treatment obligations extracted from the textual details of GATT Article III, however, is fraught with inconsistency and unpredictability; specifically, scholars argue the differences in interpretation with regard to Article III:2 and Article III:4 and the way this affects the likelihood of favorable outcomes. Within the WTO Dispute Settlement Body, decisions rendered by Panels and the Appellate Body offer no recourse for such ambiguity, meaning variations of GATT Article III citations continue to dominate requests for consultations. Using the comprehensive dataset created by Horn and Mavroidis, I investigate the variation throughout national treatment cases at the WTO and the relationship between citations invoked at the consultation and Panel request states with the outcome of the case. There is significant evidence that Article III:4 claims are less likely to yield favorable rulings at the Panel stage than Article III:2 claims. Considering the many national treatment cases lingering in the consultation stage of dispute settlement, member states may be concerned with the prospects for the eventual outcomes of their cases.
1 Introduction

Formation of the World Trade Organization (WTO) in 1995 brought Member states into a global trading order in which implementation of major tariff concessions and other trade liberalizing policies awarded with access to an integrated network of free trading partners. While the General Agreement on Tariffs and Trade (GATT) – the free-standing agreement that had previously governed international trade – became just one of many framework agreements of the WTO, two of its key principles facilitating trade without discrimination remain the underpinnings of most if not all trade agreements negotiated after its adoption: most-favored-nation and national treatment.1

1.1 A Closer Look at GATT Article III

Found in GATT Article III, national treatment instead works within borders, handling all treatment of foreign products once they have entered a local market in order to ensure that imported and domestically produced products are afforded the same treatment under local law.2 This provision of the GATT is critically important because it facilitates the aspiration of the WTO to lower barriers in order to encourage trade, even after wide-ranging tariff concessions have been granted. Specifically, countries’ failure to adhere to national treatment obligations could negate concessions achieved with regard to tariffs at the border by imposing discriminatory taxation or regulation schemes on foreign versus domestic products.3 According to Horn (2006), national treatment provides an essential “discipline on the use of internal measures” without which international agreements on tariffs would be “meaningless.”4

The first, second, and fourth paragraphs of Article III – referred hereafter as Article III:1, III:2, and III:4, respectively – are arguably the most important provisions within Article III. It is confirmed throughout WTO jurisprudence and academic literature that the Article III:1 is essential for the interpretation of both Articles III:2 and III:4 in relevant cases. In fact, Hudec identifies Article III:1 as the “starting point” for understanding the national treatment

2 Ibid.

1 Introduction

Formation of the World Trade Organization (WTO) in 1995 brought Member states into a global trading order in which implementation of major tariff concessions and other trade liberalizing policies awarded with access to an integrated network of free trading partners. While the General Agreement on Tariffs and Trade (GATT) – the free-standing agreement that had previously governed international trade – became just one of many framework agreements of the WTO, two of its key principles facilitating trade without discrimination remain the underpinnings of most if not all trade agreements negotiated after its adoption: most-favored-nation and national treatment.1

1.1 A Closer Look at GATT Article III

Found in GATT Article III, national treatment instead works within borders, handling all treatment of foreign products once they have entered a local market in order to ensure that imported and domestically produced products are afforded the same treatment under local law.2 This provision of the GATT is critically important because it facilitates the aspiration of the WTO to lower barriers in order to encourage trade, even after wide-ranging tariff concessions have been granted. Specifically, countries’ failure to adhere to national treatment obligations could negate concessions achieved with regard to tariffs at the border by imposing discriminatory taxation or regulation schemes on foreign versus domestic products.3 According to Horn (2006), national treatment provides an essential “discipline on the use of internal measures” without which international agreements on tariffs would be “meaningless.”4

The first, second, and fourth paragraphs of Article III – referred hereafter as Article III:1, III:2, and III:4, respectively – are arguably the most important provisions within Article III. It is confirmed throughout WTO jurisprudence and academic literature that the Article III:1 is essential for the interpretation of both Articles III:2 and III:4 in relevant cases. In fact, Hudec identifies Article III:1 as the “starting point” for understanding the national treatment

2 Ibid.
policy behind the subsequent provisions of Article III such that the broad objectives of national treatment are set out in this opening paragraph. Such an objective set forth in this opening paragraph is to prevent states from using internal measures in order to protect domestic production and to “ensure the equality of competitive conditions” between imported and domestic products. The phrase “so as to afford protection” is the critically important concept established in Article III:1 such that any referencing to Article III:1 by subsequent portions of the article will invoke this idea of internal measures being applied with protectionist intent. Moreover, WTO jurisprudence has illustrated the idea that national treatment obligations exist in order to prevent Members from underpinning broader commitments to international trade including tariff concessions by backfilling with internal taxes and other measures.

Subsequently, paragraphs two and four of Article III are of critical importance to understanding national treatment obligations at the WTO. The primary difference between each of these provisions is the kind of internal measures addressed in each paragraph; Article III:2 applies to internal taxes and charges while Article III:4 addresses all other internal measures through which members might achieve discrimination between like domestic and imported products. While the violation of national treatment according to each part of Article III is seemingly identical, such that imported products are afforded less favorable treatment than domestic products, the comparison of tax and non-tax measures at issue are the key difference difference between claims that refer to Articles III:2 versus Article III:4.

1.2 The Challenge of Interpretation

Despite the obvious charge and non-charge differences between Articles III:2 and III:4, the difference in practical implementation according to each of the two components of Article III is the subject of the enormous literary debate regarding the interpretation taken in dispute settlement. First examining Article III:2, there are two important characteristics of this provision that have direct implications on its interpretation with regard to dispute outcomes. Firstly, the literature identifies two separate sentences of

7 WTO Document WT/DS186/R, para. 5.10.

policy behind the subsequent provisions of Article III such that the broad objectives of national treatment are set out in this opening paragraph. Such an objective set forth in this opening paragraph is to prevent states from using internal measures in order to protect domestic production and to “ensure the equality of competitive conditions” between imported and domestic products. The phrase “so as to afford protection” is the critically important concept established in Article III:1 such that any referencing to Article III:1 by subsequent portions of the article will invoke this idea of internal measures being applied with protectionist intent. Moreover, WTO jurisprudence has illustrated the idea that national treatment obligations exist in order to prevent Members from underpinning broader commitments to international trade including tariff concessions by backfilling with internal taxes and other measures.

Subsequently, paragraphs two and four of Article III are of critical importance to understanding national treatment obligations at the WTO. The primary difference between each of these provisions is the kind of internal measures addressed in each paragraph; Article III:2 applies to internal taxes and charges while Article III:4 addresses all other internal measures through which members might achieve discrimination between like domestic and imported products. While the violation of national treatment according to each part of Article III is seemingly identical, such that imported products are afforded less favorable treatment than domestic products, the comparison of tax and non-tax measures at issue are the key difference difference between claims that refer to Articles III:2 versus Article III:4.

1.2 The Challenge of Interpretation

Despite the obvious charge and non-charge differences between Articles III:2 and III:4, the difference in practical implementation according to each of the two components of Article III is the subject of the enormous literary debate regarding the interpretation taken in dispute settlement. First examining Article III:2, there are two important characteristics of this provision that have direct implications on its interpretation with regard to dispute outcomes. Firstly, the literature identifies two separate sentences of
Article III:2: the first sentence highlighting “likeness” between goods and the second sentence introducing the idea of similar goods being “directly competitive or substitutable.” In this paragraph, these two separate ideas of “like” and “directly competitive or substitutable” goods have vastly different meanings, and therefore distinct applications with regard to national treatment obligations and violations. While there are a variety of characteristics used in order to determine whether two similar imported and domestic goods satisfy the criteria for being “like” or “directly competitive or substitutable” with one another, many of which are also subject to intense debate, WTO jurisprudence tells us that the physical characteristics of goods are critical in making a determination of likeness.

This distinction between two categories of similar goods has severe implications on finding Article III violations in dispute settlement. Within the category of “like” goods enumerated in the first sentence, there is no reference made to Article III:1. Conversely, the second sentence establishing “directly competitive or substitutable” products specifically invokes the language of Article III:1 through explicit reference. This difference between the two means that in evaluating whether internal measures constitute violations of Article III:2, second sentence, there is an evident need to prove that internal taxes and charges are applied “so as to afford protection.”

However, first sentence notably lacks any reference to this phrase in Article III:1, meaning there is no need to prove a protective application of any measure. In the second sentence, the “directly competitive or substitutable” criteria may be broader, however the burden to prove a protective application of internal measures is higher and thus might also constrain the likelihood of measures being found in violation.

Now, how do these interpretations of national treatment under Article III:2 relate to Article III:4? This requires a textual analysis of the fourth paragraph, in which only the term “like product” appears. Indeed, there is no mention of any competitive or substitutable relationship between goods. At the surface level, it seems that the concept of “likeness” should be consistent between Article III:2 and III:4, with the exception that one paragraph address charge measures and the other addresses non-charge measures. However, WTO jurisprudence explains that there is no continuity of interpretation across these provisions of Article III. The failure for such jurisprudence to offer clear interpretations of the distinct obligations conferred by Articles

---

10 WTO Document WT/DS8/AB/R
III:2 and III:4 throughout the wide body of national treatment disputes results in nothing more than confusion for Members and unpredictability throughout the system.

## 2 Literature Review

### 2.1 The Challenge of Article III in Practice

Looking outside WTO Panel and Appellate Body rulings, a vast body of literature generates consensus regarding the challenge of imprecise interpretations of GATT Article III and the provisions enshrined within it. While there is a wealth of literature offering qualitative analyses of the conundrum created by varied Article III textual interpretations, quantitative investigations of national treatment issues in international trade are limited. Economic analyses of the WTO dispute settlement process are similarly limited in number. Saggi and Sara (2008) highlight the challenge posed by the concept of “likeliness,” which presents perhaps the greatest obstacle in attempting to identify a clear interpretation of national treatment obligations. The authors argue that differentiation of otherwise homogenous goods caused by differences in country of origin and other factors makes the determination of “like products” according to Articles III:2 and III:4 extremely difficult to predict.

Zhou (2012) further wrestles with this idea of ambiguous interpretation from a qualitative perspective, arguing that common interpretation of international legal texts fails to provide an accurate understanding of the purpose of Article III, which necessarily informs the understanding of Articles III:2 and III:4. Importantly, Zhou contributes to the literature arguing that the specific textual differences between Article III:2 and Article III:4 – including the phrases “like,” “directly competitive or substitutable,” “in excess of,” “aim and effect,” and “so as to afford protection” – leads to broad or narrow interpretations from WTO adjudicators that largely affect the outcome of each relevant case.

Neven (2001) offers one of the few quantitative investigations of GATT national treatment disputes, in which the analysis of protection according to

---

Article III is proven to be extremely challenging. While the paper’s key findings discuss the robust measures of protection such as the ability to raise price and the extent to which foreign firms do so in the presence of competition and substitutability between products, or lack thereof, Neven grounds his argument in the same debate discussed in the analysis herein. Neven argues that the standards associated with each of the two sets of products – like and directly competitive or substitutable products – are unclear in their relation to one another. Neven emphasizes the challenge of establishing a clear method for evaluation of protection because of the failure of WTO Panels to clarify distinctions between Articles III:2 and III:4. He also argues that the case law offered by the Dispute Settlement Body (DSB) does not offer an adequate opinion on how best to measure protectionism according to Article III.

2.2 Hoekman, Horn, and Mavroidis (2008)

Of the quantitative literature exploring trends in WTO dispute settlement, Hoekman, Horn and Mavroidis (2008) provide the strongest foundation for the analysis executed in this paper. Hoekman at al. analyze patterns of complainants and respondents involved in disputes with respect to each WTO agreement, specifically exploring the role of developing countries within the dispute settlement system and concluding that while these countries do not frequently engage as complainants, they are disproportionately represented as respondents to disputes. Additionally, the study provides evidence that 63 percent of cases involving third parties have been GATT cases, once again reinforcing the argument that GATT and its key provisions of national treatment and most-favored nation privileges attract the most attention in dispute settlement at the WTO.

Using a comprehensive dataset created for the World Bank by Horn and Mavroidis with information on disputes through 2006, Horn et al. analyze outcomes on specific legal claims based on the rough classification of claims accepted, rejected, or not ruled upon as provided in the dataset. Their study is one of the few throughout the literature on dispute settlement that uses such a classification of individual claim outcomes to perform empirical

15 Neven, D.J. “How Should Protection Be Evaluated in Article III Disputes?”, 423.
16 Neven, D.J. “How Should Protection Be Evaluated in Article III Disputes?”, 422.
18 Ibid.

Article III is proven to be extremely challenging. While the paper’s key findings discuss the robust measures of protection such as the ability to raise price and the extent to which foreign firms do so in the presence of competition and substitutability between products, or lack thereof, Neven grounds his argument in the same debate discussed in the analysis herein. Neven argues that the standards associated with each of the two sets of products – like and directly competitive or substitutable products – are unclear in their relation to one another. Neven emphasizes the challenge of establishing a clear method for evaluation of protection because of the failure of WTO Panels to clarify distinctions between Articles III:2 and III:4. He also argues that the case law offered by the Dispute Settlement Body (DSB) does not offer an adequate opinion on how best to measure protectionism according to Article III.

2.2 Hoekman, Horn, and Mavroidis (2008)

Of the quantitative literature exploring trends in WTO dispute settlement, Hoekman, Horn and Mavroidis (2008) provide the strongest foundation for the analysis executed in this paper. Hoekman at al. analyze patterns of complainants and respondents involved in disputes with respect to each WTO agreement, specifically exploring the role of developing countries within the dispute settlement system and concluding that while these countries do not frequently engage as complainants, they are disproportionately represented as respondents to disputes. Additionally, the study provides evidence that 63 percent of cases involving third parties have been GATT cases, once again reinforcing the argument that GATT and its key provisions of national treatment and most-favored nation privileges attract the most attention in dispute settlement at the WTO.

Using a comprehensive dataset created for the World Bank by Horn and Mavroidis with information on disputes through 2006, Horn et al. analyze outcomes on specific legal claims based on the rough classification of claims accepted, rejected, or not ruled upon as provided in the dataset. Their study is one of the few throughout the literature on dispute settlement that uses such a classification of individual claim outcomes to perform empirical
analysis on factors affecting dispute outcomes. Hoekman et al. encourage the further use of this same dataset in order to investigate more detailed hypotheses regarding dispute settlement. It is within this segment of the body of literature on dispute settlement that the present analysis of national treatment claims falls.

3 Data and Methodology

The data are drawn from the comprehensive World Bank WTO dataset created by Horn and Mavroidis, which contains quantitative and qualitative information on every dispute brought before the Dispute Settlement Body since 1995. The dataset provides details at every stage of dispute settlement, from Members’ requests for consultations through to the Panel and appellate proceedings, where applicable. The data utilized herein extracts information from this Horn and Mavroidis dataset for all disputes in which a GATT III national treatment citation was invoked by the complainant in the request for consultations. Because the original dataset contained only qualitative information on disputes through DS426, initiated in August 2011, relevant information was added. Between 1995 and 2013, 150 disputes with national treatment citations were filed at the WTO, of which 98 progressed to at least the earliest stages of Panel proceedings. Within this group, the Panel issued a report in 61 disputes and the Appellate Body issued a subsequent report in 44 disputes.19 This paper is concerned only with the 61 GATT Article III disputes for which the data from Panel rulings are available. It will investigate the effects of the 110 specific GATT Article III citations made within these 61 individual disputes on the likelihood of parties achieving a favorable decision at the time of the Panel ruling.

The main outcome of interest is a dichotomous variable measuring success at the Panel stage. First, variables indicating whether parties to a dispute requested the establishment of a Panel and, subsequently, whether these Panels issued reports are coded 0 or 1. Where Panel reports provide outcome information regarding the Panels’ rulings on specific claims, the dichotomous variable indicating favorable ruling is created. While the Horn and Mavroidis dataset specifies outcomes for all claims brought forth at the Panel stage, this variable is limited to national treatment claims; such citations include a general reference to GATT Article III and more specific

19 While there are 150 GATT cases, 156 disputes are analyzed at the consultations stage and 61 disputes are analyzed at the Panel stage. DS2 and DS139 enter into the data twice each and DS27 enters five times two account for the co-complainants engaged in each case.

Data and Methodology

The data are drawn from the comprehensive World Bank WTO dataset created by Horn and Mavroidis, which contains quantitative and qualitative information on every dispute brought before the Dispute Settlement Body since 1995. The dataset provides details at every stage of dispute settlement, from Members’ requests for consultations through to the Panel and appellate proceedings, where applicable. The data utilized herein extracts information from this Horn and Mavroidis dataset for all disputes in which a GATT III national treatment citation was invoked by the complainant in the request for consultations. Because the original dataset contained only qualitative information on disputes through DS426, initiated in August 2011, relevant information was added. Between 1995 and 2013, 150 disputes with national treatment citations were filed at the WTO, of which 98 progressed to at least the earliest stages of Panel proceedings. Within this group, the Panel issued a report in 61 disputes and the Appellate Body issued a subsequent report in 44 disputes.19 This paper is concerned only with the 61 GATT Article III disputes for which the data from Panel rulings are available. It will investigate the effects of the 110 specific GATT Article III citations made within these 61 individual disputes on the likelihood of parties achieving a favorable decision at the time of the Panel ruling.

The main outcome of interest is a dichotomous variable measuring success at the Panel stage. First, variables indicating whether parties to a dispute requested the establishment of a Panel and, subsequently, whether these Panels issued reports are coded 0 or 1. Where Panel reports provide outcome information regarding the Panels’ rulings on specific claims, the dichotomous variable indicating favorable ruling is created. While the Horn and Mavroidis dataset specifies outcomes for all claims brought forth at the Panel stage, this variable is limited to national treatment claims; such citations include a general reference to GATT Article III and more specific
citations of paragraphs one, two, and four (Articles III:1, III:2, III:4). If the Panel accepts the national treatment claim, thereby finding the respondent in violation of a GATT Article III provision, the favorable outcome variable is coded with a 1. Conversely, the outcome variable is coded 0 if the Panel either rejects the complainant’s claim or exercises judicial economy by not offering a ruling. Because Panels often rule on a number of claims referencing the same GATT III article within a single dispute, and because the outcomes may differ across multiple rulings of the same citation, this outcome variable is analyzed at the sub-dispute level. Outcomes are considered to be favorable for the complainant if individual claims accepted by the Panel match the claims specified in the initiation of the dispute. In total, the Panel accepted 80 claims, and either rejected or abstained from ruling on the remaining 30 claims.

The key explanatory variables are also binary and indicate the specific GATT III articles cited in requests for consultations and Panel establishment. As this paper hypothesizes that Panels’ interpretations of GATT Article III:2 and III:4 as they vary from one dispute to another have implications on the outcomes of each dispute, it is expected that the particular citations from GATT Article III will have an impact on the decisions rendered by the Panel at the conclusion of Panel proceedings. Thus, four of these explanatory variables are created at two main stages of dispute settlement, such that references to GATT III (in general), III:1, III:2, and III:4 are transformed into binary variables taking a value of either 0 or 1. For example, if the request for consultations includes one or more of these four GATT III citations, the dichotomous variable or variables associated with those citations are coded 1. The same procedure follows for citations in the complainant’s request for establishment of a Panel, in which the binary GATT Article III citation variable takes on a value of 1 if a reference to that article of GATT III is included in the request by the complainant to establish a Panel.

3.1 Model Specification

In dispute settlement, the outcomes ruled by Panels are not explained by traditional economic theory. Rather, the events occurring throughout the dispute settlement process are determined by a variety of explanatory variables, as described in the previous section. Moreover, the underlying theoretical motivating the anticipated effect of specific national treatment article citations on the rulings issued by the Panel is extracted from textual interpretation of said articles as developed throughout the relevant literature. In this case, the progression of a dispute from consultations to the Panel
stage and, subsequently, the ruling outcome of a legal claim litigated before the Panel are affected by a variety of factors specific to the corresponding dispute. Let Panel Stage\(_i\) and Outcome\(_i\) represent variables indicating progression to the panel stage and the outcome of legal claims ruled upon by the Panel, respectively. In regressions evaluating the determinants of Panel stage, Panel Stage\(_i\) is indexed by \(i\) individual disputes, while the Outcome\(_i\) is indexed by \(i\) individual legal claims. The model of the determinants of these variables is given by Equation (1):

\[
\text{Panel Stage}_i \quad \text{(Outcome}_i) = \alpha_0 + \beta_1 (\text{GATT III}) + \beta_2 (\text{GATT III:1}) + \beta_3 (\text{GATT III:2}) + \beta_4 (\text{GATT III:4}) + \theta X_i + \mu_i, 
\]

where the first four regressors are dichotomous variables indicating the presence of particular GATT III citations at the relevant stage of the dispute; \(X\) comprises a vector of dispute-specific attributes to control for their effects on the dependent variable. Parameters \(\alpha, \beta, \text{ and } \theta\) are to be estimated and \(\mu_i\) represents the unobservable factors affecting the dependent variable of interest, including political economy factors associated with legal measures at issue and WTO Panel concerns about systemic implications of legal rulings.

In this case, the model is best estimated by a probit analysis where the dependent variable of interest is unobserved. The first application of this probit model concerns the effect of GATT III citations invoked in the request for consultations on the likelihood that the dispute will proceed to the panel stage. Let Panel Stage\(_i\) be the unobserved variable indicating the progression of a dispute to the panel stage, with Panel Stage\(_j\) as an indicator variable equal to one if a dispute moves to Panel proceedings and zero otherwise, such that \(Pr(\text{Panel Stage}_i > 0) = Pr(\text{Panel Stage}_i = 1)\). The specification is given in Equation (2) as follows:

\[
Pr(\text{Panel Stage}_i = 1) = \frac{\phi(\alpha_0 + \beta_1 (\text{GATT III}) + \beta_2 (\text{GATT III:1}) + \beta_3 (\text{GATT III:2}) + \beta_4 (\text{GATT III:4}) + \theta X_i)}{1 + \phi(\alpha_0 + \beta_1 (\text{GATT III}) + \beta_2 (\text{GATT III:1}) + \beta_3 (\text{GATT III:2}) + \beta_4 (\text{GATT III:4}) + \theta X_i)}, 
\]

in which \(\phi(\cdot)\) represents the standard normal cumulative density function and determinants of the Panel stage progression include the specific GATT III citations invoked in the request for consultations as well as dispute-specific attributes. These include the number total articles cited, presence of third parties, identity of complainants and respondents, and inclusion of
politically contentious trade topics such as health and safety, technical barriers, and agriculture. In this specification as well as those following, it is expected that different GATT Article III citations, specifically Articles III:2 and III:4, will have different effects on the dependent variables. Articles III:2 and III:4 should affect the likelihood of progression to the Panel Stage and achieving favorable outcomes in Panel rulings in different directions.

The second application of the model investigates the Panel stage of litigation, for disputes in which a Panel report has been circulated and individual rulings on GATT III claims are known. The model will estimate the effect of specific Article III citations invoked at the consultation stage on the likelihood that complainants achieve a favorable ruling from the Panel. Analyzing the effect of the citations made at the consultation state is important, because complainants in future disputes will likely be interested in how to maximize the likelihood of obtaining a favorable outcome from the outset of the dispute. This specification is estimated by Equation (3):

$$\Pr(\text{Outcome}_{i} = 1) = \phi(\alpha_0 + \beta_1 \text{(GATT III)} + \beta_2 \text{(GATT III:1)} + \beta_3 \text{(GATT III:2)} + \beta_4 \text{(GATT III:4)} + \beta_5 X)$$

The third and final application of the model is similar to Equation (3), however it involves the estimation of the effect from GATT III citations made in the request for Panel establishment on the likelihood of a favorable outcome. This specification offers perhaps the clearest logical explanation of the relationship between Article III citations and outcomes of individual GATT III rulings because complainants are referencing specific Article III provisions with full knowledge that they will be subject to litigation before a Panel. Again, this specification will control for dispute-specific attributes that may affect the direction of Panel rulings on national treatment claims, such as the influence of third parties or the capacity for disputing parties to litigate effectively before the Panel. The estimating equation is identical to Equation (3) with the exception of the GATT III citations included as regressors, which are instead drawn from the complainants' request for establishment of a Panel rather than the request for consultations.

4 Findings

Before delving into regression analysis, consideration of some descriptive statistics is relevant. First, there are obvious differences in the percentage of disputes invoking each of the GATT III citations and their progression to the Panel stage. While 60 percent of disputes invoking general GATT III citations moved on to litigation, 85 percent of disputes invoking Article III:2

politically contentious trade topics such as health and safety, technical barriers, and agriculture. In this specification as well as those following, it is expected that different GATT Article III citations, specifically Articles III:2 and III:4, will have different effects on the dependent variables. Articles III:2 and III:4 should affect the likelihood of progression to the Panel Stage and achieving favorable outcomes in Panel rulings in different directions.

The second application of the model investigates the Panel stage of litigation, for disputes in which a Panel report has been circulated and individual rulings on GATT III claims are known. The model will estimate the effect of specific Article III citations invoked at the consultation stage on the likelihood that complainants achieve a favorable ruling from the Panel. Analyzing the effect of the citations made at the consultation state is important, because complainants in future disputes will likely be interested in how to maximize the likelihood of obtaining a favorable outcome from the outset of the dispute. This specification is estimated by Equation (3):

$$\Pr(\text{Outcome}_{i} = 1) = \phi(\alpha_0 + \beta_1 \text{(GATT III)} + \beta_2 \text{(GATT III:1)} + \beta_3 \text{(GATT III:2)} + \beta_4 \text{(GATT III:4)} + \beta_5 X)$$

The third and final application of the model is similar to Equation (3), however it involves the estimation of the effect from GATT III citations made in the request for Panel establishment on the likelihood of a favorable outcome. This specification offers perhaps the clearest logical explanation of the relationship between Article III citations and outcomes of individual GATT III rulings because complainants are referencing specific Article III provisions with full knowledge that they will be subject to litigation before a Panel. Again, this specification will control for dispute-specific attributes that may affect the direction of Panel rulings on national treatment claims, such as the influence of third parties or the capacity for disputing parties to litigate effectively before the Panel. The estimating equation is identical to Equation (3) with the exception of the GATT III citations included as regressors, which are instead drawn from the complainants' request for establishment of a Panel rather than the request for consultations.

4 Findings

Before delving into regression analysis, consideration of some descriptive statistics is relevant. First, there are obvious differences in the percentage of disputes invoking each of the GATT III citations and their progression to the Panel stage. While 60 percent of disputes invoking general GATT III citations moved on to litigation, 85 percent of disputes invoking Article III:2
citations moved on to the Panel stage. In comparison, only 64 percent of disputes invoking Article III:4 citations progressed to a Panel. The same trend holds for the full progression of disputes through the issuance of a Panel report, in which disputes involving Article III:2 claims by far exceeded other citations in the publication of a Panel report. While only 29 percent of disputes invoking Article III:4 citations saw a ruling from the Panel, 62 percent of Article III:2 citations translated to the issuance of a Panel report. In the event that cases did go before the Appellate Body, cases with Article III:2 citations were again the most frequent.

The results of the regression analysis confirm these trends. In the probit estimation of the effect citations invoked at the consultations stage on the likelihood of proceeding to a Panel stage, there are significant effects from each of the Article III citations. Citations invoking Article III (generally), Article III:1, and Article III:4 have somewhat significant negative effects, decreasing the likelihood of moving to a Panel. Conversely, citations involving Article III:2 tax charges have a equally significant positive effect on progression to the Panel stage, meaning that the presence of this GATT III citation increases the likelihood of requesting the establishment of the Panel.

Examining the marginal effects of each GATT III citation from specification (2) that includes controls for dispute-specific attributes, we see that the marginal effects of each citation on the likelihood of proceeding to a panel stage are approximately equal in magnitude and significance. GATT Article III:4 has the strongest negative effect, while GATT Article III:2 has the sole positive effect. At the 10% level, the presence of an Article III:4 citation has a substantial effect of reducing the likelihood of progression to a panel stage by 0.233, or 23.3 percent. Conversely, the presence of an Article III:2 citation has the effect of increasing the likelihood of progression to a panel stage by 0.175, or 17.5 percent.

Moving on to the investigation of the relationships between consultations and Panel stage citations on the likelihood of achieving a favorable GATT III ruling, we again use a probit model to estimating the impact of these citations on the likelihood of a favorable outcome. First, the analysis looks to the citations invoked at the consultation stage. The citations invoked at the consultations stage are relevant to the outcome of the case for two reasons: (1) Members file disputes at the WTO beginning with the request for consultations, and the citations included there are quite indicative of the citations that will be invoked later, should the dispute move to litigation (2) Previous regression analysis tells us that the citations invoked at the
consultation stage affect the likelihood of a dispute moving to a Panel stage in the first place, meaning that the likelihood of achieving a favorable Panel ruling is somewhat conditional on these consultations citations and their effect on whether or not the dispute even reaches the Panel.

Examining the marginal effects of each GATT III citation from this second model, we see that only the presence of a GATT Article III:2 citation has a significant effect on the likelihood of achieving a favorable outcome. Nonetheless, this effect is significant at the 1% level in the specification including only GATT III citations, and has a substantial effect of increasing the likelihood of achieving a favorable ruling at the Panel stage. In fact, the marginal effect of an Article III:2 citation being present at the consultations stage increases the likelihood of a favorable outcome by 0.362, or 36 percent. However, the significance of this effect disappears once controls for dispute-specific attributes are taken into account, perhaps signaling that Article III:2 disputes are also associated with other factors affecting the likelihood of favorable Panel rulings, such as the number of claims cited, identity of complainants and respondents, and politically contentious issues. Indeed, there is a significant, negative impact of the inclusion of SPS citations and a significant positive impact of the inclusion of Agriculture claims on the likelihood of a favorable outcome (1% and 5% significance, respectively). Again, the marginal effects of each variable assume that all other regressors are held at their mean value.

Lastly, it is important to note that the direction of influence coming from Article III:4 citations reverses logic that was present in the previous regression. At the consultations stage, the presence of Article III:4 claims seemingly decreased the likelihood of proceeding to the Panel stage; here, the presence of an Article III:4 consultations citation has the effect of decreasing the likelihood of a favorable panel ruling. Where it may have seemed that Article III:4 claims were more easily settled at the consultation stage, it seems that they are more difficult to effectively litigate from the perspective of the complainants at the Panel stage.

The final regression analysis includes the evaluation of citations invoked during the request for the establishment of a Panel, after consultations between Members have failed. Again, the probit model strives to estimate the impact of specific GATT III citations on the likelihood of a favorable panel ruling of the national treatment issue. As stated, results from this regression are perhaps the most important in resolving the interpretational debate, given that Panels are making their decisions entirely based on the specific portions of Article III cited by complainants at this point in the
dispute settlement process. Moreover, complainants are referencing these specific parts of Article III with full knowledge that these claims will be evaluated by the Panel, and are thus compelled to be strategic with their choice of citations. Examining the first specification of this third regression analysis that includes the specific Article III citations only, there is a significant effect of each citation on the likelihood of a favorable panel ruling. These effects are also consistent with previous regression results, such that Article III:1 and Article III:4 citations negatively impact the probability of a favorable outcome, while Article III:2 has a positive effect (III:1 and III:4 significant at 1%, III:2 significant at 5%) (See Table 5).

Due to the nature of the probit analysis, examining the marginal effects of these regressors is critical. All other effects held constant, the presence of an Article III:1 or Article III:4 citation in the request for establishment of a panel decreases the likelihood of achieving a favorable outcome by 0.435 and 0.413, respectively. Alternatively, the inclusion of an Article III:2 citation increases the likelihood of achieving a favorable panel outcome by 0.167, or a full 16.7 percent. While these results are highly significant, it is important to consider their effects when taking into account other dispute-specific attributes that will also likely affect Panel decision-making. When including the relevant controls, only the effects from Article III:1 and Article III:4 citations maintain significance. At the marginal effects level, only the effect of Article III:1 remains significant at the 10% level, however the direction of all GATT III citations remains consistent with all previous regression analysis performed herein. Indeed, it does appear that the effect of control variables accounts for some of the positive impact of Article III:2 citations on panel rulings. Specifically, the marginal effects of third party involvement, participation of the EU as a respondent, and the inclusion of SPS and TBT articles are significant factors influencing the likelihood of favorable panel outcomes on national treatment issues. First, the analysis shows that the presence of third parties at the Panel stage increase the probability of a favorable outcome; this may seem counter-intuitive given the ability for third parties to complicate disputes, however it may be true that the influence of third parties is explained by the direction of their involvement on behalf of the complainants. Alternatively, the presence of the EU as a respondent in these cases decreases the probability that complainants will achieve a favorable outcome.

Lastly, the conflicting impacts of politically contentious SPS and TBT issues are notable. The presence of SPS issues has a negative marginal effect on the probability of favorable outcomes being handed down, while the presence of TBT issues has a positive marginal effect. A possible explanation for this
may be that health and safety concerns represented by SPS references may preclude the Panel from finding measures inconsistent with GATT III if there is a serious health concern associated with potentially discriminatory measures. On the other hand, the referencing of TBT in GATT III cases may signal a stronger case on behalf of the complainant due to the frequent overlap of discriminatory measures in accordance with TBT and GATT III obligations.

5 Sensitivity Analysis
5.1 Endogeneity of GATT Article III Citations

There is potential for argument that the specific GATT Article III citations invoked at various stages of the dispute settlement process if complainants are aware of this trend in dispute settlement or experienced in their involvement with litigation before the WTO. However, concerns regarding endogeneity are mitigated by the fact that Article III:2 and Article III:4 specifically target tax and non-tax measures. The applicability of the measures at issue in disputes to one article or the other is exogenous due to the fact that Members allegedly in violation of national treatment obligations impose tax and non-tax charges independently of dispute settlement proceedings. Any argument of endogeneity in the explanatory variable of interest – GATT Article III citation – is mitigated by the fact that Members must identify the exact measures in respondents’ domestic law and match them according to the provisions of WTO agreements they seemingly violate. A tax measure at issue cannot be cited as Article III:4, and conversely, a non-tax measure cannot be invoked as a violation of Article III:2. Moreover, the inclusion of a control variable for the number of total citations can further mitigate any endogeneity of GATT III citations. If a Member invokes multiple citations of GATT Article III as well as other WTO articles in order to maximize the likelihood of a favorable outcome, the effect of the “kitchen sink” strategy is accounted for. Thus, there is no concern for the endogenous decision-making of complainants in choosing GATT Article III citations given that Articles III:2 and III:4 refer to distinctly different kinds of discriminatory measures exogenously determined well in advance of disputes coming to the WTO.

6 Conclusion

The results from the present analysis are significant for the resolution of the interpretational debate regarding the relevant scope of Articles III:2 and III:4. As explained, WTO jurisprudence muddies Members’ understanding of how each provision within GATT III is to be interpreted, despite the similar
language and effective application of each Article III section. The second source of confusion across these two articles is the determination of a standard above which protection may be considered discriminatory according to each Article III section.

Regression results may clarify these issues. Consistently, it seems that Article III:4 less likely translates to a positive outcome at the Panel stage, while Article III:2 may more likely lead to such an outcome. This alludes to the differences between “like” and “directly competitive or substitutable” products as defined in each paragraph of Article III, such that the scope of products that might fall under each category seems to be the decisive factor. Recall that the Appellate Body ruled, “the product scope of Article III:4, although broader than the first sentence of Article III:2, is certainly not broader than the combined product scope of the two sentences of Article III:2.”20 The results from this analysis provide evidence that this is true in practice, but also that the combined scope of both sentences of Article III:2 may in fact be broader than the scope of Article III:4, because of the lower likelihood of a favorable ruling from Article III:4 citations as found in the present analysis. Results may reveal that, in practice, the scope of Article III:2 is in fact broader than Article III:4, thus explaining the difference in likelihood of Panels accepting complainants’ arguments of GATT III violations.

Most importantly, however, it seems that the difference between tax and non-tax charges in disputes regarding national treatment violations carries the weight. Where Article III:2 claims are made in disputes, effective litigation and favorable rulings handed down by the Panel are more likely. Conversely, Article III:4 claims are fraught with challenges in litigation and much less likely to result in favorable rulings for complainants. The explanations here are twofold. First, taxes must be easier to investigate at litigation because of the clear paper trail and potential for economic and factual analysis of their discriminatory effect. Non-tax measures, including labeling or retail regulations are much more difficult to identify as discriminatory in their application and effect. Panels do go to extra lengths to explain rulings regarding Article III:4, but do not do so with regard to Article III:2 disputes because of the clarity tax measures offer in terms of discrimination. It may be most insightful to see that results from this paper show national treatment disputes are harder to win when complainants fight measures including non-charge elements than when the measures of concern are more simply issues of taxation. Prospects for Members to win disputes...
with regard to Article III:4 no-tax issues are lower than for disputes involving Article III:2 claims of discriminatory taxation.

To conclude, the results from this analysis offer insight regarding the practical Panel interpretation of Articles III:2 and Articles III:4 that was previously unavailable in the literature, as well as the relationship between citations of these specific articles and ultimate panel rulings. This paper also contributes to the wealth of literature that highlights the importance of other dispute-specific attributes that affect outcomes at the Panel stage by confirming that the influence of the US and EU as disputing parties, the presence of third parties, and the inclusion of politically contentious issues are indeed important determinants of how disputes play out at various stages of dispute settlement.
### Table 1: GATT Article III Citations (Consultations) and Panel Proceedings

<table>
<thead>
<tr>
<th></th>
<th>(1) dy/dx</th>
<th>(2) dy/dx</th>
<th>Marginal Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Panel Stage</td>
<td>Panel Stage</td>
<td></td>
</tr>
<tr>
<td>GATT III (general)</td>
<td>– 0.960**</td>
<td>– 0.897*</td>
<td>– 0.321** – 0.267*</td>
</tr>
<tr>
<td></td>
<td>(0.462)</td>
<td>(0.509)</td>
<td>(0.128) (0.136)</td>
</tr>
<tr>
<td>GATT III:1</td>
<td>– 0.214</td>
<td>– 0.860*</td>
<td>– 0.077 – 0.272*</td>
</tr>
<tr>
<td></td>
<td>(0.503)</td>
<td>(0.509)</td>
<td>(0.186) (0.154)</td>
</tr>
<tr>
<td>GATT III:2</td>
<td>0.124</td>
<td>0.619*</td>
<td>0.045 0.174*</td>
</tr>
<tr>
<td></td>
<td>(0.310)</td>
<td>(0.367)</td>
<td>(0.106) (0.093)</td>
</tr>
<tr>
<td>GATT III:4</td>
<td>– 0.868**</td>
<td>– 0.825*</td>
<td>– 0.277** – 0.233**</td>
</tr>
<tr>
<td></td>
<td>(0.425)</td>
<td>(0.477)</td>
<td>(0.110) (0.117)</td>
</tr>
<tr>
<td>Total articles cited</td>
<td>0.00568</td>
<td>0.002</td>
<td>(0.0224) (0.007)</td>
</tr>
<tr>
<td></td>
<td>(0.0224)</td>
<td>(0.007)</td>
<td></td>
</tr>
<tr>
<td>3rd Parties in Consultations</td>
<td>0.0376</td>
<td>0.011</td>
<td>(0.0261) (0.079)</td>
</tr>
<tr>
<td></td>
<td>(0.261)</td>
<td>(0.079)</td>
<td></td>
</tr>
<tr>
<td>US complainant</td>
<td>0.00981</td>
<td>0.003</td>
<td>(0.316) (0.095)</td>
</tr>
<tr>
<td></td>
<td>(0.316)</td>
<td>(0.095)</td>
<td></td>
</tr>
<tr>
<td>US respondent</td>
<td>1.081***</td>
<td>0.266***</td>
<td>(0.378) (0.068)</td>
</tr>
<tr>
<td></td>
<td>(0.378)</td>
<td>(0.068)</td>
<td></td>
</tr>
<tr>
<td>EU complainant</td>
<td>– 0.750**</td>
<td>– 0.235**</td>
<td>(0.318) (0.097)</td>
</tr>
<tr>
<td></td>
<td>(0.318)</td>
<td>(0.097)</td>
<td></td>
</tr>
<tr>
<td>EU respondent</td>
<td>1.250***</td>
<td>0.311***</td>
<td>(0.371) (0.066)</td>
</tr>
<tr>
<td></td>
<td>(0.371)</td>
<td>(0.066)</td>
<td></td>
</tr>
<tr>
<td>SPS</td>
<td>0.00775</td>
<td>0.002</td>
<td>(0.382) (0.114)</td>
</tr>
<tr>
<td></td>
<td>(0.382)</td>
<td>(0.114)</td>
<td></td>
</tr>
<tr>
<td>TBT</td>
<td>– 0.957***</td>
<td>– 0.297***</td>
<td>(0.334) (0.097)</td>
</tr>
<tr>
<td></td>
<td>(0.334)</td>
<td>(0.097)</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>– 0.196</td>
<td>– 0.069</td>
<td>(0.319) (0.100)</td>
</tr>
<tr>
<td></td>
<td>(0.319)</td>
<td>(0.100)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.231***</td>
<td>1.078*</td>
<td>(0.432) (0.572)</td>
</tr>
<tr>
<td></td>
<td>(0.432)</td>
<td>(0.572)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>156</td>
<td>156</td>
<td>156</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

### Table 1: GATT Article III Citations (Consultations) and Panel Proceedings

<table>
<thead>
<tr>
<th></th>
<th>(1) dy/dx</th>
<th>(2) dy/dx</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Panel Stage</td>
<td>Panel Stage</td>
</tr>
<tr>
<td>GATT III (general)</td>
<td>– 0.960**</td>
<td>– 0.897*</td>
</tr>
<tr>
<td></td>
<td>(0.462)</td>
<td>(0.509)</td>
</tr>
<tr>
<td>GATT III:1</td>
<td>– 0.214</td>
<td>– 0.860*</td>
</tr>
<tr>
<td></td>
<td>(0.503)</td>
<td>(0.509)</td>
</tr>
<tr>
<td>GATT III:2</td>
<td>0.124</td>
<td>0.619*</td>
</tr>
<tr>
<td></td>
<td>(0.310)</td>
<td>(0.367)</td>
</tr>
<tr>
<td>GATT III:4</td>
<td>– 0.868**</td>
<td>– 0.825*</td>
</tr>
<tr>
<td></td>
<td>(0.425)</td>
<td>(0.477)</td>
</tr>
<tr>
<td>Total articles cited</td>
<td>0.00568</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.0224)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>3rd Parties in Consultations</td>
<td>0.0376</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.261)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>US complainant</td>
<td>0.00981</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.316)</td>
<td>(0.095)</td>
</tr>
<tr>
<td>US respondent</td>
<td>1.081***</td>
<td>0.266***</td>
</tr>
<tr>
<td></td>
<td>(0.378)</td>
<td>(0.068)</td>
</tr>
<tr>
<td>EU complainant</td>
<td>– 0.750**</td>
<td>– 0.235**</td>
</tr>
<tr>
<td></td>
<td>(0.318)</td>
<td>(0.097)</td>
</tr>
<tr>
<td>EU respondent</td>
<td>1.250***</td>
<td>0.311***</td>
</tr>
<tr>
<td></td>
<td>(0.371)</td>
<td>(0.066)</td>
</tr>
<tr>
<td>SPS</td>
<td>0.00775</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.382)</td>
<td>(0.114)</td>
</tr>
<tr>
<td>TBT</td>
<td>– 0.957***</td>
<td>– 0.297***</td>
</tr>
<tr>
<td></td>
<td>(0.334)</td>
<td>(0.097)</td>
</tr>
<tr>
<td>AG</td>
<td>– 0.196</td>
<td>– 0.069</td>
</tr>
<tr>
<td></td>
<td>(0.319)</td>
<td>(0.100)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.231***</td>
<td>1.078*</td>
</tr>
<tr>
<td></td>
<td>(0.432)</td>
<td>(0.572)</td>
</tr>
<tr>
<td>Observations</td>
<td>156</td>
<td>156</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Table 2: GATT Article III Citations (Panel Request) and Panel Rulings

<table>
<thead>
<tr>
<th></th>
<th>Marginal Effects</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>Favorable Ruling</td>
<td>dy/dx</td>
<td>dy/dx</td>
</tr>
<tr>
<td>GATT III (only)</td>
<td>– 1.172</td>
<td>1.143</td>
<td>– 0.274</td>
</tr>
<tr>
<td></td>
<td>(0.778)</td>
<td>(302.5)</td>
<td>(0.168)</td>
</tr>
<tr>
<td>GATT III:1</td>
<td>– 1.638***</td>
<td>– 1.196*</td>
<td>0.436***</td>
</tr>
<tr>
<td></td>
<td>(0.419)</td>
<td>(0.649)</td>
<td>(0.0986)</td>
</tr>
<tr>
<td>GATT III:2</td>
<td>0.813**</td>
<td>0.315</td>
<td>0.167**</td>
</tr>
<tr>
<td></td>
<td>(0.374)</td>
<td>(0.502)</td>
<td>(0.0706)</td>
</tr>
<tr>
<td>GATT III:4</td>
<td>– 2.178***</td>
<td>– 2.696***</td>
<td>0.413***</td>
</tr>
<tr>
<td></td>
<td>(0.564)</td>
<td>(1.022)</td>
<td>(0.0680)</td>
</tr>
<tr>
<td>Total articles cited</td>
<td>– 0.0201</td>
<td>– 0.00371</td>
<td>– 0.00371</td>
</tr>
<tr>
<td></td>
<td>(0.0441)</td>
<td>(0.00813)</td>
<td></td>
</tr>
<tr>
<td>3rd parties (in Panel Stage)</td>
<td>1.688</td>
<td>0.354*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.220)</td>
<td>(0.211)</td>
<td></td>
</tr>
<tr>
<td>US complainant</td>
<td>– 0.845</td>
<td>– 0.147</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.662)</td>
<td>(0.103)</td>
<td></td>
</tr>
<tr>
<td>US respondent</td>
<td>– 4.979</td>
<td>– 0.625</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(302.5)</td>
<td>(7.688)</td>
<td></td>
</tr>
<tr>
<td>EU complainant</td>
<td>– 0.581</td>
<td>– 0.116</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.671)</td>
<td>(0.139)</td>
<td></td>
</tr>
<tr>
<td>EU respondent</td>
<td>2.009*</td>
<td>– 0.438**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.061)</td>
<td>(0.179)</td>
<td></td>
</tr>
<tr>
<td>SPS</td>
<td>– 1.891</td>
<td>– 0.445*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.177)</td>
<td>(0.235)</td>
<td></td>
</tr>
<tr>
<td>TBT</td>
<td>1.044</td>
<td>0.134*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.997)</td>
<td>(0.0777)</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>0.951</td>
<td>0.128</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.934)</td>
<td>(0.0805)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.637***</td>
<td>2.512*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.593)</td>
<td>(1.341)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>110</td>
<td>110</td>
<td>110</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.0.
Bibliography


This paper examines the relationship between income inequality, ethnic fragmentation, and spending on public education in the United States. The existing literature conflicts as to whether higher inequality leads to higher or lower levels of spending, while the literature finds a consistently negative relationship between ethnic fragmentation and public goods, including education. This paper contributes to the literature by testing the combined effect of inequality and ethnic fragmentation on educational spending, using the Median Voter Theory. To address possible Tiebout sorting, this paper instruments for inequality using a predicted Gini coefficient. Drawing on a panel of school districts from 1970-2000, I find a positive relationship between inequality and school expenditures per student, a negative relationship between ethnic fragmentation and expenditures, and a positive relationship between the interaction of inequality and ethnic fragmentation, and expenditures. While consistent, these results differ in magnitude and significance across the four regions of the United States, the Northeast, the Midwest, the South, and the West. The results demonstrate the existence of an interaction effect between inequality and ethnic fragmentation, which may have implications for further policy research interested in education.
This paper investigates the impact of access to finance on firm performances based on variations across different ownership structures in China. By examining the 3,401 random samples provided by the 2003 and 2005 Investment Climate Surveys from the World Bank, I reveal how productivity responds to additional short-term capital across state-owned enterprises, foreign-invested firms, collectively-owned enterprises and domestic private companies, conditioning on size, industry and geographical features. By employing an estimation methodology that adopts two approaches, I find that domestic private firms are most effective at allocating loan resources to boost productivity but they face large capital constraints compared to other types of firms. While the state sector is on average the least productive, minority state enterprises are worse than wholly state enterprises; I theorize that this is because the general managers in minority-owned state enterprises have weaker political incentives and stronger individual interests which induce them to trade long-term productivity for higher short-run output. Furthermore, multinational companies and equity joint ventures invest less in efficiency change and technological improvements than expected. The central government should adjust the institutional arrangements as well as the financial climate to reduce the current limitations that restrict the development of the private sector.
In the last 20 years, reforms to privatize state-owned enterprise have gradually taken place to complement the increasingly market-oriented economy and to enhance productivity. Nonetheless, they still enjoy many privileges that private firms do not, such as the ability to obtain sufficient capital to fund business operations and investments. This paper contributes to the discussion regarding ownership structures and their productivity levels, specifically in the context of microeconomic research in China. Since state-owned enterprises are dominant actors in Chinese strategic industries, their distinction with other firms of different ownership structures carry strong policy significance. Their operational and management structures interact uniquely with exogenous variables both because of their specific state-driven goals and economic incentives. In order to examine the outcomes of corporatization reforms in the Chinese state sector, particularly on the efficiencies of capital allocation, we need to understand how different ownership structures promote or limit the impact that increased access to finance has on productivity.

The discussion of factor productivity in this paper targets the “technical characteristics in production process” (Nadiri 1970, p.1141). Felipe (1999) argues that change in total factor productivity (TFP) can be isolated into two major components: production efficiency changes and technological advancements. “Technical efficiency change” involves higher quality of corporate governance, the ability to lower unit costs for all factors of production, and the capacity to adopt practices that provide better outcomes with existing technology; “Technological progress” is related to the development of best practices and new techniques in the production processes (Felipe 1999, p.7). The two elements both contribute to the improvement of firm performances, but Chinese companies tend to suffer from low production efficiencies (Zheng, Bigsten and Hu 2009). In fact, Zheng, Liu and Bigsten (2003) declare that the source of productivity growth among the state enterprises during 1980-1994 was mainly through technical breakthroughs.

Considering that China’s business and financial environment is carefully regulated, ownership structures affect firms’ ability to secure external finance. Generally, while state-owned enterprises (SOEs) face softer budget constraints and foreign-owned companies (FOEs) have easier access to finance, private firms typically encounter capital limitations in China’s highly control financial markets (David and Wei 2007). Since access to finance is key to the growth of the private sector, this credit gap potentially hinders the development of the economy. Although the State Council has slowly begun addressing this issue by introducing capital market reforms to
ensure more financing for small-medium enterprises, the government overlooks the importance of ownership structures in determining the effect of this policy. This paper will seek to reveal that private enterprises are the best at allocating available credit to productivity growth.

The main empirical results are as follows. Based on 3,401 observations, I find that when provided with more credit, domestic private firms can channel additional financial resources to raise productivity by a greater extent than state-owned enterprises and foreign invested firms. Specifically, short-term capital has less influence on the productivity of minority-owned state enterprises than wholly-owned state enterprises. Since general managers from minority state enterprises have fewer political incentives to achieve long-term growth and are rewarded based on annual performance, they make investment decisions to raise short-run output instead of long-term productivity. Also, while firms with foreign investments should use their capital productively, some multinationals and joint ventures clustered in the manufacturing sectors are not as efficient as their domestic counterparts. These results suggest that foreign companies in specific labor-intensive industries suffer from efficiency loss when entering the Chinese market and the domestic business environment is not conducive enough for them to invest in technological progress.

2 Literature Review

In the context of China, the issue of access to finance as a development challenge focuses on small-medium enterprises (SMEs). Mu (2002) analyzes China’s SME lending to evaluate the impediments blocking SME access to finance and credit guarantee schemes. Beck and Kunt (2006) consider access to finance across countries as a growth constraint for SMEs and suggest financing tools that can alleviate the problem of weak financial infrastructure. The authors believe that ease in access to finance contributes to a business environment that is “conducive to competition and private commercial transactions” and SMEs are prevented from reaching their optimal size due to a lack of external financial resources.

In terms of firm ownership structures, Chinese state-owned enterprises are notorious for the unfair advantages they gain through state support. Research usually reinforces the concept that state-owned enterprises and foreign firms experience weaker credit limitations than private firms (Poncet, Steingress and Vadenbussche 2010). According to Bai and Wang (1997), the existing state bureaucracy allows most SOEs to face a soft budget constraint where enterprises still receive funding even when they are running at a loss. These firms are thus shielded from market competition to the point that SOEs are ensuring more financing for small-medium enterprises, the government overlooks the importance of ownership structures in determining the effect of this policy. This paper will seek to reveal that private enterprises are the best at allocating available credit to productivity growth.

The main empirical results are as follows. Based on 3,401 observations, I find that when provided with more credit, domestic private firms can channel additional financial resources to raise productivity by a greater extent than state-owned enterprises and foreign invested firms. Specifically, short-term capital has less influence on the productivity of minority-owned state enterprises than wholly-owned state enterprises. Since general managers from minority state enterprises have fewer political incentives to achieve long-term growth and are rewarded based on annual performance, they make investment decisions to raise short-run output instead of long-term productivity. Also, while firms with foreign investments should use their capital productively, some multinationals and joint ventures clustered in the manufacturing sectors are not as efficient as their domestic counterparts. These results suggest that foreign companies in specific labor-intensive industries suffer from efficiency loss when entering the Chinese market and the domestic business environment is not conducive enough for them to invest in technological progress.
generally found to be resource-rich and operationally inefficient (Girma, Gong and Gorg 2008). Naughton (2007) describes that the government, from 1979 – 1993, tried to enhance SOE productivities by linking the managers’ incentives to firm profitability and altering the pricing system. Even as the state encouraged SOEs to gain competitiveness in the market by transforming their objective functions, their internal organization remained unchanged. When the authority of SOE managers grew from privatization, corruption became another major concern because the government had less control over managers and opportunities for rent-seeking increased (Naughton 2007). While various types of enterprise reforms have been implemented since the mid-1990s to further corporatization with harder budget constraints, transparency and effective external oversight remains a challenge for the state sector (Naughton 2007). Thus, the corporate governance models for state-owned enterprises differ from domestic private companies and foreign firms, leading them to make diverging investment decisions.

In general, there are few micro-level assessments on the imperfections of capital allocation in China. Currently, most of the literature emphasizes the effect of financial access in relation to small-medium enterprises, and the policy implications of the research are focused on size-oriented measures. Therefore, this paper expands on the topic by investigating how ownership structure determines the impact of access to finance on firm performance in China.

It also attempts to evaluate whether the institutional changes resulting from decades of enterprise reforms have generated positive changes in corporate governance and total factor productivity. The paper categorizes firms by ownership structure while controlling for size to compare which types of firms can gain the most from improvements in access of finance. Depending on the nature of firm ownership, the companies’ governance structures dictate their investment decisions; consequently, the effects of lowering credit constraints on productivity differ across firms. Therefore, assessing how ownership structure matters to firm performance from a microeconomic perspective is important even when a majority of the literature has already revealed the macro implications for investment inefficiencies.

3 Empirical Methodology

The empirical model involves a set of regressions that uses two approaches to estimate the Cobb-Douglas production function. The ordinary least squares (OLS) models adopted from Levinson and Petrin (2003) and Fernandes (2007) employ an intermediate input, raw materials usage, to estimated the Cobb-Douglas production function. The ordinary least squares (OLS) models adopted from Levinson and Petrin (2003) and Fernandes (2007) employ an intermediate input, raw materials usage, to
adjust for the simultaneity bias between inputs and productivity. OLS estimation assumes that inputs are exogenous but they are in fact endogenous because managers decide on their plant productivity based on the availability of factors. The intermediate input can be used for correcting the OLS estimation bias as it carries positive value and is subjected to productivity shocks (Fernandes 2007).

3.1 Direct Approach

The one-step approach includes the main regressor loan quota in the production function equation along with other firm-specific covariates. The equations determine the link between credit available and productivity by analyzing outputs of firms across all industries and geographic regions in China. For the base specification:

\[ \pi_i = \alpha + \beta_1 \ln(K_i) + \beta_2 \ln(M_i) + \beta_3 \ln(L_i) + \beta_4 \ln(L_Qi) + \beta_5 SSi + \beta_6 Si + \beta_7 Xi + \beta_8 RD + \beta_9 I_i + \epsilon_i \]  

(1)

Where the dependent variable \( \pi_i \) is defined as the log of output for firm i as measured by the core business income per unit of employment. The variables \( K_i, M_i, \) and \( L_i \) are capital, raw materials and labor, which are basic attributes for the production function. \( L_Qi \) is the volume of the loan quota reported by the firms in 2003 and 2005. SSi is the percentage of state share in the specific firm, which is a proxy for government influence on production, investment and labor decisions. RD, is the level of research and development expenditure which is related to productivity because R&D spending initiates technical progress. I, is the measure for interest expenditure which affects the management’s decision on what to invest; when the cost of capital is high, the firm may spend on projects that have lower risk and faster returns instead of on structural changes. Si records the size of the firm in the form of total employment and Xi is a dummy variable for firms that export. These variables act as controls because larger firms tend to become more bureaucratic and firms that export are often very efficient.

Next, the ownership dummy variables and interaction terms are introduced to the basic production function equation:

\[ \pi_i = \alpha + \beta_1 \ln(K_i) + \beta_2 \ln(M_i) + \beta_3 \ln(L_i) + \beta_4 \ln(L_Qi) + \beta_5 SSi + \beta_6 Si + \beta_7 Xi + \beta_8 RD + \beta_9 I_i + \beta_{10}S{OSi} + \beta_{11}D{FOE} + \beta_{12}D{COE} + \beta_{13}LD{COE} + \beta_{14}SSi + \beta_{15}Si + \beta_{16}RD + \beta_{17}I_i + \epsilon_i \]  

(2)

1Size is divided into small, medium and large according to the 2003 definition from China’s Ministry of Industry and Information Technology.

adjust for the simultaneity bias between inputs and productivity. OLS estimation assumes that inputs are exogenous but they are in fact endogenous because managers decide on their plant productivity based on the availability of factors. The intermediate input can be used for correcting the OLS estimation bias as it carries positive value and is subjected to productivity shocks (Fernandes 2007).

3.1 Direct Approach

The one-step approach includes the main regressor loan quota in the production function equation along with other firm-specific covariates. The equations determine the link between credit available and productivity by analyzing outputs of firms across all industries and geographic regions in China. For the base specification:

\[ \pi_i = \alpha + \beta_1 \ln(K_i) + \beta_2 \ln(M_i) + \beta_3 \ln(L_i) + \beta_4 \ln(L_Qi) + \beta_5 SSi + \beta_6 Si + \beta_7 Xi + \beta_8 RD + \beta_9 I_i + \epsilon_i \]  

(1)

Where the dependent variable \( \pi_i \) is defined as the log of output for firm i as measured by the core business income per unit of employment. The variables \( K_i, M_i, \) and \( L_i \) are capital, raw materials and labor, which are basic attributes for the production function. \( L_Qi \) is the volume of the loan quota reported by the firms in 2003 and 2005. SSi is the percentage of state share in the specific firm, which is a proxy for government influence on production, investment and labor decisions. RD, is the level of research and development expenditure which is related to productivity because R&D spending initiates technical progress. I, is the measure for interest expenditure which affects the management’s decision on what to invest; when the cost of capital is high, the firm may spend on projects that have lower risk and faster returns instead of on structural changes. Si records the size of the firm in the form of total employment and Xi is a dummy variable for firms that export. These variables act as controls because larger firms tend to become more bureaucratic and firms that export are often very efficient.

Next, the ownership dummy variables and interaction terms are introduced to the basic production function equation:

\[ \pi_i = \alpha + \beta_1 \ln(K_i) + \beta_2 \ln(M_i) + \beta_3 \ln(L_i) + \beta_4 \ln(L_Qi) + \beta_5 SSi + \beta_6 Si + \beta_7 Xi + \beta_8 RD + \beta_{10}S{OSi} + \beta_{11}D{FOE} + \beta_{12}D{COE} + \beta_{13}LD{COE} + \beta_{14}SSi + \beta_{15}Si + \beta_{16}RD + \beta_{17}I_i + \epsilon_i \]  

(2)

1Size is divided into small, medium and large according to the 2003 definition from China’s Ministry of Industry and Information Technology.
Where the D variables are binary variables for state-owned enterprises (SOE), foreign-owned enterprises (FOE), and collectively-owned enterprises (COE). Interactions between loan quota and the types of ownerships enable different effects of increased access to finance on productivity according to individual ownership structures. The interaction terms allow for the notion that firms with more efficient operational and management structures benefit more extensively from the additional available credit. The comparison group is domestic private firms; since private firms have been consistently financially strained, a positive shock in access to finance may yield a higher productivity boost than with other companies.

Then, the state-owned and foreign-owned firms are further dissected into the sub-groups to consider the differences within:

\[ \pi_i = \alpha_0 + \beta_1 \ln(K_i) + \beta_2 \ln(M_i) + \beta_3 \ln(L_i) + \beta_4 \ln(LQi) + \beta_5 \ln(LQi)*DWSOE + \beta_6 \ln(LQi)*DMASOE + \beta_7 \ln(LQi)*DMISOE + \beta_8 \ln(LQi)*DWFOE + \beta_9 \ln(LQi)*DMAFOE + \beta_{10} \ln(LQi)*DMIFOE + \beta_{11} \ln(LQi)*DCOE + \beta_{12} DWSOE + \beta_{13} DMASOE + \beta_{14} DMISOE + \beta_{15} DWSOE + \beta_{16} DMAFOE + \beta_{17} DMIFOE + \beta_{18} DCOE + \beta_0 S_i + \beta_{20} S_i + \beta_{21} X_i + \beta_{22} RD_i + \beta_{23} I_i + \varepsilon_i \]

(3) Where the D variables are binary variables for wholly state enterprises (WSOE), majority state enterprises (MASOE), minority state enterprises (MISOE), foreign-owned enterprises (WFOE), majority foreign enterprises (MAFOE) and minority foreign enterprises (MIFOE).

This direct approach assumes constant factor elasticity across sectors so that input coefficients are identical across the production function equation. In reality, however, different industries have different production processes and pooling the firms together across sectors and locations may cause an estimation bias. To address this issue, this paper utilizes the indirect approach in addition to the direct approach just described.

### 3.2 Indirect Approach

The two-step approach first estimates the production function equation without the main regressors. Then, extracting the coefficients of inputs estimated by the first equation, the total factor productivity is determined by the residuals and regressed with the key independent variables in the productivity function.
Step 1:
\[ \pi_i = \alpha_0 + \beta_1 \ln(K_i) + \beta_2 \ln(M_i) + \beta_3 \ln(L_i) + \epsilon_i \]

From the residual of equation (4), the total factor productivity (TFP), measured for each firm \( i \), is different even within the same ownership structures and industries. In the second step, the main regressor for access to finance is included.

Step 2:
\[ TFP_i = \alpha_0 + \beta_1 \ln(LQ_i) + \beta_2 SS_i + \beta_3 X_i + \beta_4 RD_i + \beta_5 I_i + \epsilon_i \]

Finally, similar to the direct approach, the interaction terms for various ownership structures are introduced for analysis.

\[ TFP_i = \alpha_0 + \beta_1 \ln(LQ_i) + \beta_2 \ln(LQ_i)*DSOE + \beta_3 \ln(LQ_i)*DFOE + \beta_4 \ln(LQ_i)*DCOE + \beta_5 DSOE + \beta_6 DFOE + \beta_7 DCOE + \beta_8 SS_i + \beta_9 S_i + \beta_{10} X_i + \beta_{11} RD_i + \beta_{12} I_i + \epsilon_i \]

The indirect approach may potentially face issues where some factor inputs from the first equation can be correlated with the key independent variables in the second regression (Udomsaph 2012). The input capital is endogenous because access to finance influences the firm’s investment capacity and managers change the assets they hold given a specific financial environment. Since the China Investment Climate Survey does not provide panel data, the covariates cannot be logged to solve the problem with endogenous variables. For this reason, the indirect approach still suffers from simultaneity bias even though it does not assume constant factor elasticity across industries.

4 Data Analysis
4.1 Comparisons Among The Four Ownership Structures

In general, findings from both approaches indicate that state-owned enterprises are the least capable of raising their productivity when experiencing an influx of capital in the form of loans. The results support the
hypothesis that domestic private firms, currently enduring stricter financial constraints, are more efficient in using additional capital for advancing their productivities.

Columns 2 for both approaches demonstrate that the impact of loan quota on productivity is the highest for privately owned enterprises. In Table 2, when a privately owned firm secures 10% more loan quota, its productivity raises by 0.96%. For SOEs, with the same increase in loan quota, the change in output is lower than that of private firms by 0.56%; the net influence is a positive 0.4% on productivity. For the indirect approach, the contrast between state-owned enterprises and domestic private firms is larger. Domestic private firms are more than three times more productive than SOEs when the firms are offered an additional 10% of loans. After accounting for unobserved fixed industry and location characteristics in columns 3-5, the productivity differences are even more pronounced and the results remain significant on the 1% level across private and state ownership structures.

The contrast between SOEs and private firms is not alarming: it is a strong reminder that SOEs are still not functioning as efficient economic actors. The central government needs to speed up financial reforms to address the credit needs of private companies, specifically through preferential deals with banking institutions. Policy prescriptions should eliminate the obstacles that domestic private companies encounter as well as reevaluate the enterprise reforms targeted at state-owned enterprises. Even though the findings on foreign invested companies seem to contradict existing literature, it is important to note that observations mainly cover firms in labor-intensive industries. As a result, the findings on foreign companies should not be generalized to foreign direct investments in other capital-intensive sectors.

4.2 Analysis on SOE Sub-Groups

Next, I look into the sub-segments within state-owned enterprises in columns 6-9. Both direct and indirect approaches reveal that wholly owned SOEs are more efficient than the minority-owned state firms and the levels of significance rise after adjusting for industry and location fixed effects. It becomes clear that wholly and minority state enterprises drive the significant results for the SOEs in columns 2-5. Referring to column 9 in Table 2, 10% more loan quota boosts the productivity by 0.24% for wholly state firms and by 0.08% for minority SOEs, creating a three times difference between the two. Moreover, the indirect approach reports that providing short-term capital to minority state companies harms productivity.
In order to better understand why wholly and minority state enterprises respond differently to financial accessibility, this paper conducts further analyses on their corporate governance structures. The analysis reveals that enterprise reforms have created a divergence on the decision-making frameworks among SOEs. Table 5 demonstrates that providing private monetary incentives for general managers to raise firm performances is effective for wholly state enterprises but not minority state enterprises; this compensation structure increases the productivity of the former by a much larger extent than the latter. Unfortunately, however the same cannot be said for minority state enterprises; the general managers from MISOEs respond weakly to the monetary stimulus provided by the state.

Instead, the above patterns can be explained by the different concerns general managers have in wholly state enterprises and minority state owned firms. Referring to Table 4, the industry compositions of WSOEs and MISOEs suggest that the former are more concentrated than the latter in strategic and special interest sectors, such as natural resources, machinery and the information technology industry. WSOEs have 58.3%, 18.2% and 11.1% production shares in the tobacco, coke and refined petroleum products, as well as IT industry while minority state firms are not present in these industries at all. The central government has pushed through many enterprise reforms in SOEs, and the Communist Party bureaucrats understand clearly their mandate to make SOEs more competitive. Shirk (1993, p.349) highlights that “the party’s power to promote and demote officials gave subordinate officials a career incentive to go along with reforms... [and] facilitated the implementation of reforms among factory managers and workers” in state enterprises. Since the wholly owned SOEs assume more essential economic responsibilities, the managers are likely to have stronger concerns for their political career in the party. This in turn gives them greater motivation to utilize the enterprise reforms to raise the firm TFP and achieve higher outputs in the long run.

5 Limitations

While the Investment Climate Surveys provide a large set of observations for analysis, the data also present challenges to investigate changes across time. Due to the limitations of the cross-sectional survey data, the endogeneity issues cannot be addressed by lagging the independent variables. For this reason, our story is one of correlation rather than causation. Had the World Bank created a more informative database, there would be great value in exploring the productivity changes of different ownership structures as they evolved during and after enterprise reforms. Moreover, the data suffers from

In order to better understand why wholly and minority state enterprises respond differently to financial accessibility, this paper conducts further analyses on their corporate governance structures. The analysis reveals that enterprise reforms have created a divergence on the decision-making frameworks among SOEs. Table 5 demonstrates that providing private monetary incentives for general managers to raise firm performances is effective for wholly state enterprises but not minority state enterprises; this compensation structure increases the productivity of the former by a much larger extent than the latter. Unfortunately, however the same cannot be said for minority state enterprises; the general managers from MISOEs respond weakly to the monetary stimulus provided by the state.

Instead, the above patterns can be explained by the different concerns general managers have in wholly state enterprises and minority state owned firms. Referring to Table 4, the industry compositions of WSOEs and MISOEs suggest that the former are more concentrated than the latter in strategic and special interest sectors, such as natural resources, machinery and the information technology industry. WSOEs have 58.3%, 18.2% and 11.1% production shares in the tobacco, coke and refined petroleum products, as well as IT industry while minority state firms are not present in these industries at all. The central government has pushed through many enterprise reforms in SOEs, and the Communist Party bureaucrats understand clearly their mandate to make SOEs more competitive. Shirk (1993, p.349) highlights that “the party’s power to promote and demote officials gave subordinate officials a career incentive to go along with reforms... [and] facilitated the implementation of reforms among factory managers and workers” in state enterprises. Since the wholly owned SOEs assume more essential economic responsibilities, the managers are likely to have stronger concerns for their political career in the party. This in turn gives them greater motivation to utilize the enterprise reforms to raise the firm TFP and achieve higher outputs in the long run.

5 Limitations

While the Investment Climate Surveys provide a large set of observations for analysis, the data also present challenges to investigate changes across time. Due to the limitations of the cross-sectional survey data, the endogeneity issues cannot be addressed by lagging the independent variables. For this reason, our story is one of correlation rather than causation. Had the World Bank created a more informative database, there would be great value in exploring the productivity changes of different ownership structures as they evolved during and after enterprise reforms. Moreover, the data suffers from
a potential estimation bias in reporting the loan quota because accountants did not directly fill out the survey. Thus, the measurement for access to finance might be an approximation rather than a number extracted from financial statements.

Likewise, as the empirical methodology used was developed primarily for examining the productivity of private firms, it may be less compatible for state-owned enterprises. Since the SOEs’ successes are dependent on more than economic gains, the TFP model may fail to evaluate whether the SOEs are “productive” in other areas such as employment. Nonetheless, given that the central government has strived to privatize the SOEs, the model based on production behaviors in the private sector is suitable to assess state-owned enterprises as economic actors.

6 Conclusion

The empirical analyses conducted in the paper shed light on how firm ownership matters for financial reform in China. The effect of improving access to finance on firm productivity depends on the type of operational and management structure because this determines how firms react to the exogenous business environment. In the current policy wave that emphasizes liberalizing capital constraints for small-medium enterprises, more attention should be paid to other firm characteristics. Ownership structure is potentially as significant, if not more, than size in deciding the external constraints that a firm faces.

This paper, using a direct and indirect approach, discovers that when more loans are employed, productivity increases by the largest degree for private firms. State-owned enterprises and foreign invested firms are not as productive as domestic private companies when they receive additional short-term capital. Within the state sector, minority-owned state enterprises are less productive than wholly owned state enterprises. I find a tradeoff of long-term productivity for short-run output in the investment decisions of general managers from minority state enterprises. In addition, external constraints present in the domestic environment make foreign firms less efficient as expected. Multinational companies and joint ventures have strong potential to increase their investments into technological progress and efficiency change if existing legal institutions improve at protecting intellectual property. Given that state-owned enterprises typically encounter the fewest credit constraints despite being the worst at efficiency utilizing new investment, these results signal a misallocation of short-term capital, in turn revealing larger inefficiencies in the Chinese financial market.

Likewise, as the empirical methodology used was developed primarily for examining the productivity of private firms, it may be less compatible for state-owned enterprises. Since the SOEs’ successes are dependent on more than economic gains, the TFP model may fail to evaluate whether the SOEs are “productive” in other areas such as employment. Nonetheless, given that the central government has strived to privatize the SOEs, the model based on production behaviors in the private sector is suitable to assess state-owned enterprises as economic actors.

6 Conclusion

The empirical analyses conducted in the paper shed light on how firm ownership matters for financial reform in China. The effect of improving access to finance on firm productivity depends on the type of operational and management structure because this determines how firms react to the exogenous business environment. In the current policy wave that emphasizes liberalizing capital constraints for small-medium enterprises, more attention should be paid to other firm characteristics. Ownership structure is potentially as significant, if not more, than size in deciding the external constraints that a firm faces.

This paper, using a direct and indirect approach, discovers that when more loans are employed, productivity increases by the largest degree for private firms. State-owned enterprises and foreign invested firms are not as productive as domestic private companies when they receive additional short-term capital. Within the state sector, minority-owned state enterprises are less productive than wholly owned state enterprises. I find a tradeoff of long-term productivity for short-run output in the investment decisions of general managers from minority state enterprises. In addition, external constraints present in the domestic environment make foreign firms less efficient as expected. Multinational companies and joint ventures have strong potential to increase their investments into technological progress and efficiency change if existing legal institutions improve at protecting intellectual property. Given that state-owned enterprises typically encounter the fewest credit constraints despite being the worst at efficiency utilizing new investment, these results signal a misallocation of short-term capital, in turn revealing larger inefficiencies in the Chinese financial market.
Therefore, the Chinese state council should focus on ownership-specific financial obstacles to complement policy solutions based on the size of the firm.

While this paper discusses some outcomes of the enterprise reforms, more research has to be done to evaluate whether they have continued to transform state-owned enterprises for the better. The findings suggest that the privatization incentives have not motivated all SOEs to become more efficient and in specific circumstances have generated counterproductive effects. Moreover, the state sector “has acted as a brake on the efficiency of resource use and has constrained reforms in other areas” (Yusuf, Nabeshina and Perkins 2006, p.5). Consequently, Chinese state leaders must garner the political will for continuous reforms if they seek to rejuvenate China’s slowed economic growth.
| Table 2 | Impact of access in finance on firm productivity | Direct effects | R&D | CEO | CFO | CEO-CFO
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital</td>
<td>$0.27*** ($0.09** *)</td>
<td>$0.29*** ($0.09**)</td>
<td>$0.26*** ($0.09**)</td>
<td>$0.29*** ($0.09**)</td>
<td>$0.28*** ($0.09**)</td>
</tr>
<tr>
<td>R&amp;D Spending</td>
<td>$0.18*** ($0.09** *)</td>
<td>$0.18*** ($0.09** *)</td>
<td>$0.18*** ($0.09** *)</td>
<td>$0.18*** ($0.09** *)</td>
<td>$0.18*** ($0.09** *)</td>
</tr>
<tr>
<td>CEO</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
</tr>
<tr>
<td>CFO</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
</tr>
</tbody>
</table>

Note: Standard errors in parentheses.

** p < 0.01; * p < 0.05; ** p < 0.1

---

| Table 2 | Impact of access in finance on firm productivity | Direct effects | R&D | CEO | CFO | CEO-CFO
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital</td>
<td>$0.27*** ($0.09** *)</td>
<td>$0.29*** ($0.09** *)</td>
<td>$0.26*** ($0.09** *)</td>
<td>$0.29*** ($0.09** *)</td>
<td>$0.28*** ($0.09** *)</td>
</tr>
<tr>
<td>R&amp;D Spending</td>
<td>$0.18*** ($0.09** *)</td>
<td>$0.18*** ($0.09** *)</td>
<td>$0.18*** ($0.09** *)</td>
<td>$0.18*** ($0.09** *)</td>
<td>$0.18*** ($0.09** *)</td>
</tr>
<tr>
<td>CEO</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
</tr>
<tr>
<td>CFO</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
<td>$0.07** ($0.05**)</td>
</tr>
</tbody>
</table>

Note: Standard errors in parentheses.

** p < 0.01; * p < 0.05; ** p < 0.1
7 Bibliography


LIBERAL OBJECTIVES: DOES THE IDEOLOGICAL DISTANCE BETWEEN THE UNITED STATES AND DEVELOPING COUNTRIES AFFECT HOW MUCH AID THEY RECEIVE?

CHRIS STROMEYER
Georgetown University

ABSTRACT
Does the ideological distance between a donor and recipient country of aid affect how much aid is committed? This paper uses a Latin America dataset of U.S. Bilateral, World Bank and Inter-American Development Bank aid data between 1975 and 2012 to determine whether the economic ideology of a government, as measured on a left to right spectrum, affects how much aid is committed. Furthermore, a country’s commitment to political rights and civil liberties, another strong ideological position of the United States, is also considered. We find that “leftist” and “unfree” governments do not systematically receive less bilateral aid, but that “leftist” countries which a low regard for political rights and civil liberties are allocated substantially less. Furthermore, the political ideology of the U.S. Executive and Legislative branches has no effect on how aid is allocated. Lastly, the World Bank is found to be a better promoter of U.S. interests than the Inter-American Development Bank.

1 Introduction
Ostensibly, the sole aim of foreign aid is to promote development. Despite this, the literature is rich with anecdotal and empirical evidence that political aims have influenced aid allocation as well. Even if the United States views capitalism, human rights and democracy as means to development, aid committed on those grounds rather than according to direct need indicates a measure of political influence.

This paper will thus ask whether the ideological distance between the United States and recipient countries affects how much aid is committed to those countries. Ideology will be measured in two particular ways. The first is according to the economic ideology, operationalized on a “left” to “right” spectrum, which proxies for the adherence to free markets. The second, in an
attempt to measure democracy and human rights, is measured in terms of political rights and civil liberties. Ideological distance can thus increase in two ways: (a) if a particular recipient country regime moves further away from the United States, by, for example, shifting to the left or developing more autocratic means of governance and (b) if the ruling government in the United States shifts to the right. The literature has explored a number of political determinants of foreign aid, ranging from vote buying for the UN Security Council and General Assembly to the promotion of commercial interests. Despite this, there has been no study focused on measuring ideological distance with a focus on the domestic politics of donor and recipient countries.

The study is focused to Latin America in the years 1975 to 2012. There are two primary reasons for this geographic restriction. The first is the long history of concern the United States has had for the region. Beginning with the Monroe Doctrine in 1823, which established a sphere of influence in the continent, the United States has taken an active role in the region’s domestic politics. There is thus reason to posit that if ideological distance matters in foreign aid decisions anywhere, the effect should be strongest in Latin America. Second, the region has a long history of political participation and has thus had a number of political transitions since 1975 both between democracies and autocracies and between “leftist” and “rightist” regimes. No other region in the world demonstrates such variability within countries during this period. As a result, the coverage of the Database of Political Institutions (Beck et al. 2013), which I rely on to code regimes as “leftist” or “rightist”, is particularly extensive in Latin America.2

Two types of aid will be considered. The main focus will be on bilateral aid allocated by the U.S. government directly to recipient countries. Despite this, bilateral aid is not the only tool the United States has to promote its liberal objectives. Multilateral aid has become an integral part of the international system since the Bretton Woods Conference in 1944. I therefore also consider aid from the World Bank, which was conceived at Bretton Woods, and the Inter-American Development Bank (IADB), which was established in the Bretton Woods spirit in 1959. The literature is ripe with examples of U.S. influence over these institutions, both because they embody the liberal objectives of American foreign policy and because they are frequently considered to be liberal institutions.

1 This study will presume that “rightist” (ie. Republican) governments in the United States are more aligned with the liberal ideals of capitalism, democracy and human rights and are thus more fervent supporters of them.

2 The many limitations of this dataset, and reasons why they are best mitigated when we focus only on Latin America, will be further discussed in Section IV.

attempt to measure democracy and human rights, is measured in terms of political rights and civil liberties. Ideological distance can thus increase in two ways: (a) if a particular recipient country regime moves further away from the United States, by, for example, shifting to the left or developing more autocratic means of governance and (b) if the ruling government in the United States shifts to the right. The literature has explored a number of political determinants of foreign aid, ranging from vote buying for the UN Security Council and General Assembly to the promotion of commercial interests. Despite this, there has been no study focused on measuring ideological distance with a focus on the domestic politics of donor and recipient countries.

The study is focused to Latin America in the years 1975 to 2012. There are two primary reasons for this geographic restriction. The first is the long history of concern the United States has had for the region. Beginning with the Monroe Doctrine in 1823, which established a sphere of influence in the continent, the United States has taken an active role in the region’s domestic politics. There is thus reason to posit that if ideological distance matters in foreign aid decisions anywhere, the effect should be strongest in Latin America. Second, the region has a long history of political participation and has thus had a number of political transitions since 1975 both between democracies and autocracies and between “leftist” and “rightist” regimes. No other region in the world demonstrates such variability within countries during this period. As a result, the coverage of the Database of Political Institutions (Beck et al. 2013), which I rely on to code regimes as “leftist” or “rightist”, is particularly extensive in Latin America.

Two types of aid will be considered. The main focus will be on bilateral aid allocated by the U.S. government directly to recipient countries. Despite this, bilateral aid is not the only tool the United States has to promote its liberal objectives. Multilateral aid has become an integral part of the international system since the Bretton Woods Conference in 1944. I therefore also consider aid from the World Bank, which was conceived at Bretton Woods, and the Inter-American Development Bank (IADB), which was established in the Bretton Woods spirit in 1959. The literature is ripe with examples of U.S. influence over these institutions, both because they embody the liberal objectives of American foreign policy and because they are frequently considered to be liberal institutions.

1 This study will presume that “rightist” (ie. Republican) governments in the United States are more aligned with the liberal ideals of capitalism, democracy and human rights and are thus more fervent supporters of them.

2 The many limitations of this dataset, and reasons why they are best mitigated when we focus only on Latin America, will be further discussed in Section IV.
spirit of the international system established Bretton Woods and due to the large voting share the U.S. holds at both institutions. I find tepid evidence that ideological distance matters in U.S. aid allocation decisions. In particular, “leftist” governments and those who pay little attention to political rights and civil liberties are not systematically allocated less bilateral aid. Despite this, there is some evidence that when a government is both “leftist” and “unfree” it is allocated considerably less aid. The World Bank, on the other hand, has systematically allocated less aid to both “leftist” and “unfree” regimes. This contrasts with the Inter-American Development Bank, which does not seem to make consistent decisions according to the “left” and “unfree” criteria. This result reinforces the emerging literature showing that the U.S. has more influence over the former institution than the latter. In contrast, shifts to the “right” in the United States have no measurable effect on the allocation of bilateral aid. This lends credence to the notion that Republicans and Democrats follow a similar foreign policy.

2 Empirical Strategy

2.1 Estimation Method

There is considerable debate in the literature about how to best model aid data (Berthelemy 2006). The crux of the problem is that aid data often has a lower-bound corner solution, since in a number of country-year observations no aid is allocated. This raises a particular problem for linear models since a 0 observation could be an indication of an intention to allocate “negative aid”. The corner solution is thus hiding information. The corner solution problem is evident in reference to all three donors, but is particularly stark in the case of the World Bank, where 35% of observations are clustered at 0. My data thus seems a ripe example for the use of a Tobit model, a non-linear limited dependent variable model. Nevertheless, the Tobit model has biased estimators when fixed effects are introduced (Berthelemy 2006). This is a particular problem for our study, since the countries in our dataset suffer
from a high degree of heterogeneity. Fixed effects are thus a natural test in order to take unobserved recipient specificities into account. This intuition was validated when a Hausman test resoundingly rejected the null hypothesis that a random effects model was a plausible alternative to a fixed effects one.

Since both fixed effects linear models and random effect Tobit models are employed frequently in the literature, I will present results from both while remaining cognizant that each has its own limitations. Following Kilby (2006) my fixed effects linear model will allow for a common first order autoregressive, AR(1), process across panels to correct for possible autocorrelation in the data.

2.2 Hypothesis $H_3$ Specification

The first hypothesis, $H_3$, will be modeled with the following specification:

$$
\ln \left( \frac{Aid_{it}}{Pop_{it}} \right) = \alpha + \beta_1 Left_{it} + \beta_2 Unfree_{it} + \beta_3 (Left_{it} \times Unfree_{it-1}) + \beta_4 \ln \left( \frac{Aid_{it-1}}{Pop_{it-1}} \right) + E_{it} + S_{it} + C_{it} + P_{it} + \gamma_1 + \delta_t + \epsilon_{it}
$$

Where $Left$ is the dummy for the instances in which recipient country $i$ in year $t$ is coded as “leftist” and $Unfree$, which can represent either $Unfree_1$ or $Unfree_2$, is a dummy for regimes with low regard for political and civil liberties. Due to differences in when the $Left$ and $Unfree$ variables are

---

6 Another model which makes a frequent appearance in the literature is a two-step one, where the selection equation is separated from the allocation one. This method suffers from the grave risk of introducing a selection bias in the second step, since there might be unobserved factors that affect both whether one receives aid and how much aid one receives.

7 There are a number of different results one can pull from the Tobit model. There is considerable debate in the literature about which model is appropriate under what circumstances (Wooldridge 2009). This paper will use the latent dependent variable, which allows for negative values when fitted. Choosing this outcome means that marginal effects can be interpreted the same way as those of OLS regressions.

8 The main reason to expect autocorrelation in the data is institutional budgeting creates inertia for bureaucratic reasons, leading one to believe that the aid a country receives one year is not independent from that which it received the previous year (Kilby 2006).

9 Honoré (1992) developed a fixed effects estimator for semiparametric Tobit models. I tested the regression software add-on he created for this model on a number of my regressions and found the model to be surprisingly internally inconsistent and very sensitive to slight changes in the inclusion of variables. I thus decided not to pursue analysis with this model.

---

from a high degree of heterogeneity. Fixed effects are thus a natural test in order to take unobserved recipient specificities into account. This intuition was validated when a Hausman test resoundingly rejected the null hypothesis that a random effects model was a plausible alternative to a fixed effects one.

Since both fixed effects linear models and random effect Tobit models are employed frequently in the literature, I will present results from both while remaining cognizant that each has its own limitations. Following Kilby (2006) my fixed effects linear model will allow for a common first order autoregressive, AR(1), process across panels to correct for possible autocorrelation in the data. My Tobit model, on the other hand, will attempt to partly correct for the same problem by including a lagged dependent variable and clustering standard errors at the country level.

2.2 Hypothesis $H_3$ Specification

The first hypothesis, $H_3$, will be modeled with the following specification:

$$
\ln \left( \frac{Aid_{it}}{Pop_{it}} \right) = \alpha + \beta_1 Left_{it} + \beta_2 Unfree_{it} + \beta_3 (Left_{it} \times Unfree_{it-1}) + \beta_4 \ln \left( \frac{Aid_{it-1}}{Pop_{it-1}} \right) + E_{it} + S_{it} + C_{it} + P_{it} + \gamma_1 + \delta_t + \epsilon_{it}
$$

Where $Left$ is the dummy for the instances in which recipient country $i$ in year $t$ is coded as “leftist” and $Unfree$, which can represent either $Unfree_1$ or $Unfree_2$, is a dummy for regimes with low regard for political and civil liberties. Due to differences in when the $Left$ and $Unfree$ variables are
measured, the latter is lagged by one period\textsuperscript{10}. These two variables are interacted to see whether there is a larger effect when a regime is both “leftist” and not “free”. $\ln(\text{Aid}_{it-1})$ represents the lagged dependent variable which only enters the Tobit model. On the other hand, $\gamma$ represents country fixed effects, which are included in order to control for heterogeneous conditions in countries or between the donor and recipient government that are not already covered by other controls. In a similar vein, $b$ represents time fixed effects, which are particularly important because they control for the budget of bilateral and multilateral aid in any given year. These are only present in the linear specification of the model.

E, S, C and P are all vectors of controls. The first two represent the needs (“loan demand”) of recipient countries, with E catching any economic or developmental needs country $i$ might have in year $t$ and S being a vector of substitute options to U.S. bilateral and multilateral influenced aid for these countries in any given year. In addition, C and P represent any commercial and political interests the U.S. might have in recipient countries. It is worth exploring each of these vectors in turn.

The economic needs vector ($E$) is the most extensive and diverse. It has become standard in the literature to include population, GDP per capita and, to a lesser extent, infant mortality rates for each country in each year. I will do the same as they each capture a different dimension of need. I will also add to the literature by including dummies for years in which a natural disaster of significant magnitude occurs. The substitutes vector ($S$) allows for the lending institutions not presently being examined, and for other lenders, to enter as either substitutes or complements for the recipient country. In essence, they are a measure of loan demand by recipient countries. The consensus in the literature seems to be that multilateral aid enters as a complement, as bilateral aid agencies allow them to do the due diligence on countries and/or projects, while bilateral aid enters as a substitute.

Commercial interests ($C$) have most often been represented by U.S. exports to a certain country in a given year. Some have disaggregated it further, showing that the positive effect is strongest in countries to which the U.S. exports capital goods (Younas 2008). I will maintain the more commonly used basic measure.

The political interests vector ($P$) controls for other political factors relevant to the United States. I therefore include a dummy for years in which a certain disaster of significant magnitude occurs. The substitutes vector ($S$) allows for the lending institutions not presently being examined, and for other lenders, to enter as either substitutes or complements for the recipient country. In essence, they are a measure of loan demand by recipient countries. The consensus in the literature seems to be that multilateral aid enters as a complement, as bilateral aid agencies allow them to do the due diligence on countries and/or projects, while bilateral aid enters as a substitute.

Commercial interests ($C$) have most often been represented by U.S. exports to a certain country in a given year. Some have disaggregated it further, showing that the positive effect is strongest in countries to which the U.S. exports capital goods (Younas 2008). I will maintain the more commonly used basic measure.

The political interests vector ($P$) controls for other political factors relevant to the United States. I therefore include a dummy for years in which a certain disaster of significant magnitude occurs. The substitutes vector ($S$) allows for the lending institutions not presently being examined, and for other lenders, to enter as either substitutes or complements for the recipient country. In essence, they are a measure of loan demand by recipient countries. The consensus in the literature seems to be that multilateral aid enters as a complement, as bilateral aid agencies allow them to do the due diligence on countries and/or projects, while bilateral aid enters as a substitute.

Commercial interests ($C$) have most often been represented by U.S. exports to a certain country in a given year. Some have disaggregated it further, showing that the positive effect is strongest in countries to which the U.S. exports capital goods (Younas 2008). I will maintain the more commonly used basic measure.

The political interests vector ($P$) controls for other political factors relevant to the United States. I therefore include a dummy for years in which a certain disaster of significant magnitude occurs. The substitutes vector ($S$) allows for the lending institutions not presently being examined, and for other lenders, to enter as either substitutes or complements for the recipient country. In essence, they are a measure of loan demand by recipient countries. The consensus in the literature seems to be that multilateral aid enters as a complement, as bilateral aid agencies allow them to do the due diligence on countries and/or projects, while bilateral aid enters as a substitute.

Commercial interests ($C$) have most often been represented by U.S. exports to a certain country in a given year. Some have disaggregated it further, showing that the positive effect is strongest in countries to which the U.S. exports capital goods (Younas 2008). I will maintain the more commonly used basic measure.

The political interests vector ($P$) controls for other political factors relevant to the United States. I therefore include a dummy for years in which a certain disaster of significant magnitude occurs. The substitutes vector ($S$) allows for the lending institutions not presently being examined, and for other lenders, to enter as either substitutes or complements for the recipient country. In essence, they are a measure of loan demand by recipient countries. The consensus in the literature seems to be that multilateral aid enters as a complement, as bilateral aid agencies allow them to do the due diligence on countries and/or projects, while bilateral aid enters as a substitute.

Commercial interests ($C$) have most often been represented by U.S. exports to a certain country in a given year. Some have disaggregated it further, showing that the positive effect is strongest in countries to which the U.S. exports capital goods (Younas 2008). I will maintain the more commonly used basic measure.

The political interests vector ($P$) controls for other political factors relevant to the United States. I therefore include a dummy for years in which a certain disaster of significant magnitude occurs. The substitutes vector ($S$) allows for the lending institutions not presently being examined, and for other lenders, to enter as either substitutes or complements for the recipient country. In essence, they are a measure of loan demand by recipient countries. The consensus in the literature seems to be that multilateral aid enters as a complement, as bilateral aid agencies allow them to do the due diligence on countries and/or projects, while bilateral aid enters as a substitute.

Commercial interests ($C$) have most often been represented by U.S. exports to a certain country in a given year. Some have disaggregated it further, showing that the positive effect is strongest in countries to which the U.S. exports capital goods (Younas 2008). I will maintain the more commonly used basic measure.

The political interests vector ($P$) controls for other political factors relevant to the United States. I therefore include a dummy for years in which a certain disaster of significant magnitude occurs. The substitutes vector ($S$) allows for the lending institutions not presently being examined, and for other lenders, to enter as either substitutes or complements for the recipient country. In essence, they are a measure of loan demand by recipient countries. The consensus in the literature seems to be that multilateral aid enters as a complement, as bilateral aid agencies allow them to do the due diligence on countries and/or projects, while bilateral aid enters as a substitute.

Commercial interests ($C$) have most often been represented by U.S. exports to a certain country in a given year. Some have disaggregated it further, showing that the positive effect is strongest in countries to which the U.S. exports capital goods (Younas 2008). I will maintain the more commonly used basic measure.

The political interests vector ($P$) controls for other political factors relevant to the United States. I therefore include a dummy for years in which a certain disaster of significant magnitude occurs. The substitutes vector ($S$) allows for the lending institutions not presently being examined, and for other lenders, to enter as either substitutes or complements for the recipient country. In essence, they are a measure of loan demand by recipient countries. The consensus in the literature seems to be that multilateral aid enters as a complement, as bilateral aid agencies allow them to do the due diligence on countries and/or projects, while bilateral aid enters as a substitute.
country serves as a temporary member on the UNSC, an exogenous factor that has been widely linked to higher aid in the literature.

2.3 Hypothesis H Specification

The second hypothesis, $H_h$, will be modeled with the following specification:

$$
\ln\left( \frac{A_{id_h}}{Pop_{it}} \right) = \alpha + \beta_1 \text{Left}_{it} + \beta_2 \text{Unfree}_{it-1} + \beta_3 \text{R Exec}_{it} + \beta_4 \text{R Congress}_{it} \\
+ \beta_5 (\text{Left}_{it} \times \text{R Exec}_{it}) \\
+ \beta_6 (\text{Left}_{it} \times \text{R Exec} \times \text{R Congress}_{it}) \\
+ \beta_7 (\text{Unfree}_{it-1} \times \text{R Exec}) \\
+ \beta_8 (\text{Unfree}_{it-1} \times \text{R Exec} \times \text{R Congress}_{it}) \\
+ \beta_9 (\text{Left}_{it} \times \text{Unfree}_{it-1} \times \text{R Exec} \times \text{R Congress}_{it}) \\
+ \beta_{10} \ln\left( \frac{A_{id-1}}{Pop_{it-1}} \right) + E_{it} + S_{it} + C_{it} + P_{it} + \gamma_t + \delta_t + \varepsilon_{it}
$$

Where the only additions are $R_\text{Exec}$ and $R_\text{Congress}$, variables measuring ideological shifts in year $t$ in the U.S. The interactions terms will allow us to interpret whether a shift further to the right by the U.S., represented by Republican control of government, affects how much aid is committed to “leftist” and/or “unfree” countries.

3 Results

3.1 Hypothesis A – Recipient Country Ideological Shifts

Table 1 shows the results for the $H_A$ specification, which tests whether “leftist” and/or “unfree” regimes receive less aid in per capita terms. I find no statistically significant evidence that either “leftist” or “unfree” regimes receive systematically less aid. This holds even when the definition of “unfree” is restricted to those countries which Freedom House views as being “Not Free” (Unfree_2). Coefficients are consistently negative but they do not reach conventional levels of significance. I do find some evidence that regimes that are both “leftist” and “unfree” under the more stringent classification receive considerably less aid than those which are “rightist” and “free”. The magnitude of this effect is very high. Column 4 shows that these countries see a decrease in aid of over 160% relative to the control group. This means that there is an overall intent by the U.S. to, ceteris paribus, grant “negative aid” to these countries. The same interaction term in country serves as a temporary member on the UNSC, an exogenous factor that has been widely linked to higher aid in the literature.
the Tobit model (column 8) has an even larger magnitude, but it just fails to achieve significance at the 10% level.\textsuperscript{11}

A few comments are in order about the control variables as well. \textit{GDP per Capita} is significant and takes on its expected negative sign, although significance is much larger under the Tobit specifications. In general, a 1% increase in GDP per capita is associated with a 1% decrease in aid. The \textit{Infant Mortality Rate} is surprisingly negative, but is not significant.\textsuperscript{12} Meanwhile the \textit{Natural Disaster} dummy is positive and highly significant in all Tobit specifications. \textit{U.S. Trade}, on the other hand, is not significant in either model.\textsuperscript{13} Furthermore, there is weak evidence that bilateral and IADB aid are complements, but no relationship of note with World Bank aid. The two most surprising results are those of the \textit{Population} and \textit{UNSC} variables. The former is positive and significant under the linear model but negative and insignificant in the Tobit model. This contradicts theory and most of the literature, which usually finds that when the dependent variable is measured in per capita terms, there is a bias towards smaller countries since donors see more value in giving more aid to them in order to demonstrate a presence, rather than trying to achieve an equitable distribution in per capita terms. A possible reason for this is that the fixed effects regression is simply picking up population growth over time. The highly significant values on the \textit{UNSC} variable strongly contradict what has been found in the literature (Kuziemko & Werker 2006; Dreher & Vreeland 2014). There is a plausible explanation for this: Dreher & Vreeland (2014) find that the Latin American region has the least turn-taking for UNSC seats of any region in the world. Instead, the temporary spots most often go to the richest countries. The countries which have spent the most time on the council in our time period of interest are Brazil (10 years), Argentina (8 years) and Colombia, Mexico, Panama and Venezuela each with 6 years. Together this constitutes 55% of all available UNSC year-seats from 1975 to 2012. Since these countries also receive temporary spots most often go to the richest countries. The countries which have spent the most time on the council in our time period of interest are Brazil (10 years), Argentina (8 years) and Colombia, Mexico, Panama and Venezuela each with 6 years. Together this constitutes 55% of all available UNSC year-seats from 1975 to 2012. Since these countries also receive

\textsuperscript{11}Due to the bias inherent in both the linear and Tobit models, due to the corner solution in the former and the lack of fixed effects in the latter, means that we need to interpret magnitudes of coefficients with extreme caution.

\textsuperscript{12}Even though it fails to attain significance on its own, it is jointly significant with other non-significant variables in some of the specifications and was therefore not dropped from the regression. In general, F-tests were conducted on all non-significant variables and I found that each of them is at least jointly significant with one of the other non-significant variables.

\textsuperscript{13}\textit{U.S. Trade} does achieve significance if we drop our two substitute variables from the regression, \textit{IADB Aid Per Capita} and \textit{World Bank Aid Per Capita}. This indicates that there might be high collinearity between these variables. Our results are generally robust when we exclude these two variables, which also has the added benefit of adding another 4 countries to our sample.
much less aid in per capita terms, the result might simply be driven by sample bias. Still, the country fixed effects should be controlling for this. However, Argentina and Brazil have in recent years received more aid than the negligible amounts they were receiving historically. Since most of their UNSC terms occurred in earlier years, the country fixed effects might not be fully correcting for the sample bias.

A general comparison of the linear AR(1) fixed effects model with the random effects Tobit model renders a few insights. Significance is in general much lower in the linear fixed effects model, which is not surprising for two reasons. First, a number of the control variables are relatively time-invariant. Variables such as GDP per Capita are thus not surprisingly more significant in the Tobit model. Second, some of the variables, such as Natural Disaster, are only significant in their effect on aid when compared across countries and not within them, as a random effects model does. Given these explanations, one can be reasonably confident that both models are providing useful and consistent results for our analysis.

For the World Bank, the control variables are generally more consistent between the linear and Tobit models. A notable difference from the bilateral results is that the Natural Disaster variable is now consistently positive and significant, which indicates a larger focus on natural disasters by the MDB. Meanwhile, the UNSC variable is positive yet insignificant. In addition, the World Bank has less of a bias against richer countries than the U.S. with its bilateral aid. This further validates our previous conclusion.

---

14 The lack of fixed effects in the Tobit model is likely the reason for the higher significance under those specifications.
15 Column 8 also has a highly significant interaction effect, but even this does not change the significance of the Unfree measure.
16 This is likely the case because it has two lending arms that we consider as developmental aid. While the United States has to carefully allocate its grant budget, the World Bank can safely lend to rich countries such as Argentina and Brazil through the IBRD with the expectation that it will receive its money back.

---

14 The lack of fixed effects in the Tobit model is likely the reason for the higher significance under those specifications.
15 Column 8 also has a highly significant interaction effect, but even this does not change the significance of the Unfree measure.
16 This is likely the case because it has two lending arms that we consider as developmental aid. While the United States has to carefully allocate its grant budget, the World Bank can safely lend to rich countries such as Argentina and Brazil through the IBRD with the expectation that it will receive its money back.
that the surprising result regarding the UNSC variable was likely due to sample bias. Lastly, it is evident that World Bank and IADB aid are strong complements, with an increase in one leading to a 0.4% to 0.7% increase in the other.

As a robustness check for all three donors, our Unfree variables derived from Freedom House are replaced with an Autocracy variable from the PolityIV project (Gurr & Marshall 2013). They code all countries on a range from -10 (most autocratic) to +10 (most democratic). Any country with a score at or below -6 is considered an Autocracy. The results for bilateral aid confirm what has been found in the literature so far, namely that autocracies receive less aid, but do not fully corroborate our results. The most important difference is that the Autocracy variable is found to be consistently negative and significant even when not interacted with Left. This lends credence to the notion that political and civil liberties are a broader measure than just “democracy” and that the U.S. does not treat the two as the same. The results for the Multilateral Development Banks are even further from those of the Unfree measure: in some specifications the World Bank and IADB allocate more aid to autocracies. It could thus be posited that the MDB’s are more interested in broader measures of political and civil rights than in shallow measures of democracy. This short discussion demonstrates that the Freedom House and PolityIV indices should not be used interchangeably in the literature, but rather one index should be chosen for deliberate reasons, as this paper has done.

that the surprising result regarding the UNSC variable was likely due to sample bias. Lastly, it is evident that World Bank and IADB aid are strong complements, with an increase in one leading to a 0.4% to 0.7% increase in the other.

In the analysis for the Inter-American Development Bank, significant results for the main explanatory variables of interest only appear under Tobit model specifications, which casts doubts on the robustness of the conclusions. The results seem to indicate that regimes that are Unfree do not receive less aid, and in one instance receive more aid at standard significance levels, but that regimes that are both Unfree and leftist, as measured by the interaction term, receive significantly less aid. In this way, these results are most similar to the bilateral conclusions, except the latter did not find any evidence that being Unfree increases aid.

The one surprising control variable is US Trade, which is both positive and significant. Nevertheless, this result is not robust across all specifications. In addition, the UNSC variable is negative and insignificant. This provides tepid evidence that the World Bank loans are more beholden to temporary membership of the UNSC, as concluded by Dreher & Vreeland (2014), and indicates that our initial surprising result for bilateral aid was due to sample bias.

As a robustness check for all three donors, our Unfree variables derived from Freedom House are replaced with an Autocracy variable from the PolityIV project (Gurr & Marshall 2013). They code all countries on a range from -10 (most autocratic) to +10 (most democratic). Any country with a score at or below -6 is considered an Autocracy. The results for bilateral aid confirm what has been found in the literature so far, namely that autocracies receive less aid, but do not fully corroborate our results. The most important difference is that the Autocracy variable is found to be consistently negative and significant even when not interacted with Left. This lends credence to the notion that political and civil liberties are a broader measure than just “democracy” and that the U.S. does not treat the two as the same. The results for the Multilateral Development Banks are even further from those of the Unfree measure: in some specifications the World Bank and IADB allocate more aid to autocracies. It could thus be posited that the MDB’s are more interested in broader measures of political and civil rights than in shallow measures of democracy. This short discussion demonstrates that the Freedom House and PolityIV indices should not be used interchangeably in the literature, but rather one index should be chosen for deliberate reasons, as this paper has done.

In the analysis for the Inter-American Development Bank, significant results for the main explanatory variables of interest only appear under Tobit model specifications, which casts doubts on the robustness of the conclusions. The results seem to indicate that regimes that are Unfree do not receive less aid, and in one instance receive more aid at standard significance levels, but that regimes that are both Unfree and leftist, as measured by the interaction term, receive significantly less aid. In this way, these results are most similar to the bilateral conclusions, except the latter did not find any evidence that being Unfree increases aid.

The one surprising control variable is US Trade, which is both positive and significant. Nevertheless, this result is not robust across all specifications. In addition, the UNSC variable is negative and insignificant. This provides tepid evidence that the World Bank loans are more beholden to temporary membership of the UNSC, as concluded by Dreher & Vreeland (2014), and indicates that our initial surprising result for bilateral aid was due to sample bias.

As a robustness check for all three donors, our Unfree variables derived from Freedom House are replaced with an Autocracy variable from the PolityIV project (Gurr & Marshall 2013). They code all countries on a range from -10 (most autocratic) to +10 (most democratic). Any country with a score at or below -6 is considered an Autocracy. The results for bilateral aid confirm what has been found in the literature so far, namely that autocracies receive less aid, but do not fully corroborate our results. The most important difference is that the Autocracy variable is found to be consistently negative and significant even when not interacted with Left. This lends credence to the notion that political and civil liberties are a broader measure than just “democracy” and that the U.S. does not treat the two as the same. The results for the Multilateral Development Banks are even further from those of the Unfree measure: in some specifications the World Bank and IADB allocate more aid to autocracies. It could thus be posited that the MDB’s are more interested in broader measures of political and civil rights than in shallow measures of democracy. This short discussion demonstrates that the Freedom House and PolityIV indices should not be used interchangeably in the literature, but rather one index should be chosen for deliberate reasons, as this paper has done.
Overall, there is some evidence across all three donors that when recipient countries shift farther away ideologically from the U.S., they are allocated less aid. Sometimes, as in the case of bilateral and IADB aid, this is only true when a regime is both “leftist” and has a low regard for political rights and civil liberties. In the case of the World Bank, on the other hand, only satisfying one of those conditions is enough. The results are generally the most consistent in the World Bank case, indicating that this Bretton Woods institution has been able to make aid decisions according, in part, to the liberal ideals it was founded on.

3.2 Hypothesis B – United States Ideological Shifts

Having found some evidence that when ideological distance increases due to policy and regime shifts in recipient countries aid allocation decreases, I now turn my attention to shifts on the donor side. Table 2 shows the results for this analysis. The first result to note is that under the linear specification, our results corroborate what Younas (2008) found, namely that Republican dominated legislatures normally allocate less aid overall. However, I also find that Republican Presidents allocate more aid in the Tobit regressions. Although this could indeed be true, it might also be a result of the lack of year fixed effects in that regression. Overall, the results consistently show that ideological shifts in the United States, while possibly having an impact on the overall size of the aid budget, do not influence how the budget is allocated. The one exception to this result is found in column 4, where the interaction term of $R_{\text{Exec}}$ and $Left$ seems to indicate that when a Republican is in power, “leftist” governments receive less aid. Despite this, the result is not robust when I exclude the substitution variables and thereby consider the full dataset.\footnote{Similarly to the other regressions, all other estimators are robust to the exclusion of these variables.}

Although the results were omitted due to space considerations, I also tested whether a Republican Executive and/or Congress has an additive effect on the interaction terms that were significant in previous models, namely when a regime is both $\text{Unfree}$ and $Left$. The results were not conclusive nor statistically significant. My overall results are also robust to an alternative way of measuring $R_{\text{Congress}}$, namely as a dummy coded 1 when Republicans are in control of the House, Senate or both and 0 otherwise.

The lack of an impact of U.S. domestic ideology on the allocation of aid to specific recipient countries lends credence to the notion that, at least when it involves foreign policy, the Republican and Democratic parties share a largely common ideology.

17 Similarly to the other regressions, all other estimators are robust to the exclusion of these variables.
3.3 Endogeneity Issues

A number of endogeneity issues remain unresolved in this paper. It is worth considering each of them in turn in order to determine their importance and the bias, if any, they could have had on our results.

There could be a reverse causality problem between US Trade and Aid. This is particularly true due to the existence of tied aid, where aid is only disbursed conditional on an agreement that the recipient country use it to buy goods from the donor. There are two arguments that indicate why this risk is limited. The first is that tied aid has decreased significantly over the last two decades as a result of an OECD goal to eliminate it altogether (OECD 2013). Secondly, we measure aid in terms of commitments and not disbursements. Aid disbursements lag significantly behind commitments, particularly in the case of the project aid that is at the heart of the tied aid debate (Berthelemey 2006). There is thus a semi-random lag of when purchases, if any, of donor country goods are made compared to when a commitment to allocate aid was made. Although these two facts do not completely eliminate the possibility of an endogeneity problem, they demonstrate that the risk is not that large.

The second endogeneity issue arises from a possible simultaneity issue between aid and democracy, a large part of what is being measured by Freedom House’s political liberties and civil rights indices. In order to establish that this is a problem, one has to argue that aid is successful in promoting democracy in the first place. Two arguments can be made about the causal link between aid and democracy. The first is that general development assistance, by increasing a country’s income, creates the structural change necessary for democracy to flourish. Knack (2004) finds no evidence of this being the case in general. 18 A second argument relates to targeted democracy assistance, meant to empower individuals and groups fighting for democracy. The U.S. has been active in this since the early 1990s when U.S. Agency for International Development’s (USAID) Democracy and Governance Office were created (Finkel et al. 2007). Since then, USAID has spent money yearly on programs whose direct intent is to promote the establishment of democracy. There are currently 217 active programs in Latin America alone (USAID 2014). Finkel et al. (2007) found this kind of aid to be highly successful at promoting Democracy. This lends credence to the notion that there exists here an important simultaneity problem, even though the literature, which often controls for Democracy, has never addressed it. In order to determine what impact USAID’s democracy

---

18 This does not mean that it cannot have been the case in particular instances. A few exceptions to the rule would be enough to lead to a simultaneity problem.
office has had on aid, I split my dataset into two parts and ran the same regression for all years before 1990 and another for all observations including 1990 and since then. Unsurprisingly, although the results that Left and Unfree regimes receive less aid hold in both time periods, they are only significant before 1990. Whether this is a result of USAID’s democracy promotion programs or the Cold War is difficult to ascertain. In either case, my findings after splitting the dataset along with the nature of USAID’s work, seem to indicate that this simultaneity problem is biasing my results in the opposite direction of my findings. In other words, USAID’s program entails more aid for non-democratic countries, while my findings indicate that they receive less. This simultaneity problem thus remains important, but there is evidence to believe that my findings regarding the Unfree variable would hold even if it were resolved.

The last, and most vexing, endogeneity issue concerns reverse causality between the Left variable and aid. As stated in Section III, the U.S. views aid allocation as a foreign policy tool since it is able to encourage countries to remain ideologically aligned (a “carrot”) or punish them from straying from this ideal (a “stick”). In either of these scenarios there is potential for reverse causality, since it is unclear whether the U.S. is reacting to the political ideology of recipient countries or recipient countries are reaction to U.S. aid policy. In an attempt to overcome this issue, a number of instrumental variables were tested in an attempt to find a variable that is correlated with a country’s decision to go “left” yet uncorrelated with the decision to allocate aid\textsuperscript{19}. Unfortunately none of the instruments proved to be significantly correlated with Left to proceed with a full two-staged least squares regression. This paper thus follows the literature in finding it difficult to find instruments for political variables that are (1) correlated with the variable of interest, (2) uncorrelated with the dependent variable and (3) for which there is available data across the time period of interest.

4 Conclusions, Limitations and Possible Extensions

This study finds that when the ideological distance between the United States and a recipient country in Latin America increases, less bilateral aid is allocated to the country in question. However, in order for this phenomenon to kick in, the increase in ideological distance needs to be large, entailing both a regime shift towards the left and towards less regard for political rights and civil liberties. On the other hand, shifts to the right in the United States, measured as instances when Republicans are in power in the

---

\textsuperscript{19} Among the instruments examined were: unemployment rate, different measures of income inequality, the presence of financial crises and various permutations of lags of these variables.

---

office has had on aid, I split my dataset into two parts and ran the same regression for all years before 1990 and another for all observations including 1990 and since then. Unsurprisingly, although the results that Left and Unfree regimes receive less aid hold in both time periods, they are only significant before 1990. Whether this is a result of USAID’s democracy promotion programs or the Cold War is difficult to ascertain. In either case, my findings after splitting the dataset along with the nature of USAID’s work, seem to indicate that this simultaneity problem is biasing my results in the opposite direction of my findings. In other words, USAID’s program entails more aid for non-democratic countries, while my findings indicate that they receive less. This simultaneity problem thus remains important, but there is evidence to believe that my findings regarding the Unfree variable would hold even if it were resolved.

The last, and most vexing, endogeneity issue concerns reverse causality between the Left variable and aid. As stated in Section III, the U.S. views aid allocation as a foreign policy tool since it is able to encourage countries to remain ideologically aligned (a “carrot”) or punish them from straying from this ideal (a “stick”). In either of these scenarios there is potential for reverse causality, since it is unclear whether the U.S. is reacting to the political ideology of recipient countries or recipient countries are reaction to U.S. aid policy. In an attempt to overcome this issue, a number of instrumental variables were tested in an attempt to find a variable that is correlated with a country’s decision to go “left” yet uncorrelated with the decision to allocate aid\textsuperscript{19}. Unfortunately none of the instruments proved to be significantly correlated with Left to proceed with a full two-staged least squares regression. This paper thus follows the literature in finding it difficult to find instruments for political variables that are (1) correlated with the variable of interest, (2) uncorrelated with the dependent variable and (3) for which there is available data across the time period of interest.

4 Conclusions, Limitations and Possible Extensions

This study finds that when the ideological distance between the United States and a recipient country in Latin America increases, less bilateral aid is allocated to the country in question. However, in order for this phenomenon to kick in, the increase in ideological distance needs to be large, entailing both a regime shift towards the left and towards less regard for political rights and civil liberties. On the other hand, shifts to the right in the United States, measured as instances when Republicans are in power in the

---

\textsuperscript{19} Among the instruments examined were: unemployment rate, different measures of income inequality, the presence of financial crises and various permutations of lags of these variables.
Executive and Legislative branches, while theoretically increasing ideological distance, do not decrease the amount of aid committed. The study was also extended to the World Bank and the Inter-American Development Bank. The former was found to have an even stronger preference for "liberal ideals" than the United States, while the results for the latter corroborated recent literature that the IADB is an exception among MDBs in that it is not beholden to its largest donor (Hernandez 2013).

Although these findings paint a somewhat consistent picture, there are three major factors that limit the conclusions that can be drawn. The first relates to the dataset itself. The Database of Political Institutions codes all governments compared to others across all countries in the world. There is reason to believe that the United States is more interested in shifts to the "left" or "right" relative to previous governments in the same country. Our dataset therefore understates political transitions and limits the variability of a key explanatory variable of interest. Second, the literature has yet to agree on a good way to model aid data and overcome the corner solution problem. This paper relied on a linear model with fixed effects and a random effects Tobit model; both of these modelling approaches have inherent limitations that lead to biased results. Lastly, a number of endogeneity issues remain unresolved, particularly related to the crucial measures of democracy and political ideology. The difficulty of finding instrumental variables for political factors is a major limitation to the conclusions that can be drawn.

Given these three limitations, and in particular the latter two, the results from this paper should be interpreted as correlations rather than direct causations.

The study of the political determinants of foreign aid is ripe with opportunities for further research. The first priority should be the construction of a more extensive and complete database that codes regimes in developing countries across the world. A reliable dataset that does not suffer from sample bias would promote a flurry of fascinating research. Second, more work needs to be done to establish causality between political/commercial variables and aid decisions. Until now, only membership of the UNSC has been demonstrated to be truly exogenous. Much work is left to be done to find reliable instruments, or other methods, to determine causality for variables such as democracy, trade, and political ideology. Aid remains a powerful tool, both as a means to development and as a political instrument for large donors. Understanding the interplay between these two objectives remains a challenge but a worthwhile area of further study.

Although these findings paint a somewhat consistent picture, there are three major factors that limit the conclusions that can be drawn. The first relates to the dataset itself. The Database of Political Institutions codes all governments compared to others across all countries in the world. There is reason to believe that the United States is more interested in shifts to the "left" or "right" relative to previous governments in the same country. Our dataset therefore understates political transitions and limits the variability of a key explanatory variable of interest. Second, the literature has yet to agree on a good way to model aid data and overcome the corner solution problem. This paper relied on a linear model with fixed effects and a random effects Tobit model; both of these modelling approaches have inherent limitations that lead to biased results. Lastly, a number of endogeneity issues remain unresolved, particularly related to the crucial measures of democracy and political ideology. The difficulty of finding instrumental variables for political factors is a major limitation to the conclusions that can be drawn. Given these three limitations, and in particular the latter two, the results from this paper should be interpreted as correlations rather than direct causations.

The study of the political determinants of foreign aid is ripe with opportunities for further research. The first priority should be the construction of a more extensive and complete database that codes regimes in developing countries across the world. A reliable dataset that does not suffer from sample bias would promote a flurry of fascinating research. Second, more work needs to be done to establish causality between political/commercial variables and aid decisions. Until now, only membership of the UNSC has been demonstrated to be truly exogenous. Much work is left to be done to find reliable instruments, or other methods, to determine causality for variables such as democracy, trade, and political ideology. Aid remains a powerful tool, both as a means to development and as a political instrument for large donors. Understanding the interplay between these two objectives remains a challenge but a worthwhile area of further study.
<table>
<thead>
<tr>
<th>Model</th>
<th>Linear</th>
<th>Linear</th>
<th>Linear</th>
<th>Linear</th>
<th>Tobit</th>
<th>Tobit</th>
<th>Tobit</th>
<th>Tobit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>-0.0463</td>
<td>-0.0380</td>
<td>-0.104</td>
<td>0.00571</td>
<td>0.0169</td>
<td>-0.00565</td>
<td>0.0333</td>
<td>0.0522</td>
</tr>
<tr>
<td>(0.226)</td>
<td>(0.226)</td>
<td>(0.233)</td>
<td>(0.212)</td>
<td>(0.226)</td>
<td>(0.125)</td>
<td>(0.189)</td>
<td>(0.125)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>L.Unfree_1</td>
<td>-0.374</td>
<td>-0.427</td>
<td>-0.219</td>
<td>-0.206</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.239)</td>
<td>(0.264)</td>
<td>(0.168)</td>
<td>(0.191)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L.Unfree_2</td>
<td>-0.421</td>
<td>-0.0520</td>
<td>-0.614</td>
<td>-0.356</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.361)</td>
<td>(0.399)</td>
<td>(0.546)</td>
<td>(0.479)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LeftXL.Unfree_1</td>
<td>0.188</td>
<td>-0.0402</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.397)</td>
<td>(0.300)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LeftXL.Unfree2</td>
<td>-1.695**</td>
<td>-2.826</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.790)</td>
<td>(1.769)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP_per_Cap (log)</td>
<td>-1.338*</td>
<td>-1.207*</td>
<td>-1.322*</td>
<td>-1.107</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.712)</td>
<td>(0.701)</td>
<td>(0.714)</td>
<td>(0.698)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infant_Mortality</td>
<td>-0.00824</td>
<td>-0.00805</td>
<td>-0.00799</td>
<td>-0.00823</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.0135)</td>
<td>(0.0133)</td>
<td>(0.0135)</td>
<td>(0.0123)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural_Disaster</td>
<td>0.0672</td>
<td>0.0779</td>
<td>0.0666</td>
<td>0.302**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.149)</td>
<td>(0.150)</td>
<td>(0.150)</td>
<td>(0.168)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population (log)</td>
<td>0.918**</td>
<td>0.848**</td>
<td>0.912**</td>
<td>0.909**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.366)</td>
<td>(0.360)</td>
<td>(0.367)</td>
<td>(0.385)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US_Trade (log)</td>
<td>0.319</td>
<td>0.306</td>
<td>0.316</td>
<td>0.299</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.234)</td>
<td>(0.223)</td>
<td>(0.234)</td>
<td>(0.235)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNSC</td>
<td>-0.419</td>
<td>-0.409</td>
<td>-0.418</td>
<td>-0.385</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.231)</td>
<td>(0.223)</td>
<td>(0.232)</td>
<td>(0.237)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WB_Aid_per_Cap (log)</td>
<td>-0.00489</td>
<td>-0.00464</td>
<td>-0.00504</td>
<td>-0.00431</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.0123)</td>
<td>(0.0124)</td>
<td>(0.0124)</td>
<td>(0.0124)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IADB_Aid_per_Cap (log)</td>
<td>0.0492**</td>
<td>0.0494**</td>
<td>0.0494**</td>
<td>0.0494**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.0207)</td>
<td>(0.0206)</td>
<td>(0.0207)</td>
<td>(0.0206)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L.Bilateral_Aid_per_Cap (log)</td>
<td>0.848**</td>
<td>0.842**</td>
<td>0.842**</td>
<td>0.842**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.0700)</td>
<td>(0.0688)</td>
<td>(0.0714)</td>
<td>(0.0682)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Year Fixed Effects: Yes
Country Fixed Effects: Yes
Observations: 727
Number of Countries: 24
Adjusted R-squared: 0.282
Pseudo-R2: 0.282

Table 1: HA - Bilateral Aid per Capita (log)
### Table 2: HB – Bilateral Aid per Capita (log)

<table>
<thead>
<tr>
<th>Model</th>
<th>(1) Linear</th>
<th>(2) Linear</th>
<th>(3) Linear</th>
<th>(4) Tobit</th>
<th>(5) Tobit</th>
<th>(6) Tobit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>-0.0298</td>
<td>-0.0135</td>
<td>-0.0083</td>
<td>0.229</td>
<td>0.0212</td>
<td>-0.00681</td>
</tr>
<tr>
<td></td>
<td>(0.296)</td>
<td>(0.227)</td>
<td>(0.226)</td>
<td>(0.154)</td>
<td>(0.120)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>R_Exec</td>
<td>0.118</td>
<td>0.134</td>
<td>0.153</td>
<td>0.278**</td>
<td>0.265**</td>
<td>0.126</td>
</tr>
<tr>
<td></td>
<td>(0.205)</td>
<td>(0.223)</td>
<td>(0.185)</td>
<td>(0.0874)</td>
<td>(0.130)</td>
<td>(0.0903)</td>
</tr>
<tr>
<td>R_Congress</td>
<td>-4.285*</td>
<td>-4.413*</td>
<td>-4.687*</td>
<td>1.924</td>
<td>2.189</td>
<td>1.049</td>
</tr>
<tr>
<td></td>
<td>(2.371)</td>
<td>(2.382)</td>
<td>(2.400)</td>
<td>(1.324)</td>
<td>(1.430)</td>
<td>(1.453)</td>
</tr>
<tr>
<td>LeftXR_Exec</td>
<td>0.000712</td>
<td>-0.345*</td>
<td></td>
<td>0.000712</td>
<td>-0.345*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.316)</td>
<td>(0.188)</td>
<td></td>
<td>(0.316)</td>
<td>(0.188)</td>
<td></td>
</tr>
<tr>
<td>LeftXR_ExecXR_Congress</td>
<td>3.891</td>
<td>1.440</td>
<td>(4.921)</td>
<td>(2.734)</td>
<td>0.012</td>
<td>(0.311)</td>
</tr>
<tr>
<td></td>
<td>(3.891)</td>
<td>(1.440)</td>
<td>(4.921)</td>
<td>(2.734)</td>
<td>0.012</td>
<td>(0.311)</td>
</tr>
<tr>
<td>L.Unfree_1</td>
<td>-0.358</td>
<td>-0.397</td>
<td>-0.188</td>
<td>-0.0285</td>
<td>-0.0271</td>
<td>0.0026</td>
</tr>
<tr>
<td></td>
<td>(0.240)</td>
<td>(0.163)</td>
<td>(0.167)</td>
<td>(0.167)</td>
<td>(0.167)</td>
<td>(0.167)</td>
</tr>
<tr>
<td>L.Unfree_1XR_Exec</td>
<td>0.0482</td>
<td>0.271</td>
<td>(0.311)</td>
<td>(0.194)</td>
<td>0.012</td>
<td>0.012</td>
</tr>
<tr>
<td>L.Unfree_1XR_ExecXR_Congress</td>
<td>3.891</td>
<td>1.440</td>
<td>(4.921)</td>
<td>(2.734)</td>
<td>0.012</td>
<td>(0.311)</td>
</tr>
<tr>
<td>L.Unfree_2</td>
<td>-0.714</td>
<td>-1.121**</td>
<td>(0.576)</td>
<td>(0.494)</td>
<td>0.012</td>
<td>0.012</td>
</tr>
<tr>
<td>L.Unfree_22XR_Exec</td>
<td>0.401</td>
<td>0.825</td>
<td>(0.663)</td>
<td>(0.726)</td>
<td>0.012</td>
<td>(0.311)</td>
</tr>
<tr>
<td>L.Unfree_2XR_ExecXR_Congress</td>
<td>5.440</td>
<td>0.148</td>
<td>(4.654)</td>
<td>(2.307)</td>
<td>0.012</td>
<td>(0.311)</td>
</tr>
</tbody>
</table>

**Year Fixed Effects:** Yes
**Country Fixed Effects:** Yes
**Observations:** 727
**Number of Countries:** 24
**Adjusted R-squared:** 0.042
**Pseudo-R:** 0.283

---

### Table 2: HB – Bilateral Aid per Capita (log)

<table>
<thead>
<tr>
<th>Model</th>
<th>(1) Linear</th>
<th>(2) Linear</th>
<th>(3) Linear</th>
<th>(4) Tobit</th>
<th>(5) Tobit</th>
<th>(6) Tobit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>-0.0298</td>
<td>-0.0139</td>
<td>-0.0083</td>
<td>0.229</td>
<td>0.0212</td>
<td>-0.00681</td>
</tr>
<tr>
<td></td>
<td>(0.296)</td>
<td>(0.227)</td>
<td>(0.226)</td>
<td>(0.154)</td>
<td>(0.120)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>R_Exec</td>
<td>0.118</td>
<td>0.134</td>
<td>0.153</td>
<td>0.278**</td>
<td>0.265**</td>
<td>0.126</td>
</tr>
<tr>
<td></td>
<td>(0.205)</td>
<td>(0.223)</td>
<td>(0.185)</td>
<td>(0.0874)</td>
<td>(0.130)</td>
<td>(0.0903)</td>
</tr>
<tr>
<td>R_Congress</td>
<td>-4.285*</td>
<td>-4.413*</td>
<td>-4.687*</td>
<td>1.924</td>
<td>2.189</td>
<td>1.049</td>
</tr>
<tr>
<td></td>
<td>(2.371)</td>
<td>(2.382)</td>
<td>(2.400)</td>
<td>(1.324)</td>
<td>(1.430)</td>
<td>(1.453)</td>
</tr>
<tr>
<td>LeftXR_Exec</td>
<td>0.000712</td>
<td>-0.345*</td>
<td></td>
<td>0.000712</td>
<td>-0.345*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.316)</td>
<td>(0.188)</td>
<td></td>
<td>(0.316)</td>
<td>(0.188)</td>
<td></td>
</tr>
<tr>
<td>LeftXR_ExecXR_Congress</td>
<td>3.891</td>
<td>1.440</td>
<td>(4.921)</td>
<td>(2.734)</td>
<td>0.012</td>
<td>(0.311)</td>
</tr>
<tr>
<td></td>
<td>(3.891)</td>
<td>(1.440)</td>
<td>(4.921)</td>
<td>(2.734)</td>
<td>0.012</td>
<td>(0.311)</td>
</tr>
<tr>
<td>L.Unfree_1</td>
<td>-0.358</td>
<td>-0.397</td>
<td>-0.188</td>
<td>-0.0285</td>
<td>-0.0271</td>
<td>0.0026</td>
</tr>
<tr>
<td></td>
<td>(0.240)</td>
<td>(0.163)</td>
<td>(0.167)</td>
<td>(0.167)</td>
<td>(0.167)</td>
<td>(0.167)</td>
</tr>
<tr>
<td>L.Unfree_1XR_Exec</td>
<td>0.0482</td>
<td>0.271</td>
<td>(0.311)</td>
<td>(0.194)</td>
<td>0.012</td>
<td>0.012</td>
</tr>
<tr>
<td>L.Unfree_1XR_ExecXR_Congress</td>
<td>3.891</td>
<td>1.440</td>
<td>(4.921)</td>
<td>(2.734)</td>
<td>0.012</td>
<td>(0.311)</td>
</tr>
<tr>
<td>L.Unfree_2</td>
<td>-0.714</td>
<td>-1.121**</td>
<td>(0.576)</td>
<td>(0.494)</td>
<td>0.012</td>
<td>0.012</td>
</tr>
<tr>
<td>L.Unfree_22XR_Exec</td>
<td>0.401</td>
<td>0.825</td>
<td>(0.663)</td>
<td>(0.726)</td>
<td>0.012</td>
<td>(0.311)</td>
</tr>
<tr>
<td>L.Unfree_2XR_ExecXR_Congress</td>
<td>5.440</td>
<td>0.148</td>
<td>(4.654)</td>
<td>(2.307)</td>
<td>0.012</td>
<td>(0.311)</td>
</tr>
</tbody>
</table>

**Year Fixed Effects:** Yes
**Country Fixed Effects:** Yes
**Observations:** 727
**Number of Countries:** 24
**Adjusted R-squared:** 0.042
**Pseudo-R:** 0.283

---

Standard errors in parentheses (clustered for Tobit models)

*** p<0.01, ** p<0.05, * p<0.1
5 Bibliography


ABSTRACT

My distinguished majors thesis utilizes hypothetical gamble questions within the 1979 National Longitudinal Survey of Youth over the period 1993-2010 to investigate risk tolerance. Using maximum likelihood estimation, I estimate the sample mean and standard deviation of risk tolerance within my sample and examine whether or not risk tolerance remains constant over the period in the face of health and income shocks. Previous literature using the Health and Retirement Study (HRS) finds little or no change under health and income shocks, but a decrease in risk tolerance with age. The high retention rates and seventeen-year period of interest within the NLSY79 allow for an in-depth analysis of risk tolerance stability. Studying the NLSY79 cohort provides an opportunity to examine still developing and possibly more malleable risk preferences. My paper focuses specifically on the determinants of risk tolerance, variation across demographic groups and if and how risk tolerance changes over time when exposed to adverse shocks such as job displacement.
UNEMPLOYMENT BENEFIT INCREASES AND MORTALITY

ERIC S. ALDENHOFF
University of Maryland

ABSTRACT

I examine the impact of mass layoffs on mortality and test whether unemployment insurance benefits provide protection against the negative impact of job loss. I use country-level data for all 50 US states and the District of Columbia from the National Center for Health Statistics and the Department of Labor. Using a difference-in-difference approach with state and year fixed effects, I estimate that a one percent increase in mass layoffs leads to a 0.35 percent increase in mortality. Males are more vulnerable to mass layoffs than females, black and Hispanic workers are more vulnerable to mass layoffs than whites, and younger and older workers respond less to a mass layoff than the rest of the working populations. I also estimate that an increase in extended unemployment benefits has a protective effect on the increase in mortality due to mass layoffs. A thousand dollars per unemployed person increase in extended benefits reduces the mortality due to mass layoffs by 0.007 percent. In the context of sequestration, budget cuts, and the increasing scrutiny of federal expenditures, these results have policy implications regarding optimal allocation of funds as well as the optimal levels of benefit generosity. Unfortunately, I cannot rule out the possibility that my results are driven by time trends or omitted variables that cannot be controlled for due to the limitations of the data.
ABSTRACT

Political risk is an important factor in multinational corporations’ decisions to invest abroad. While foreign direct investment typically promises high returns, such investments also carry the risk that the host government will renege on contracts in order to expropriate the profits and assets of the foreign investor. In countries with federal political systems, subnational governments often play a major role in negotiating—and potentially breaching—long-term contracts with foreign investors. Therefore, the decision to invest will involve considerations of subnational political conditions. This study examines the effects of subnational politics on foreign direct investment across 27 Indian states and union territories from 1996-2009. Results suggest that lower levels of political competition—represented by the dominance of one party over states’ Legislative Assemblies—are negatively correlated with FDI. An increase of 1% of seats held by the ruling party in the Legislative Assembly is correlated with a decrease in annual FDI inflows of approximately 4 million US Dollars, an amount worth approximately 5% of the average annual value of state-level FDI inflows. Further, movement from a non-majority to a majority government is correlated with a decrease in annual FDI inflows of approximately 80 million USD, an amount almost equivalent to average annual state-level FDI inflows. These results suggest that additional veto players in states’ Legislative Assemblies constrain ruling parties’ abilities to unilaterally breach investment contracts, enabling such states to offer a credible commitment to foreign investors and thus to attract FDI. This study presents evidence that the institutionalized gridlock of fragmented governments actually attracts FDI. Party regionalization and the number of changes of ruling party are not found to be significantly correlated with FDI.
1 Introduction

From higher-paying jobs to increased labor productivity, technological spillovers, and reduced income inequality, the potential benefits of foreign direct investment (FDI) have forged a global landscape in which the attraction of FDI has become a priority for developing and middle-income countries. Although these benefits remain subject to debate, national governments of all shapes and sizes have striven to generate conditions perceived by foreign investors as delivering high returns on investment. Many governments have been highly successful in their attempts to attract foreign investment. In 2012—for the first time ever—developing economies absorbed more FDI than developed countries, accounting for 52 percent of $1.45 trillion worth of global FDI. In response to this increasingly intense global competition for FDI, scholars have sought to determine the conditions that attract—and repel—foreign investors. Most studies have emphasized the effects of economic factors such as exchange rates and taxation levels. Multinational corporations’ primary consideration is generating an acceptable return on investment, though overtly economic factors alone do not shape the decision to invest. Political factors, ranging from federalism to corruption, can affect returns and therefore play a role in MNCs’ decisions to invest abroad. While this study acknowledges the critical role played by economic factors, it will examine the less explored side of the FDI equation—the political side.

Foreign direct investment typically flows unevenly to different regions within recipient countries (Ortega Gomez, Cruz, & Alcaraz Vargas, 2010). Existing scholarship, however, has tended to focus on national levels of FDI and to ignore the distribution of FDI inflows within countries. While factors such as macroeconomic policy or political systems are major determinants of FDI at the national level, subnational conditions play a significant role in determining where FDI locates within a country. Insofar as subnational governments are responsible for policies or conditions that affect the viability of foreign investments, investors will theoretically be sensitive to

---

local political conditions. This study will explore the interaction between subnational politics and foreign direct investment. To what extent do subnational political conditions affect the geographic distribution of FDI within countries? For example, are potential investors deterred or encouraged by local political competition and regional political parties? Such questions are particularly relevant in the case of federal democracies, in which subnational political entities often play critical roles in negotiating and maintaining long-term contracts with foreign investors. By examining the effects of subnational politics on FDI, these countries will be able to understand—and influence—the location choices of foreign investors within their borders.

A federal democracy and a major destination for foreign direct investment, India is an ideal case study of the interaction between subnational politics and FDI. Foreign direct investment flows to India increased rapidly following the country’s implementation of liberal economic reforms in the early 1990s. The stock of FDI in India is now quite large—some $220 billion, or 12% of GDP, according to the Reserve Bank of India (RBI). In most sectors, Indian states compete to attract these employment-creating and growth-stimulating foreign firms. Some states have been much more successful than others in their attempts to attract foreign investors—FDI inflows to India have been defined by extreme regional concentration in a small group of states. FDI has accrued disproportionately to the states of Maharashtra, Tamil Nadu, Gujarat, Andhra Pradesh, Karnataka, Uttar Pradesh, and Haryana. The top three states in terms of average FDI inflow received approximately 57% of the total average inflow, while the top seven states received approximately 90% of all inflows. This study will move beyond the economic factors attracting FDI to certain states, exploring the effects of subnational political conditions on state-level FDI inflows. The sharp disparity in state-level FDI inflows threatens to exacerbate regional economic inequalities in India, as relatively high-income states dominate the FDI landscape while the least developed states receive very little foreign investment. An analysis of the political factors underlying the uneven geographical distribution of FDI flows to India therefore carries important regional implications.

A federal democracy and a major destination for foreign direct investment, India is an ideal case study of the interaction between subnational politics and FDI. Foreign direct investment flows to India increased rapidly following the country’s implementation of liberal economic reforms in the early 1990s. The stock of FDI in India is now quite large—some $220 billion, or 12% of GDP, according to the Reserve Bank of India (RBI). In most sectors, Indian states compete to attract these employment-creating and growth-stimulating foreign firms. Some states have been much more successful than others in their attempts to attract foreign investors—FDI inflows to India have been defined by extreme regional concentration in a small group of states. FDI has accrued disproportionately to the states of Maharashtra, Tamil Nadu, Gujarat, Andhra Pradesh, Karnataka, Uttar Pradesh, and Haryana. The top three states in terms of average FDI inflow received approximately 57% of the total average inflow, while the top seven states received approximately 90% of all inflows. This study will move beyond the economic factors attracting FDI to certain states, exploring the effects of subnational political conditions on state-level FDI inflows. The sharp disparity in state-level FDI inflows threatens to exacerbate regional economic inequalities in India, as relatively high-income states dominate the FDI landscape while the least developed states receive very little foreign investment. An analysis of the political factors underlying the uneven geographical distribution of FDI flows to India therefore carries important regional implications.

5 “Foreign Companies in India: The Rewards of Royalty”, The Economist, Jan 26 2013.
7 According to Centre for Monitoring the Indian Economy (CMIE) data on greenfield majority foreign-owned investment projects in India from 1996-2009.
implications for Indian economic development, and for regional patterns of development in other federal democracies.

The origins of post-colonial India’s federalist structure lie in the States Reorganization Act of 1956, which defined the boundaries of states and territories along linguistic lines. As of April 2014, India is comprised of 28 states and 7 union territories. All states and two union territories (Delhi and Puducherry) are governed by directly elected parliamentary-style legislatures called “Legislative Assemblies”. According to the Constitution of India, both Legislative Assemblies and the national legislature can legislate on the 47 items on the “Concurrent List”, which include contracts, commercial and industrial monopolies, and any industry producing goods over which state control would be “expedient in the public interest”.

Subnational political entities have played a major role in the Indian FDI landscape in recent years. For example, the Government of Tamil Nadu is often credited with the creation of one of the most favored destinations for foreign investment in India. However, some state governments have acted opportunistically, altering contracts with foreign investors in order to pander to populist interests. A particularly well-known example is the Government of Maharashtra, which gained populist support by cancelling Enron’s Dabhol power project and later agreeing to an advantageous renegotiation of the terms of the contract. The breach of contract resulted in hundreds of million dollars’ worth of delays for Enron. While relatively rare, government decisions to breach contracts are often the most important risks for firms operating abroad and can therefore act as major deterrents of FDI. How can governments demonstrate a credible commitment to uphold contracts with foreign investors? This study will use the logic of credible commitments via veto players—political actors or institutions that can block policy change—to analyze Indian states’ abilities to attract FDI. Subnational politics, namely state-level political competition and party regionalization, are expected to influence state-level FDI inflows through credible commitments to investment contracts via veto players.

9 “Enron Development Corp.: The Dabhol Power Project in Maharashtra, India”, Harvard Business Publishing.
10 “Enron Development Corp.: The Dabhol Power Project in Maharashtra, India”, Harvard Business Publishing.
2 Literature Review

In the context of foreign direct investment, political risk refers to the probability that political decisions or events will harm business transactions. Changes in government policies, political unrest, or economic sanctions are all political factors that could negatively impact multinational corporations’ investments. This study, however, will focus on one type of political risk: the risk that the host government reneges on a contract with a foreign investor. Given the large sunk costs and relative immobility associated with most FDI, multinational investors are extremely vulnerable to these risks (Vernon, 1971; 1980). According to Vernon’s “obsolescing bargain” model, the host government gains bargaining power relative to the MNC after the decision to invest has been made. Host governments, aware of the costs of relocation or liquidation for MNCs, are tempted to renegotiate contracts or expropriate assets after investors’ costs are sunk (Williamson, 1996). Today, most political risks take the form of “creeping expropriation”—policy changes that negatively affect the operations of an MNC rather than outright nationalization. Such changes include defaulting on tax deals and subsidies, restricting capital flows, and attempting to capture MNCs’ income streams (Kobrin, 1982; Henisz, 2000). As in the case of the Government of Maharashtra, host governments can use the threat of cancellation of a project with already-sunk costs to extract an advantageous renegotiation from the foreign investor. Host governments face significant time-inconsistency problems, potentially valuing the short-term political benefits of breaching contracts over higher long-term levels of FDI if they did not renege (Jensen & MacGillivray, 2005). MNCs are therefore sensitive to the probability that a host government will renege on a contract, and in the face of a high probability of reneging may divert their investment to a complementary market with less political risk (Jensen & MacGillivray, 2005). Host governments perceived as being likely to renege on contracts are therefore expected to attract lower levels of FDI than those offering a credible commitment to investment contracts.

Much of the political risk literature has emphasized the role of veto players in lowering political risk and thus in attracting FDI. Veto players are political actors and/or institutions that can block legislation (Jensen & MacGillivray, 2005). The number of veto players is a key determinant of policy stability, as veto players constrain policymaking and limit the ability of governments to enact sudden policy changes (Tsebelis, 2000). Henisz (2000; 2002) found

---

that institutions with multiple veto players constrain policy change and hence attract more FDI by ensuring policy predictability. Li (2005) found that political constraints through veto players decrease the likelihood of expropriation of foreign assets in both democracies and autocracies. Jensen and McGillivray (2005) found that federalism, through the presence of veto players in multiple jurisdictions, increases governments’ credibility to commit to investment contracts and thus increases FDI inflows. Using India as the example of a highly regionalized political party system, Simmons et al (2011) found that countries with regionalized party systems attract less FDI than nationalized systems, as the presence of additional veto players allows national-level parties to credibly commit to pro-FDI policies.

However, these studies have not considered the effects of political competition within a democracy—that is, the differing degree to which subnational democratic governments are constrained by institutions. The extent of ruling parties’ control over Indian states’ Legislative Assemblies varies widely, suggesting that state governments face differing numbers of veto players and thus differing degrees of policy constraints. Further, the varying geographical presence of regional and national parties suggests that some types of parties may face more veto players and thus policy constraints than others.

This study will utilize the logic of credible commitment to contracts via veto players to examine the effect of two major trends in Indian subnational politics on state-level FDI inflows. The first trend is increased state-level political competition, as evidenced by the success of non-INC political parties in many states and the preponderance of coalition governments in state Legislative Assemblies since the early 1990s. Because they are responsible for all state-level lawmaking—and have the capability of breaching contracts with foreign investors—this study will emphasize Legislative Assemblies as the critical political institutions in the subnational Indian FDI landscape. The degree of competition within Indian states’ Legislative Assemblies varies widely—in some states, a ruling party holds over 95% of the total seats, while the ruling party may hold just over 20% of seats in other states. A majority government holding over 50% of seats can pass legislation without the support of other parties, while the leading party of a fragmented coalition government must gain the support of numerous other parties. The number of actors capable of blocking legislation—veto players—therefore varies widely in different states’ Legislative Assemblies. According to the logic of credible commitments via veto players, state-level governments facing more veto players in the Assembly will offer a more credible commitment to contracts with foreign investors. Ruling parties that institutions with multiple veto players constrain policy change and hence attract more FDI by ensuring policy predictability. Li (2005) found that political constraints through veto players decrease the likelihood of expropriation of foreign assets in both democracies and autocracies. Jensen and McGillivray (2005) found that federalism, through the presence of veto players in multiple jurisdictions, increases governments’ credibility to commit to investment contracts and thus increases FDI inflows. Using India as the example of a highly regionalized political party system, Simmons et al (2011) found that countries with regionalized party systems attract less FDI than nationalized systems, as the presence of additional veto players allows national-level parties to credibly commit to pro-FDI policies.

However, these studies have not considered the effects of political competition within a democracy—that is, the differing degree to which subnational democratic governments are constrained by institutions. The extent of ruling parties’ control over Indian states’ Legislative Assemblies varies widely, suggesting that state governments face differing numbers of veto players and thus differing degrees of policy constraints. Further, the varying geographical presence of regional and national parties suggests that some types of parties may face more veto players and thus policy constraints than others.

This study will utilize the logic of credible commitment to contracts via veto players to examine the effect of two major trends in Indian subnational politics on state-level FDI inflows. The first trend is increased state-level political competition, as evidenced by the success of non-INC political parties in many states and the preponderance of coalition governments in state Legislative Assemblies since the early 1990s. Because they are responsible for all state-level lawmaking—and have the capability of breaching contracts with foreign investors—this study will emphasize Legislative Assemblies as the critical political institutions in the subnational Indian FDI landscape. The degree of competition within Indian states’ Legislative Assemblies varies widely—in some states, a ruling party holds over 95% of the total seats, while the ruling party may hold just over 20% of seats in other states. A majority government holding over 50% of seats can pass legislation without the support of other parties, while the leading party of a fragmented coalition government must gain the support of numerous other parties. The number of actors capable of blocking legislation—veto players—therefore varies widely in different states’ Legislative Assemblies. According to the logic of credible commitments via veto players, state-level governments facing more veto players in the Assembly will offer a more credible commitment to contracts with foreign investors. Ruling parties
constrained by more competitors in the Assembly are less able to implement sudden changes to investment contracts. State-level political competition is therefore expected to be positively correlated with FDI inflows.

The second trend examined in this study is the regionalization of Indian political parties. Although India is the “canonical example of a regionalized party system”, with a proliferation of state-based political parties in power who frequently do not run candidates in states outside their own, no studies have examined the impact of regional political parties on state-level FDI inflows. This study will seek to fill this gap in the literature, using the logic of credible policy commitments via veto players to examine the effect of regional political parties on the amount of state-level FDI inflows. Regional parties’ relatively narrow power base (in just one to three states, according to the Election Commission of India) connotes the presence of fewer veto players than in the case of national parties, whose “veto base” consists of party affiliates in many states and the central government. While external party affiliates cannot formally “block” legislation in states’ Legislative Assemblies, it is likely that party affiliates in other states and the central governments play a major role in encouraging or discouraging certain policies in individual states. Regional parties’ policy decisions are in theory less constrained by veto players, making them less able to credibly commit to contracts as compared to national parties, whose nation-wide set of veto players renders them less likely to implement sudden changes to investment contracts. It is therefore expected that state rule by regional parties is negatively associated with FDI inflows.

This study will utilize the findings of the existing literature on subnational determinants of FDI to control for non-political determinants of FDI. In the context of the United States, Coughlin, Terza, and Aromdee (1991) found that state per capita income, manufacturing density, transportation infrastructure, and FDI-attracting expenditures are positively linked to state-level FDI inflows. In the Chinese context, Luo (2005) found that natural resources and low labor costs did not attract FDI to China’s hinterland provinces, while policy incentives and industrial agglomeration did. Cheng and Kwan (2000) found that the presence of a large regional market, good infrastructure, existing FDI, and preferential policy had a positive effect on provincial FDI inflows and labor cost had a negative effect.

In the Indian context, Siddarthan (2006) found that states with high per capita income, high industrial output, and coastal locations attracted high

levels of FDI. Mukherjee (2011) found market size, agglomeration, and size of the services/manufacturing base to be positively correlated with state-level FDI inflows, and tax rates and the cost of labor to be negatively correlated. Morris (2004) found that the presence of a metropolitan city in a state has a significant positive effect on FDI inflows. Meanwhile, Nunnenkamp and Stracke (2007) found a positive correlation of FDI with per capita income, population density, per capita bank deposits, telephone density, level of education and per capita net value added in manufacturing. Aggarwal (2005) found that rigid labor markets in Indian states discourage FDI. Bajpal and Sachs (1999) examined the effect of economic reforms, finding that the most reform-oriented states attracted the most FDI. Chatterjee et al (2013) found the level and variability of profitability of existing firms to have a significant effect on FDI inflows. One of the few studies examining the role of political factors in attracting FDI at the state level was conducted by Ramachandran and Goebel (2002), who found that the state of Tamil Nadu had emerged as one of the most favored investment destinations in India due to factors that included strong and stable state-level government with proactive government policies and investor-friendly and transparent decision-making processes.

3 Data and Methodology
3.1 Data and Relevant Variables

This study utilizes a panel dataset of 25 Indian states and 2 union territories across the years 1996 to 2009. The dataset includes all major Indian states: namely Maharashtra, Tamil Nadu, Gujarat, Andhra Pradesh, Karnataka, Uttar Pradesh, Haryana, West Bengal, Rajasthan, Uttarakhand, Punjab, Madhya Pradesh, Himachal Pradesh, Goa, Chhattisgarh, Kerala, Jharkhand, Orissa, Jammu and Kashmir, Tripura, Bihar, Assam, Arunachal Pradesh, Meghalaya, and Sikkim. The union territories included in the dataset are Delhi and Puducherry, which are the only union territories possessing state-level legislative bodies. The union territories of Daman and Diu and Dadra and Nagar Haveli, included in the original dataset, were excluded from this study due to their lack of state-level governments. This extensive list of states and union territories was determined by the availability of data on the dependent variable, which is the value of state-level annual greenfield FDI inflow given in units of 10 million Rupees (to be converted into US Dollars in the Results section). Thus the 55.44 worth of annual greenfield FDI inflow to the state of Andhra Pradesh in the year 1996 would represent a value of 554,400,000 Rupees (or approximately USD 9.1 million). The data on state-level FDI inflows was sourced from the Centre for Monitoring the Indian Economy (CMIE), an organization widely perceived as the most accurate and comprehensive source of data on the Indian economy. Specifically, the

levels of FDI. Mukherjee (2011) found market size, agglomeration, and size of the services/manufacturing base to be positively correlated with state-level FDI inflows, and tax rates and the cost of labor to be negatively correlated. Morris (2004) found that the presence of a metropolitan city in a state has a significant positive effect on FDI inflows. Meanwhile, Nunnenkamp and Stracke (2007) found a positive correlation of FDI with per capita income, population density, per capita bank deposits, telephone density, level of education and per capita net value added in manufacturing. Aggarwal (2005) found that rigid labor markets in Indian states discourage FDI. Bajpal and Sachs (1999) examined the effect of economic reforms, finding that the most reform-oriented states attracted the most FDI. Chatterjee et al (2013) found the level and variability of profitability of existing firms to have a significant effect on FDI inflows. One of the few studies examining the role of political factors in attracting FDI at the state level was conducted by Ramachandran and Goebel (2002), who found that the state of Tamil Nadu had emerged as one of the most favored investment destinations in India due to factors that included strong and stable state-level government with proactive government policies and investor-friendly and transparent decision-making processes.
data is from CMIE’s CapEx database, which provides data on volumes of majority foreign-owned greenfield investment (investment into new ventures rather than existing facilities or projects) into India. The CMIE data provides 313 observations of the dependent variable, annual state-level greenfield FDI inflows, with an average value worth approximately USD 82 million.

This study seeks to determine the effects of two broad aspects of subnational politics on state-level FDI inflows, using five primary independent variables. All data on state-level politics used to construct the primary independent variables is sourced from the Election Commission of India’s statistical reports on elections to state Legislative Assemblies. The first aspect of subnational politics is the degree of political competition within states, which will be described by three variables. *Assembly seats* is a continuous variable, measuring the percentage of seats held by the ruling party in the state Legislative Assembly. The value of *assembly seats* ranges from a minimum of 22.99% of seats held by a coalition government in Jammu & Kashmir to a maximum of 96.88% of seats held by a dominant majority in Sikkim.

Lower values of *assembly seats* describe fragmented state legislative assemblies, with numerous viable parties competing for power and a ruling party that must cooperate with other parties to pass legislation. Specifically, a value of *assembly seats* of 50% or less indicates a coalition or minority government, which requires the support of other parties (whether within a formal coalition arrangement or not) to secure the majority of votes required to pass legislation. In contrast, higher values of *assembly seats* describe state assemblies with one dominant party and relatively less political competition. The higher the value of *assembly seats*, the more power the ruling party holds in the assembly. Specifically, an *assembly seats* value greater than 50% indicates an absolute majority government, which does not require the support of any other parties to pass legislation. *Assembly seats* provides an estimate of the effect of a 1 percent increase in seats held in the state legislative assembly by the ruling party on the volume of annual FDI inflows. Because increased political competition is expected to be positively associated with FDI inflows via increased veto players, it is expected that *assembly seats* will be negatively associate with FDI inflows.

The second political competition variable, *majority*, is similar to *assembly seats*, but specifically attempts to capture the effect of majority governments versus minority or coalition governments on FDI inflows. *Majority* is a dummy variable, which equals 1 if the ruling party holds an absolute majority (with more than 50% of the total seats) and 0 if the ruling party does not hold an absolute majority. Some governments (such as those
in Sikkim and Tripura) have enjoyed absolute majorities over the entire 14-year period, while some state legislative assemblies did not experience an absolute majority government at all over the period (such as those of Maharashtra and Kerala). A notable difference exists between these types of governments due to the ability of absolute majority governments to pass legislation without the support of other parties. In contrast, ruling parties holding 50% of seats or less must gain the support of their coalition allies or parties outside the coalition in order to pass legislation. Governments possessing an absolute majority in the legislative assembly thus face few veto players relative to minority or coalition governments, rendering it easier for them to alter investment contracts. Majority provides an estimate of the effect of absolute majority governments on FDI inflows, as compared to minority or coalition governments. Because increased values of both assembly seats and majority describe lower levels of subnational political competition, both variables are expected to be negatively associated with FDI inflows.

The second aspect of subnational politics to be explored in this study is the regionalization of Indian political parties, which will be measured by two dummy variables. The two variables, regional1 and regional2, both equal 1 if the state’s ruling party is a regional party and 0 if the ruling party is a national party. Thus both variables estimate the effect of state rule by regional political parties on annual state-level FDI inflows, as compared to national governments. The sole difference between the two regionalization variables is in their definitions of “regional” parties. The first variable, regional1, defines “regional” parties according to the Election Commission of India’s definition of regional parties as any party registered in fewer than four states. According to regional1, therefore, national parties are parties registered in at least four states. Under this definition, states such as West Bengal, which was ruled by the Communist Party of India (Marxist) from 1996-2009, were never ruled by regional parties (as the CPM is registered in more than three states). However, the categorization of the CPM in West Bengal changes from “national” to “regional” under the second regionalization variable, regional2. Instead of the Election Commission definitions, regional2 narrows the definition of “national” party to the only two parties possessing the vast geographical presence required to be elected to national office in India, the Indian National Congress (INC) and the Bharatiya Janata Party (BJP). Therefore, all parties other than the INC and BJP are coded as regional parties with a value of 1, while the INC and BJP are coded as 0. The regional2 definition considers whether policy stability via veto players is only taken into account by foreign investors at the level of the two truly “national” parties, the INC and BJP. While parties such as the
3.2 Methodology

This study will estimate a series of random-effects generalized least squares (GLS) regression models, using annual greenfield FDI inflow to each state as the dependent variable in all specifications. Previous analyses of subnational FDI have discussed the likelihood of unmodeled heterogeneity between panels and have therefore used a Hausman specification test to determine whether fixed effects or random effects models should be used (Samford and Gomez 2014). In the context of this study, the results of the Hausman test (p=0.198) fail to reject the null hypothesis and recommend the use of random-effects over fixed-effects models. Following the methodologies of other studies on subnational FDI determinants, all independent variables are lagged by one year (i.e. t-1, or from 1995 to 2008) to clarify the direction of causation and to elucidate the fact that FDI flows in a particular year are generally determined by the economic and political conditions existing in the previous year. This study will utilize numerous additional variables to control for non-political determinants of state-level FDI inflows. The control variables in this study are similar to those used throughout the literature on the subnational determinants of FDI. Though this body of literature is somewhat small, certain factors—such as states’ wealth, size, economic development, infrastructure quality, labor quality, labor cost, tax burden, and geographical location—have emerged repeatedly as potentially affecting subnational FDI inflows. As mentioned, these factors are of an overly economic or geographic nature as opposed to the political determinants emphasized in this study. Given the lack of a centralized source of Indian state-level data, these variables originate from various sources but were all secured via the website IndiaStat, a reputable and reliable source for India-related data.¹⁴

¹⁴ IndiaStat, <www.indiastat.com>

CPM and the Bahujan Samaj Party (BSP) have a presence in numerous states, their primary support bases are still generally limited to one or two states (West Bengal in the case of CPM and Uttar Pradesh in the case of BSP). It may therefore be the case that the logic of credible commitments to contracts and pro-FDI policies via veto players only applies to the two parties with genuine national-level support and international recognition. Due to their limited geographical presence under both definitions, regional parties are assumed to face fewer veto players than national parties. It is therefore expected that regional parties are less able to offer a credible commitment to stable, pro-FDI policies than national parties and therefore that regional parties attract less foreign investment. This study will utilize numerous additional variables to control for non-political determinants of state-level FDI inflows. The control variables in this study are similar to those used throughout the literature on the subnational determinants of FDI. Though this body of literature is somewhat small, certain factors—such as states’ wealth, size, economic development, infrastructure quality, labor quality, labor cost, tax burden, and geographical location—have emerged repeatedly as potentially affecting subnational FDI inflows. As mentioned, these factors are of an overly economic or geographic nature as opposed to the political determinants emphasized in this study. Given the lack of a centralized source of Indian state-level data, these variables originate from various sources but were all secured via the website IndiaStat, a reputable and reliable source for India-related data.¹⁴

³² Methodology

This study will estimate a series of random-effects generalized least squares (GLS) regression models, using annual greenfield FDI inflow to each state as the dependent variable in all specifications. Previous analyses of subnational FDI have discussed the likelihood of unmodeled heterogeneity between panels and have therefore used a Hausman specification test to determine whether fixed effects or random effects models should be used (Samford and Gomez 2014). In the context of this study, the results of the Hausman test (p=0.198) fail to reject the null hypothesis and recommend the use of random-effects over fixed-effects models. Following the methodologies of other studies on subnational FDI determinants, all independent variables are lagged by one year (i.e. t-1, or from 1995 to 2008) to clarify the direction of causation and to elucidate the fact that FDI flows in a particular year are generally determined by the economic and political conditions existing in the previous year. This study will utilize numerous additional variables to control for non-political determinants of state-level FDI inflows. The control variables in this study are similar to those used throughout the literature on the subnational determinants of FDI. Though this body of literature is somewhat small, certain factors—such as states’ wealth, size, economic development, infrastructure quality, labor quality, labor cost, tax burden, and geographical location—have emerged repeatedly as potentially affecting subnational FDI inflows. As mentioned, these factors are of an overly economic or geographic nature as opposed to the political determinants emphasized in this study. Given the lack of a centralized source of Indian state-level data, these variables originate from various sources but were all secured via the website IndiaStat, a reputable and reliable source for India-related data.¹⁴

¹⁴ IndiaStat, <www.indiastat.com>
previous year. Because of the many zero value observations occurring in both the dependent and independent variables, this study will follow studies such as Jensen and McGillivray (2005) in not using log-transformed variables. Log transformation in this case results in many lost observations and does not properly the model the expected additive relationship between the dependent variable (FDI) and the independent variables. In lieu of log-transformed variables, robust standard errors have been used to control for heteroskedasticity in all regressions.

Initial regression models in Table 1 include each of the primary independent variables, along with the full set of controls for states’ economic and geographical/geological characteristics. Thus model 1 examines the potential correlation between assembly seats and annual state-level FDI inflows, model 2 examines the potential correlation between majority and FDI inflows, and so forth. Regression models in Table 2 perform robustness checks on assembly seats across five additional specifications including the other primary independent variables and the full set of controls. Models in Table 3 perform robustness checks on majority across five additional specifications, also including the other primary independent variables and the full set of controls. In the Results section, all amounts given in millions of US Dollars have been converted from Rupees using the USD-INR exchange rate as of late March 2012, 60.18 Rupees to the Dollar.

4 Results

The results of all 15 regression models in this study suggest that the degree of political competition within Indian states’ Legislative Assemblies is strongly correlated with state-level FDI inflows. The dominance of one political party over the Legislative Assembly, as represented by the assembly seats and majority variables, is found to be negatively associated with FDI across all models. In particular, an increase of 1% of seats held by the ruling party in the Legislative Assembly is correlated with a decrease in annual FDI inflows of approximately 4 million US Dollars, a decrease worth approximately 5% of the USD 82 million of average annual inflow across all states. The effect of moving from a non-majority to a majority government in the Legislative Assembly is a decrease in annual FDI inflows of approximately USD 80 million, an amount almost equivalent to the average annual inflow across all states. In contrast, the insignificance of the party instability, regional1, and regional2 variables across all models suggests that changes of ruling parties and party regionalization have little effect on state-level FDI inflows.
whether the critical threshold level of seats—the over 50% required for a majority environments. However, we must look to the commitment to foreign investors who perceive such states as politically risky easily alter investment contracts, and thus do not offer a credible competition. An increase in just one seat held in the Assembly represents the almost identical size as in model 1. Throughout the models, a 1% increase in the percentage of seats held by the ruling party in the Legislative Assembly is associated with a USD 4.03 million (model 7) to a USD 4.09 million (model 8) decrease in annual state-level FDI inflows. As in model 1, these decreases are worth approximately 5% of the average state-level inflow to all states. All models therefore estimate that a 1% increase in seats held by the ruling party in the state Legislative Assembly is correlated with a decrease in annual state-level FDI inflows of USD 4 million, worth approximately 5% of the average annual FDI state-level inflow across all states from 1996-2009.

This relatively large predicted effect suggests that the logic of credible commitments via veto players applies to the case of Indian states’ Legislative Assemblies. States with Legislative Assemblies dominated by one party—those with high values for assemblyseats—are estimated to attract less FDI than states in which the ruling party faces greater levels of political competition. An increase in just one seat held in the Assembly represents the loss of one veto player, an actor that would previously have constrained the policy decisions of the ruling party. The more seats held by the ruling party in the Legislative Assembly, the fewer veto players the government faces in passing legislation. Governments possessing large percentages of seats can easily alter investment contracts, and thus do not offer a credible commitment to foreign investors who perceive such states as politically risky environments. However, we must look to the majority variable to determine whether the critical threshold level of seats—the over 50% required for a
majority government—plays a role in foreign investors’ decision to invest in particular states.

In model 2, majority has a large negative coefficient (-484.3), which is significant at the 5% level. According to this coefficient, movement from a non-majority (meaning a coalition or minority government with 50% or less of the seats) to a majority government is correlated with a USD 80.48 million decrease in annual state-level FDI inflow. This decrease is worth approximately 97% of the average annual FDI inflow across all states. The results presented in model 2 are similar to those in all models containing the majority variable. In models 11-15, the coefficients on the majority variable remain significant at the 5% level, and of similar size as in model 2. Throughout the models, a change from a non-majority to a majority government is correlated with decreases in annual state-level FDI inflows of USD 77.9 million (model 14) to USD 80.6 million (model 13). As in model 2, these amounts are almost equivalent to the average annual FDI inflow across all states, USD 82 million. The scale of this predicted effect suggests that majority governments are strongly negatively correlated with state-level FDI inflows. As in the case of the assemblyseats variable, this result follows from the logic of credible commitments via veto players. Majority governments, possessing the ability to pass legislation without the support of other parties in the Legislative Assembly, face few (if any) veto players in the Assembly and can unilaterally breach investment contracts with relative ease. In contrast, coalition and minority governments must gain the support of other parties before they can pass legislation and therefore offer a more credible commitment to foreign investors.

The results of both the assemblyseats and majority variables suggest that political competition in states’ Legislative Assemblies plays a major role in offering foreign investors a credible commitment to investment contracts and thus in attracting FDI. Higher degrees of political competition provide additional veto players in states’ Legislative Assemblies, constraining ruling parties’ abilities to engage in opportunistic breaches of contracts. In states with low degrees of political competition, the party dominating the Assembly can breach contracts with relative ease. Such states are seen by foreign investors as politically risky, and they will prefer to invest in states whose governments offer a more credible commitment to contracts—those states experiencing high levels of political competition. These results hold when controlling for the number of changes of ruling party and for the geographical presence of the ruling party (regional versus national), as seen in models 6-15. In contrast to the oft-held belief that slow-moving coalition and minority governments repel foreign investors, this study presents majority government—plays a role in foreign investors’ decision to invest in particular states.

In model 2, majority has a large negative coefficient (-484.3), which is significant at the 5% level. According to this coefficient, movement from a non-majority (meaning a coalition or minority government with 50% or less of the seats) to a majority government is correlated with a USD 80.48 million decrease in annual state-level FDI inflow. This decrease is worth approximately 97% of the average annual FDI inflow across all states. The results presented in model 2 are similar to those in all models containing the majority variable. In models 11-15, the coefficients on the majority variable remain significant at the 5% level, and of similar size as in model 2. Throughout the models, a change from a non-majority to a majority government is correlated with decreases in annual state-level FDI inflows of USD 77.9 million (model 14) to USD 80.6 million (model 13). As in model 2, these amounts are almost equivalent to the average annual FDI inflow across all states, USD 82 million. The scale of this predicted effect suggests that majority governments are strongly negatively correlated with state-level FDI inflows. As in the case of the assemblyseats variable, this result follows from the logic of credible commitments via veto players. Majority governments, possessing the ability to pass legislation without the support of other parties in the Legislative Assembly, face few (if any) veto players in the Assembly and can unilaterally breach investment contracts with relative ease. In contrast, coalition and minority governments must gain the support of other parties before they can pass legislation and therefore offer a more credible commitment to foreign investors.

The results of both the assemblyseats and majority variables suggest that political competition in states’ Legislative Assemblies plays a major role in offering foreign investors a credible commitment to investment contracts and thus in attracting FDI. Higher degrees of political competition provide additional veto players in states’ Legislative Assemblies, constraining ruling parties’ abilities to engage in opportunistic breaches of contracts. In states with low degrees of political competition, the party dominating the Assembly can breach contracts with relative ease. Such states are seen by foreign investors as politically risky, and they will prefer to invest in states whose governments offer a more credible commitment to contracts—those states experiencing high levels of political competition. These results hold when controlling for the number of changes of ruling party and for the geographical presence of the ruling party (regional versus national), as seen in models 6-15. In contrast to the oft-held belief that slow-moving coalition and minority governments repel foreign investors, this study presents
evidence that the “institutionalized gridlock” of fragmented governments actually attracts FDI. While investors seek swift government action as they decide to invest, it is commitment to contracts they seek in the long term—and it is this commitment that keeps them coming back.

However, the logic of credible commitments to contracts via veto players does not appear to extend to party affiliates outside states’ Legislative Assemblies. Across all models in which they appear, the coefficients on the party regionalization variables, $regional1$ and $regional2$, are negative (as expected) and sizeable, though insignificant. These results can be seen in models 4 and 5, and repeatedly throughout Tables 2 and 3. Party regionalization does not appear correlated with state-level FDI inflows. This result likely occurs because party affiliates outside the Legislative Assembly do not act as true veto players in the eyes of foreign investors. Though regional parties necessarily face fewer veto players than national parties overall, they do not necessarily face fewer veto players within the relevant legislative body, the Legislative Assembly. Because it is within the Legislative Assembly that breaches of investment contracts would occur—through the passage of laws to increase the taxes on a project, for example—it is the number of veto players in the Assembly that matters.

5 Conclusion

While economic and geographic factors play a major role in determining the destination of foreign direct investment in India, this study presents evidence that subnational politics also influence the location choices of foreign investors. Increased political competition in Indian states’ legislatures is found to be positively correlated with FDI inflows. Governments whose policy decisions are constrained by veto players in Legislative Assemblies can less easily renge on contracts than governments facing little political opposition. The significant political constraints facing minority and coalition governments enable them to provide foreign investors with credible commitments to long-term contracts. The credibility of these commitments lowers the level of political risk perceived by foreign investors, and provides an incentive to invest. In contrast to the oft-held view that slow-moving coalition governments discourage foreign investors, the institutional gridlock in such governments actually attracts FDI by rendering breach of contract less likely. In the realm of foreign direct investment, the fragmentation of India’s governments has the potential to be a blessing rather than a curse.

15 “India’s Upcoming Elections: Will the States and the Economy Decide?” National Bureau of Asian Research.
## Table 1: Initial Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FDI</td>
<td>FDI</td>
<td>FDI</td>
<td>FDI</td>
<td>FDI</td>
</tr>
<tr>
<td>GSDP</td>
<td>0.00836**</td>
<td>0.00873**</td>
<td>0.00739**</td>
<td>0.00690**</td>
<td>0.00711**</td>
</tr>
<tr>
<td></td>
<td>(0.00337)</td>
<td>(0.00345)</td>
<td>(0.00342)</td>
<td>(0.00345)</td>
<td>(0.00346)</td>
</tr>
<tr>
<td>percitasdp</td>
<td>0.0309</td>
<td>0.0375</td>
<td>0.0379</td>
<td>0.0352</td>
<td>0.0341</td>
</tr>
<tr>
<td></td>
<td>(0.0266)</td>
<td>(0.0268)</td>
<td>(0.0272)</td>
<td>(0.0273)</td>
<td>(0.0278)</td>
</tr>
<tr>
<td>FDlstock</td>
<td>0.0898**</td>
<td>0.0773*</td>
<td>0.0691*</td>
<td>0.0683*</td>
<td>0.0678*</td>
</tr>
<tr>
<td></td>
<td>(0.0403)</td>
<td>(0.0401)</td>
<td>(0.0391)</td>
<td>(0.039)</td>
<td>(0.0391)</td>
</tr>
<tr>
<td>population</td>
<td>0.00693</td>
<td>0.00747</td>
<td>0.00521</td>
<td>0.00384</td>
<td>0.00328</td>
</tr>
<tr>
<td></td>
<td>(0.00528)</td>
<td>(0.00537)</td>
<td>(0.00529)</td>
<td>(0.00532)</td>
<td>(0.00557)</td>
</tr>
<tr>
<td>valueofindustrialoutput</td>
<td>0.0473*</td>
<td>0.0428*</td>
<td>0.0473*</td>
<td>0.0438*</td>
<td>0.0438*</td>
</tr>
<tr>
<td></td>
<td>(0.00287)</td>
<td>(0.00284)</td>
<td>(0.00274)</td>
<td>(0.00259)</td>
<td>(0.00291)</td>
</tr>
<tr>
<td>percapitasdp</td>
<td>0.00237</td>
<td>0.00444</td>
<td>0.00734</td>
<td>0.00568</td>
<td>0.00541</td>
</tr>
<tr>
<td></td>
<td>(0.00729)</td>
<td>(0.00728)</td>
<td>(0.0079)</td>
<td>(0.00743)</td>
<td>(0.00747)</td>
</tr>
<tr>
<td>roaddensity</td>
<td>0.515</td>
<td>0.637</td>
<td>0.308</td>
<td>0.289</td>
<td>0.361</td>
</tr>
<tr>
<td></td>
<td>(1.617)</td>
<td>(1.68)</td>
<td>(1.612)</td>
<td>(1.627)</td>
<td>(1.626)</td>
</tr>
<tr>
<td>literacyrate</td>
<td>7.777</td>
<td>9.642</td>
<td>4.912</td>
<td>5.171</td>
<td>5.867</td>
</tr>
<tr>
<td></td>
<td>(14.9)</td>
<td>(15.12)</td>
<td>(15.28)</td>
<td>(15.66)</td>
<td>(16.08)</td>
</tr>
<tr>
<td>reform</td>
<td>1.197**</td>
<td>1.226***</td>
<td>996.8***</td>
<td>997.7***</td>
<td>997.7***</td>
</tr>
<tr>
<td></td>
<td>(339.7)</td>
<td>(356.4)</td>
<td>(335.9)</td>
<td>(339.3)</td>
<td>(342.7)</td>
</tr>
<tr>
<td>coastal</td>
<td>1.386**</td>
<td>1.511***</td>
<td>1.015**</td>
<td>1.065**</td>
<td>1.080**</td>
</tr>
<tr>
<td></td>
<td>(550.7)</td>
<td>(600.4)</td>
<td>(546.9)</td>
<td>(549.8)</td>
<td>(562.4)</td>
</tr>
<tr>
<td>minerals</td>
<td>-577.0**</td>
<td>-591.6**</td>
<td>-750.0***</td>
<td>-740.6***</td>
<td>-726.3***</td>
</tr>
<tr>
<td></td>
<td>(274.7)</td>
<td>(279.9)</td>
<td>(276.3)</td>
<td>(278.4)</td>
<td>(278.4)</td>
</tr>
<tr>
<td>assemblyseats</td>
<td>-24.88***</td>
<td>-484.3**</td>
<td>81.55</td>
<td>81.55</td>
<td>81.55</td>
</tr>
<tr>
<td></td>
<td>(9.571)</td>
<td>(254)</td>
<td>(109.7)</td>
<td>(109.7)</td>
<td>(109.7)</td>
</tr>
<tr>
<td>majority</td>
<td>-209.3</td>
<td>-97.37</td>
<td>-209.3</td>
<td>-97.37</td>
<td>-97.37</td>
</tr>
<tr>
<td></td>
<td>(266.4)</td>
<td>(245.8)</td>
<td>(266.4)</td>
<td>(254.8)</td>
<td>(254.8)</td>
</tr>
</tbody>
</table>

Observations: 240

R² (overall): 0.44

Robust standard errors in parentheses.

Table 1: Initial Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FDI</td>
<td>FDI</td>
<td>FDI</td>
<td>FDI</td>
<td>FDI</td>
</tr>
<tr>
<td>GSDP</td>
<td>0.00836**</td>
<td>0.00873**</td>
<td>0.00739**</td>
<td>0.00690**</td>
<td>0.00711**</td>
</tr>
<tr>
<td></td>
<td>(0.00337)</td>
<td>(0.00345)</td>
<td>(0.00342)</td>
<td>(0.00345)</td>
<td>(0.00346)</td>
</tr>
<tr>
<td>percitasdp</td>
<td>0.0309</td>
<td>0.0375</td>
<td>0.0379</td>
<td>0.0352</td>
<td>0.0341</td>
</tr>
<tr>
<td></td>
<td>(0.0266)</td>
<td>(0.0268)</td>
<td>(0.0272)</td>
<td>(0.0273)</td>
<td>(0.0278)</td>
</tr>
<tr>
<td>FDlstock</td>
<td>0.0898**</td>
<td>0.0773*</td>
<td>0.0691*</td>
<td>0.0683*</td>
<td>0.0678*</td>
</tr>
<tr>
<td></td>
<td>(0.0403)</td>
<td>(0.0401)</td>
<td>(0.0391)</td>
<td>(0.039)</td>
<td>(0.0391)</td>
</tr>
<tr>
<td>population</td>
<td>0.00693</td>
<td>0.00747</td>
<td>0.00521</td>
<td>0.00384</td>
<td>0.00328</td>
</tr>
<tr>
<td></td>
<td>(0.00528)</td>
<td>(0.00537)</td>
<td>(0.00529)</td>
<td>(0.00532)</td>
<td>(0.00557)</td>
</tr>
<tr>
<td>valueofindustrialoutput</td>
<td>0.0473*</td>
<td>0.0428*</td>
<td>0.0473*</td>
<td>0.0438*</td>
<td>0.0438*</td>
</tr>
<tr>
<td></td>
<td>(0.00287)</td>
<td>(0.00284)</td>
<td>(0.00274)</td>
<td>(0.00259)</td>
<td>(0.00291)</td>
</tr>
<tr>
<td>percapitasdp</td>
<td>0.00237</td>
<td>0.00444</td>
<td>0.00734</td>
<td>0.00568</td>
<td>0.00541</td>
</tr>
<tr>
<td></td>
<td>(0.00729)</td>
<td>(0.00728)</td>
<td>(0.0079)</td>
<td>(0.00743)</td>
<td>(0.00747)</td>
</tr>
<tr>
<td>roaddensity</td>
<td>0.515</td>
<td>0.637</td>
<td>0.308</td>
<td>0.289</td>
<td>0.361</td>
</tr>
<tr>
<td></td>
<td>(1.617)</td>
<td>(1.68)</td>
<td>(1.612)</td>
<td>(1.627)</td>
<td>(1.626)</td>
</tr>
<tr>
<td>literacyrate</td>
<td>7.777</td>
<td>9.642</td>
<td>4.912</td>
<td>5.171</td>
<td>5.867</td>
</tr>
<tr>
<td></td>
<td>(14.9)</td>
<td>(15.12)</td>
<td>(15.28)</td>
<td>(15.66)</td>
<td>(16.08)</td>
</tr>
<tr>
<td>reform</td>
<td>1.197**</td>
<td>1.226***</td>
<td>996.8***</td>
<td>997.7***</td>
<td>997.7***</td>
</tr>
<tr>
<td></td>
<td>(339.7)</td>
<td>(356.4)</td>
<td>(335.9)</td>
<td>(339.3)</td>
<td>(342.7)</td>
</tr>
<tr>
<td>coastal</td>
<td>1.386**</td>
<td>1.511***</td>
<td>1.015**</td>
<td>1.065**</td>
<td>1.080**</td>
</tr>
<tr>
<td></td>
<td>(550.7)</td>
<td>(600.4)</td>
<td>(546.9)</td>
<td>(549.8)</td>
<td>(562.4)</td>
</tr>
<tr>
<td>minerals</td>
<td>-577.0**</td>
<td>-591.6**</td>
<td>-750.0***</td>
<td>-740.6***</td>
<td>-726.3***</td>
</tr>
<tr>
<td></td>
<td>(274.7)</td>
<td>(279.9)</td>
<td>(276.3)</td>
<td>(278.4)</td>
<td>(278.4)</td>
</tr>
<tr>
<td>assemblyseats</td>
<td>-24.88***</td>
<td>-484.3**</td>
<td>81.55</td>
<td>81.55</td>
<td>81.55</td>
</tr>
<tr>
<td></td>
<td>(9.571)</td>
<td>(254)</td>
<td>(109.7)</td>
<td>(109.7)</td>
<td>(109.7)</td>
</tr>
<tr>
<td>majority</td>
<td>-209.3</td>
<td>-97.37</td>
<td>-209.3</td>
<td>-97.37</td>
<td>-97.37</td>
</tr>
<tr>
<td></td>
<td>(266.4)</td>
<td>(245.8)</td>
<td>(266.4)</td>
<td>(254.8)</td>
<td>(254.8)</td>
</tr>
</tbody>
</table>

Observations: 240

R² (overall): 0.44

Robust standard errors in parentheses.
### Table 2: Robustness Checks on Assemblyseats Variable

<table>
<thead>
<tr>
<th>Variables</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(9.657)</td>
<td>(9.677)</td>
<td>(9.817)</td>
<td>(9.76)</td>
<td>(9.87)</td>
<td></td>
</tr>
<tr>
<td>majority</td>
<td>50.18</td>
<td>25.99</td>
<td>32.71</td>
<td>25.99</td>
<td>32.71</td>
</tr>
<tr>
<td>(108.7)</td>
<td>(122.6)</td>
<td>(115.8)</td>
<td>(122.6)</td>
<td>(115.8)</td>
<td></td>
</tr>
<tr>
<td>partyinstability</td>
<td>-177</td>
<td>-179.5</td>
<td>-98.22</td>
<td>-98.22</td>
<td></td>
</tr>
<tr>
<td>(262.6)</td>
<td>(290)</td>
<td>(254.5)</td>
<td>(265.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>regional1</td>
<td>-147.8</td>
<td>-131.5</td>
<td>-23.85</td>
<td>-23.85</td>
<td></td>
</tr>
<tr>
<td>(266.7)</td>
<td>(293.8)</td>
<td>(256)</td>
<td>(266.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>regional2</td>
<td>61.99</td>
<td>16.62</td>
<td>39.49</td>
<td>39.49</td>
<td></td>
</tr>
<tr>
<td>(109.4)</td>
<td>(123.2)</td>
<td>(116.7)</td>
<td>(123.2)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Controls for State Characteristics | YES | YES | YES | YES | YES |

Observations | 240 | 235 | 235 | 235 | 235 |

R² (overall) | 0.46 | 0.45 | 0.45 | 0.45 | 0.45 |

Robust standard errors in parentheses *** p<0.01, ** p<0.05

---

### Table 3: Robustness Checks on Majority Variable

<table>
<thead>
<tr>
<th>Variables</th>
<th>(11)</th>
<th>(12)</th>
<th>(13)</th>
<th>(14)</th>
<th>(15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>majority</td>
<td>-470.2**</td>
<td>-470.8**</td>
<td>-485.2**</td>
<td>-468.8**</td>
<td>-479.5**</td>
</tr>
<tr>
<td>(255.7)</td>
<td>(258.6)</td>
<td>(259.9)</td>
<td>(259.8)</td>
<td>(261.1)</td>
<td></td>
</tr>
<tr>
<td>partyinstability</td>
<td>61.99</td>
<td>16.62</td>
<td>39.49</td>
<td>39.49</td>
<td></td>
</tr>
<tr>
<td>(109.4)</td>
<td>(123.2)</td>
<td>(116.7)</td>
<td>(123.2)</td>
<td>(116.7)</td>
<td></td>
</tr>
<tr>
<td>regional1</td>
<td>-147.8</td>
<td>-131.5</td>
<td>-23.85</td>
<td>-23.85</td>
<td></td>
</tr>
<tr>
<td>(266.7)</td>
<td>(293.8)</td>
<td>(256)</td>
<td>(266.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>regional2</td>
<td>61.99</td>
<td>16.62</td>
<td>39.49</td>
<td>39.49</td>
<td></td>
</tr>
<tr>
<td>(109.4)</td>
<td>(123.2)</td>
<td>(116.7)</td>
<td>(123.2)</td>
<td>(116.7)</td>
<td></td>
</tr>
</tbody>
</table>

Controls for State Characteristics | YES | YES | YES | YES | YES |

Observations | 240 | 235 | 235 | 235 | 235 |

R² (overall) | 0.44 | 0.44 | 0.44 | 0.44 | 0.44 |

Robust standard errors in parentheses *** p<0.01, ** p<0.05
6 Bibliography


MARKET REACTIONS TO THE US SHALE GAS BOOM

NATALIYA LANGBURD

Yale University

ABSTRACT

I employ an event study methodology to assess the economic impact of shale gas on commodity prices, US upstream sectors, and US downstream sectors, using announcements of shale reserve and technology discoveries, mergers, and legislation between 2008 and 2013. I find that, first, shale reserve and technology announcements are positive supply shocks to the commodity market and do not initially benefit shale producers, associated with negative mean abnormal returns to future natural gas prices and to upstream stocks averaging -6%. Second, merger announcements benefit shale producers, but are not a shock to future commodity supplies, associated with positive abnormal returns to upstream stocks and no significant abnormal returns to US natural gas future prices. Third, Federal legislation cutting shale research is a negative demand shock to the market, associated with negative abnormal returns to future natural gas prices. The findings explain shale gas’s current effects on the market and predict how they will percolate through the economy.
DEMOCRACY AND THE EFFECTS OF NEGATIVE EXOGENOUS INCOME SHOCKS: EVIDENCE FROM INDIVIDUAL AND HOUSEHOLD DATA IN SUB-SAHARAN AFRICA

SAMSUN KNIGHT

Oberlin College

ABSTRACT

Are citizens of democracies cushioned from the vagaries of the business cycle? Using a comprehensive micro dataset of citizens from 24 countries of sub-Saharan Africa, I measure the survival of infants born to the same mother and the school attendance of individuals as they vary across different institutional regimes and different economic contexts. I find that democracy is correlated with significantly increased sensitivity to exogenous aggregate income shocks (as measured by drought) of the infant mortality rate and significantly reduced sensitivity to exogenous aggregate income shocks for school enrollment. The most consistently salient aspect of democracy with respect to these effects appears to be competitiveness of executive recruitment. The infant mortality findings are consistent with a political economy model, based on Acemoglu and Robinson (2005), wherein more autocratic governments respond more robustly to drought mainly due to fewer roadblocks in immediate implementation of policy.

1 Introduction

Do democracies offer their citizens inferior protection from negative exogenous aggregate income shocks? In the wake of the recent 2008 financial crisis, there was considerable talk of an emerging ‘Beijing consensus’, an idea that state-led economies unhindered by democracy were better equipped to for the fast-paced, globalized modern world (New York Times 2011). This intellectual movement rode strong anecdotal evidence of the West’s slow recovery from recession versus China’s immediate bounce.
back and quickly gained credence in both media and academic circles. It was not long in arriving in sub-Saharan Africa, where the choice between democracy and autocracy is real and pressing. At least one of the attractive aspects of Chinese-style autocracy is the idea that autocratic governments better respond to recession.

The differing ramifications of the financial crisis in the West versus in China are hardly rock-solid proof that one political model is more or less robust to negative shocks than the other. Regarding the relationship between democracy and economic outcomes, the conclusions of theory are far from clear and the empirical evidence is, at best, mixed (Sirowy and Inkeles 1990, Przeworski 1991, Ross 2006, Acemoglu et al 2008). Anecdotally, this uncertainty is well illustrated by the observation that all of the developed countries of the OECD are democracies, but a majority of these modern industrial economies first developed under autocratic regimes. Compounding this difficulty is the fact that democracy is obviously not randomly distributed at the national level, and so it is hard to distinguish its effects from those of other factors.

However, by mirroring the empirical approach of Kudamatsu (2012), it is possible to sidestep the problems inherent in the macroeconomic data by using the individual and household micro data of the Demographic and Health Surveys, collected with uniform methodology by USAID, to approach the issue. I focus on sub-Saharan Africa both for its large variance in democratic institutions, widely dispersed across disparate geographical regions, and for the immediate relevance of this particular question to this particular region. I focus on infant mortality and schooling both because of the availability of the data and their salience in the quality of contemporaneous well-being (with respect to infant mortality) and prospective well-being (with respect to schooling). While the challenge of establishing causation remains elusive, I am able to combine the data with the comprehensive democracy metrics of the Polity IV project to robustly estimate the correlations between democracy, drought and well-being. I find that the infant mortality rate is worse protected from drought under greater democracy, while school enrollment is better preserved under greater democracy.

2 Theoretical Model

In their 2005 book, “Economic Origins of Dictatorship and Democracy”, Acemoglu and Robinson describe two models of political economy under democracy and under dictatorship, respectively, to facilitate the analysis of
economic outcomes under either regime type. These models will serve as the starting point for the theoretical model proposed by this paper.

First, Acemoglu and Robinson assume that voters have single-peaked preferences – each voter has an ideal policy preference, and the further away from that ideal, the less they prefer the policy. Each voter is also assumed to carry equal political power. In both direct democracy and in a party representative democracy, this leads to political outcomes that are the ideal policy preference of the median voter (Downs 1957). In poor countries, where the median voter is poorer than the mean voter, one would expect this to lead to redistributive policymaking, either through direct redistribution or greater provision of public goods or both.

For non-democracies, Acemoglu and Robinson assume that the ruling elite are trying to enrich themselves as much as possible while still minimizing the risk of a revolution against them. In general, the autocratic government will implement regressive policies to enrich themselves at the expense of the general populace. However, the revolution constraint is highly contingent on the opportunity cost of revolution, which is largely determined by the everyday income of one who might otherwise become a revolutionary. If the oligarchs have maximized their personal gain in a normal year, then a negative exogenous shock will take them too far from their desired balance between risk of revolution and profits, and the average citizen will have a sub-optimal (from the perspective of the rulers) opportunity cost of revolting. The ruling class will thus have an immediate incentive to quickly attenuate the negative impacts of a negative aggregate income shock in order to better secure their advantageous standing in society. In the empirical literature, this threat is confirmed as quite salient: Brückner and Ciccone found in a 2011 paper that exogenous negative shocks are significantly correlated with political concessions in the form of greater democratization. Here, I am investigating the obverse phenomenon, of economic concessions.

The above theory is then augmented with a “policy lag”, or the speed with which policies are decided upon, passed into law, and subsequently implemented. The assumption is made that this pace of policymaking is maximized under a purely autocratic government, where laws are changed by the decision of a single leader versus in a system of checks and balances. As both autocracies and democracies have strong incentives to respond to negative shocks, as outlined above, the difference in their effective reactions is theorized to be a result of differences in the policy lag. However, since their motivations for attenuating negative shocks' impact differs, we may expect the focuses of their relief to differ as well.
It is worth noting that, according to Acemoglu and Robinson, autocracies would target methods to raise the income of the citizenry that directly address immediate sources of malcontent, such as food and health, while the model predicts that an autocracy would emphasize less the negative impacts of a negative shock that did not drive people so immediately to revolt. Additionally, infant mortality is likely more exclusively responsive to contemporaneous conditions, while school enrollment would incorporate expectations of the future.

The theoretical connection between economic contractions or expansions and infant mortality and school enrollment rates is direct and supported by the empirics, particularly with regard to sub-Saharan Africa. Infant mortality is broadly used as an indicator of chronic hunger by policy makers and international organizations (Zweifel and Navia 2000). The reasoning behind a positive correlation between hunger and personal income is clear: along with the ability to purchase any good, income shocks may cut into the ability to purchase food. In addition, a negative income shock may affect child mortality in myriad other ways: lower personal income may lead to fewer private expenditures on child health inputs and could lead to interruptions and failures in public health services.

With regards to school enrollment, a simple intertemporal choice model considering schooling as a form of investment shows that school enrollment may increase or decrease in the face of a drop in personal income depending on how the income and substitution effects interact. An economic contraction may necessitate child labor to support the household, or it could reduce the potential child wage and hence the opportunity cost of attending school. Recent studies of sub-Saharan Africa suggest that the relationship is uniformly positive: a drop in income leads to a drop in enrollment (Ferreira and Schady 2009, Grimm 2011).

Therefore, both of these metrics are indicators of the negative income effects of a negative aggregate income shock on infant mortality and school enrollment, respectively. To the extent that autocracies are faster at responding to negative income shocks, they will be more effective in combating said shocks’ negative impacts. However, since autocracies have a more limited incentive to help their citizens only to the point to which they are no longer on the verge of revolt, this aid may similarly be more limited in its scope.
3 Data and Metrics

For data on infant mortality and school enrollment, this paper uses the Demographic and Health Surveys. These data are utilized particularly for their broad regional reach, encompassing all of the nations of sub-Saharan Africa, made invaluable by their methodological consistency. Additionally, the surveys collect individual data on mothers and births, constructing retrospective histories of childbirth for each mother interviewed. Women report when each of their children were born and, if they died, when they died. The mean number of births per mother is four, so almost every mother provides a small window into the effects of any pertinent changes between her subsequent births. I reformat the surveys so that each individual birth is a single observation in the year of birth, with a dummy variable that is equal to one if the baby dies within thirteen months and zero otherwise. In this way, I am able to perform fixed effects analysis around each individual mother to control for myriad confounding factors that are particular to each individual mother.

The household surveys afford fewer opportunities for viewing a long window of time, but they still retain the advantage of yielding individual micro observations of school enrollment. Each survey gathers statistics pertaining to individual households including the number of residents attending school this year and last. Each of the surveys therefore offers a snapshot of school enrollment at the individual level within a given context of certain income levels and governing institutions.

For a measure of democracy, following in the footsteps of many other authors studying democracy with empirical methods, I rely on the Polity IV project of Marshall and Gurr (2013). The Polity IV measurements have been shown to be highly consistent with other indices of democracy and are complemented by the ready availability of disaggregated democratic characteristics, allowing for a piecemeal analysis of the separate aspects of popular governance (Munck and Verkuilen 2002).

The aggregate democracy measure is a weighted combination of individual measures of the regulation of executive elections and (separately) of party participation in elections, the openness of executive elections, the competitiveness of executive recruitment and (separately) of party elections, and constraints on executive power.

For data on exogenous income shocks, I rely on rainfall data from Miguel, Satyanath and Sergenti (2004), originally collected from the Global
Precipitation Climatology Project by NASA, which uses data from gauge stations and satellite data to estimate global precipitation. I use log rainfall, and specifically deviation from mean country log rainfall, to serve as a metric of exogenous income shocks, following in the footsteps of Bruckner and Ciccone (2011) who clearly demonstrated that rainfall shocks are a valid instrument for exogenous GDP shocks in the largely agricultural economies of sub-Saharan Africa.

4. **Empirics**

To investigate the relationship between infant mortality, democracy and the business cycle, I estimate the following linear probability model for each individual mother:

\[
Y_{it} = \alpha_{it} + \beta_{it} + \chi_{it} + \delta_{it} + y_{1i} * D_{it} + y_{2i} * RA_{it} + y_{3i} * D_{it} * RA_{it} + \kappa_{it} + \epsilon_{it}
\]

(1)

Alternatively, a quadratic term may also be included:

\[
Y_{it} = \alpha_{it} + \beta_{it} + \chi_{it} + \delta_{it} + y_{1i} * D_{it} + y_{2i} * RA_{it} + y_{3i} * RA_{it}^{2} + y_{4i} * D_{it} * RA_{it} + y_{5i} * RA_{it}^{2} + y_{6i} * D_{it} * RA_{it}^{2} + \kappa_{it} + \epsilon_{it}
\]

(2)

The dependent variable \( Y \) is a dummy that equals 1 if baby \( i \) born to mother \( m \) of country \( c \) in year \( t \) dies within 13 months of childbirth, and 0 otherwise. \( \alpha_{it} \) measures mother fixed effects, \( \beta_{it} \) measures country trend fixed effects, \( \chi_{it} \) measures birth order fixed effects, and \( \delta_{it} \) measures birth year fixed effects. \( D_{it} \) is the democracy variable, and in the case of regressions on the disaggregated democracy component variables, may represent a column vector of democracy characteristics. (In such a case, \( \gamma \) would constitute a row vector, respectively.) \( RA_{it} \) represents deviation of log rainfall from the respective country’s mean log rainfall, where such rainfall is above the mean. \( RB_{it} \) represents the opposite, deviation from mean country log rainfall when said deviation is below the mean. \( \kappa_{it} \) is a vector of exogenous covariates, including a dummy for gender, a dummy for multiple births (twins, triplets, etc.) a dummy for small birth space (born within 24 months
of mother’s last child) and a measure of the mother’s age. \( \varepsilon_{\text{ovr}} \) is the error term.

The primary results for infant mortality are seen in table 1. The positive correlation between infant mortality and drought is augmented by increased democracy, consistent with a world where autocracies act swifter to attend to the most important needs of victims of droughts. Without drought, democracy significantly reduces infant mortality by a small amount, consistent with the results of Kudamatsu (2012).

In table 2, the democracy index is broken down into its component variables. Executive recruitment competitiveness is clearly the driver of the observed weaker response of democracies, corroborating the theory that policy lag is a large driver of this difference – in line with the example from the 2011 Kenya drought, the power of executive to damp down on the jockeying for power that hampers relief efforts. The next most salient aspects are all measures of the consolidation of executive power, further corroborating this theory as well as the “big man” theory of politics in Africa, that the executive carries more weight than other bodies.

To investigate the relationship between school enrollment, democracy and the business cycle, I estimate the following linear probability model using straightforward OLS:

\[
Y = X' \alpha + y_1 * R + y_2 * D + y_3 * D * R + \varepsilon
\]  

(3)

The dependent variable \( Y \) is equal to one if the child has attended school this year. \( \alpha \) is a vector of control variables, including sex of child, age of child (within 5-18 years), age of head of household, sex of head of household, family ownership of durables, country log per capita GDP, and % of GDP that is Overseas Direct Aid. \( R \) measures rainfall below mean, and \( D \) is the democracy indicator, and may be a single variable or a vector of variables. \( \varepsilon \) is the error term.

In table 3 the primary results for school enrollment is described. Here the correlation with democracy and drought is inverted: a greater degree of democracy is associated with a lesser sensitivity to negative economic shocks. The relatively severe drop in school enrollment under autocracy suggests, somewhat paradoxically, that even though autocracies are better addressing citizens’ health and food needs as measured by the infant mortality rate, they do not more effectively insulate citizens from the overall
negative income shock. More children drop out of school in a drought in autocracies.

Table 4 details the same regression with democracy broken down into its component variables. Since the schooling dataset is less extensive than the infant mortality data, the high correlation between the components made impossible a regression with all the variables involved. However, by regressing on each component one by one, it is clear that the most consistently important component is again the competitiveness of executive recruitment. This component may be interpreted as indicating the level to which the government is truly inclusive of the majority of voters, which in the median voter model of Acemoglu and Robinson reinforces the government’s commitment to improving the lot of the average citizen but does not have as direct an impact on policy lag as the component for executive constraints or the component for the competitiveness of parties.

The component for executive constraints is found to be significant, but it is not the most important variable. Therefore, it appears that a swift response is less important for school enrollment than is an expectation for an eventual response, which does not contradict the predictions set forth in this paper’s model.

Absent changes in log per capita GDP, the results clearly suggest that more democracy is associated with increased school enrollment, consistent with the findings of Harding and Stasavage (2012) and Stasavage (2005).

4 Conclusion

From the empirical findings of this paper, it appears that the “Beijing Consensus” is not wrong, but neither is it particularly strong. Yes, more autocratic governments may respond more robustly to negative income shocks to address immediate concerns of life and death, but the overall economy does not fare better for it—the fact that school enrollment weathers economic storms better under democracies demonstrates that overall individual incomes are better protected under democracy. This makes sense given the disparate goals of countercyclical policy under democracy versus under autocracy: in democracy, the policy is truly trying to improve wellbeing, while under autocracy, the policy is just trying to address the kind of immediate concerns that might break the camel’s back. In addition to all of this, the advantage of dictatorships over democracies rests largely in the ability of dictatorships to act swiftly and strictly in the national interest. This implies that when a crisis is truly massive and political will is united, a democracy may be just as effective if not more so than an autocracy. This
reconciles the findings of this paper with such standards of the literature as Sen’s point that no famine has ever occurred in a democracy (Sen 1999). If a crisis is truly severe, a democracy retains the wherewithal to address it.

The policy implications of this paper are clear: democracies need better health aid in drought to prevent their citizens from receiving real, tangible punishment for an innately sclerotic political process. And since democracies are correlated with tangible health and education gains for a populace in the absence of a negative income shock, spreading democracy is still an important mission for global leaders. Aid to democracies in times of negative income shocks would provide a dual use of supporting vulnerable populations and helping stem the ideological advance of autocracy across the world.

The theoretical implications, however, are less clear. The mixed findings, with school enrollment reacting in the opposite manner to infant mortality suggest that more steps must be taken to clarify whether the model put forth here actually holds. While it is possible to reconcile the findings with the given theory, the reasoning is less than airtight. More data must be collected. More work must be done.
5 Bibliography


<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>10175</th>
<th>10176</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.825</td>
<td></td>
</tr>
<tr>
<td>Family Below Mean/Democracy Index Interaction Term</td>
<td>1.179</td>
<td></td>
</tr>
<tr>
<td>Family Below Mean</td>
<td>3.284</td>
<td>3.014</td>
</tr>
<tr>
<td>Total GDP % of GDP</td>
<td>0.008</td>
<td>0.006</td>
</tr>
<tr>
<td>Expenditure (public services)</td>
<td>0.007</td>
<td>0.006</td>
</tr>
<tr>
<td>Age</td>
<td>0.006</td>
<td>0.005</td>
</tr>
<tr>
<td>Sex</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Attended School Last Year</td>
<td>0.003</td>
<td>0.003</td>
</tr>
</tbody>
</table>

**Note:** *P < 0.05, **P < 0.01, ***P < 0.001.*
DEMONCACY ON DEMAND? LOCAL GOVERNMENT REPRESENTATION AND ACCESS TO BPL CARDS IN INDIA

BRIAN GOGGIN
Georgetown University

ABSTRACT

In this paper I test whether attendance at local government meetings by the poor affects their likelihood of receiving government benefits in India. More specifically, I test whether the attendance of disadvantaged groups—mainly the poor and lower castes—at local Gram Panchayat meetings increases their likelihood of receiving below the poverty line (BPL) cards, which allow disadvantaged groups to buy food at a subsidized price. I find that while there are little to no effects of attendance by itself at these local government meetings, personal connections with the government, and, in particular, personal acquaintance with a government official, increases the likelihood of possessing a BPL card by 15 to 24 percentage points. These results show that while political participation at the local level may be desirable for promoting civil society, there is little evidence of its effects for accurate targeting of government benefits. Instead, other factors, especially personal connections with government officials, likely overshadow the efficacy of political participation alone.

1 Introduction

Introduced by the 73rd Constitutional Amendment Act in 1993, Gram Panchayats (GP) represents the lowest level of democratic governance in India. GPs typically have between 500 to 10,000 constituents spread out through a number of villages. Members from these constituencies usually elect 5-20 members to serve on the GP council based on electoral wards divided evenly by population. These bodies main duties are to administer local infrastructure projects and to identify welfare recipients of state and
national development programs. In particular, Gram Panchayats are responsible for
surveying their constituents in order to determine who qualifies a below the poverty
line (BPL) ration card. In this paper, I test whether attendance at local government
meetings has a positive effect on the proper targeting of BPL ration cards.

BPL ration cards form a backbone of the nation’s public distribution campaign,
in which the federal government buys up grain from farmers, distributes to the states
based on their share of poor, and then the states decide how to distribute the grain to
public distribution shops (shops at which villagers can buy reduced price food with BPL
cards, often known as PDS shops). These cards also make people eligible for a number
of state-sponsored welfare schemes, including subsidized food and free hospitalization.
Qualification requirements for BPL ration cards differ on state-specific criteria. For
example, in Punjab the necessary requirements are an annual income less than 20,000
rupees, as well as families not having more than 2 hectares of land or a TV or a refrigerator.

Panchayats bear the large brunt of conducting village surveys to determine eligible
beneficiaries and the preliminary list of BPL card recipients is supposed to be finalized
at a Gram Sabha meeting. These are periodic (2-4 times annually) meetings of the
GP council at which all voters from a panchayat constituency are invited to audit and
supervise the workings of the board, including BPL card selection and ratification of the
GP budget. After the GP surveys the eligible households in its constituency for BPL
cards, the government then distributes these cards to the households. Many believe
GP politicians hold undue control over the process, and villager monitoring is limited.
Some have estimated that up to two-fifths of all BPL cards in India are with non-poor
families.

In this paper, I use a OLS, and probit regression techniques to test the possible
effect of attendance at a GP meeting on accurate BPL card targeting and find little
to no significant effects. On the other hand, I do find significant effects for attend-
ance interacted with having a personal connection with government officials. These
results provide evidence against the efficacy of direct political participation and support
theories of political patronage in India.

The rest of this paper will proceed as follows. Section 2 reviews previous literature
on political representation and government decentralization in the developing world.
Section 3 discusses the data used, as well as the empirical strategy while Section 5
discusses the results. Section 6 concludes.

1 Department of Food & Public Distribution, New Delhi Grc. of India. See also Grc. of Uttarak-
hand, Department of Food and Civil Supplies
2 As Grain Piles Up, India’s Poor Still go Hungry, New York Times
3 Department of Planning, Grc. of Pahjab
4Besley, Pande, and Rao (2007)
5Ram, Mohanty, and Ram (2009)

In this paper, I use a OLS, and probit regression techniques to test the possible
effect of attendance at a GP meeting on accurate BPL card targeting and find little
to no significant effects. On the other hand, I do find significant effects for attend-
ance interacted with having a personal connection with government officials. These
results provide evidence against the efficacy of direct political participation and support
theories of political patronage in India.

The rest of this paper will proceed as follows. Section 2 reviews previous literature
on political representation and government decentralization in the developing world.
Section 3 discusses the data used, as well as the empirical strategy while Section 5
discusses the results. Section 6 concludes.

1 Department of Food & Public Distribution, New Delhi Grc. of India. See also Grc. of Uttarak-
hand, Department of Food and Civil Supplies
2 As Grain Piles Up, India’s Poor Still go Hungry, New York Times
3 Department of Planning, Grc. of Pahjab
4Besley, Pande, and Rao (2007)
5Ram, Mohanty, and Ram (2009)
2 Literature Overview

The rise of government decentralization in the developing world has seen a concomitant interest in the academic world. There have been several attempts to empirically test the effects of decentralization in the past. At a macro level, Triesman (2002) used cross country data to find that federal systems of government and government expenditures at the subnational level were associated with poorer levels of sanitation and higher rates of corruption. More micro-level studies have been more favorable, showing increases in access to sanitation, basic schooling, and public expenditures for the poor (Santos 1998, Fugnet 2004). Several household-level survey studies of local meetings in the developing world are focused on Indian Gram Panchayats in particular. Chattopadhyay and Duflo (2001) show that random assignment of women to GP council heads in West Bengal leads to increased investment in infrastructure for rural women. Similarly, Bardhan and Moookherjee (2010) find that reservations on GP councils for scheduled castes and tribes (SC and ST) were associated with increases in benefits provided to these groups.

Aside from empirical research on decentralization, several political scientists have studied the efficacy of political participation and its relationship to government benefits. In the United States and in much of the developed world Verba et al. (1995) find that there are lower rates of participation, as measured by voting, campaign working, and protesting, among economically disadvantaged groups—and in particular, recipients of means-tested benefits. On the other hand, participation by disadvantaged groups in the developing world is often equal to, or greater than, that of the rest of the population. For example, Montavalo (2008) shows that in Latin America, unequal participation by socioeconomic characteristics is not nearly as pronounced as in the developed world. In particular, he finds that wealth is negatively associated with attendance at municipal level meetings, although there remains an expected positive correlation with education. Rates of participation among the poor are even higher in India, and Chandra (2004) offers a theory of patronage democracy that may explain this phenomenon with regards to voting. That is, high rates of voting among disadvantaged groups in India are likely driven by the desire for political rewards from politicians, who distribute goods more to their followers in a clientalistic fashion, particularly along caste or ethnic lines.

There have been a few qualitative studies to combine the questions of participation and the effectiveness of decentralization. Ghatak and Ghatak (2002) find qualitative evidence from survey responses in West Bengal that participants at Gram Sabha meetings repeatedly emphasize their monitoring role. In particular, they document many instances of complaints about the failures of development program administration, badly needed public infrastructure projects, accusations of corruption in local expenditures, and the lack of targeting of government programs to women, lower castes, or the neediest constituents.

Perhaps the closest empirical study to this paper comes from Besley, Pande, and Rao (2005). Here the authors use household-level survey data from South India to find that the landless and lower castes are more likely to attend Gram Sabha meetings. Furthermore, the authors find that landless and lower castes are more likely to receive benefits.
BPL cards in villages where a Gram Sabha meeting was held. Although holding a
meeting is significant for giving disadvantaged groups access to BPL card benefits, it is
unclear whether their attendance at these meetings was the driving force behind this
outcome. For this reason, I hope to provide an extension on their findings with my
paper in examining whether the attendance by the poor at meetings is instrumental
in the functions of Gram Panchayat’s distribution of BPL cards. In approaching my
question, I hope to draw from much of the work done by Besley, Pande, and Rao (2005)
by emulating their controls and model. This question not only adds to the literature on
effective decentralization in India, but also to the functioning of democratic exchange
and monitoring within decentralized institutions.

3 A Simple Model of Village Participation

3.1 The Main Hypothesis and Controls

I distill the research question of this paper into two related and testable hypotheses
about whether attendance at Gram Sabha meetings matters for receipt of BPL cards.
These hypotheses are motivated by both the literature on participation and decentral-
ization above. In each hypothesis below, I use three independent variables. First, I
use the the village-level measures of number of villagers attending and percentage of
villagers attending from a given household’s survey sample. Secondly, I use a household-
level dummy variable for household attendance if anyone from a given household has
attended a meeting. The purpose behind the use of all three independent variables is to
test all possible effects of participation. For example, perhaps household attendance is
not as relevant as the percentage of other villagers that attend, or perhaps it is actually
more relevant because a self-interested household first and foremost ensures that their
name is on the list of BPL card recipients.

1. Representation: The first hypothesis is inspired by literature on participation.
   This hypothesis states that representation from one’s village at a Gram Sabha
   meeting could increase the visibility, and therefore, allocation of BPL cards to
   that village. That is to say, perhaps there is a unambiguous positive impact of the
   number and percentage of villagers attending, as well as household attendance on
   the probability of being noticed by a GP council, which increases the probability a
given household receives a BPL card.

2. Monitoring: Relatedly, I will test whether the attendance of people from one’s
   own village increases the likelihood that people below the poverty line receive
   a BPL card. That is, is there a monitoring role whereby participation by peo-
   ple from one’s own village increases the proper targeting of BPL card benefits?
   After all, the purpose of Gram Sabha meetings is to have the citizenry audit
   and supervise the working of the GP council. If this system is working as it is
   supposed to, increased participation should increase the accurate targeting of

BPL cards in villages where a Gram Sabha meeting was held. Although holding a
meeting is significant for giving disadvantaged groups access to BPL card benefits, it is
unclear whether their attendance at these meetings was the driving force behind this
outcome. For this reason, I hope to provide an extension on their findings with my
paper in examining whether the attendance by the poor at meetings is instrumental
in the functions of Gram Panchayat’s distribution of BPL cards. In approaching my
question, I hope to draw from much of the work done by Besley, Pande, and Rao (2005)
by emulating their controls and model. This question not only adds to the literature on
effective decentralization in India, but also to the functioning of democratic exchange
and monitoring within decentralized institutions.
disadvantaged groups. Therefore, this hypothesis predicts that increased village-level attendance in the form of number attending and percentage attending, as well as increased household-level attendance, will increase the household-level probability of receiving a BPL card for poor households.

4 Data

The Indian Human Development Survey (2005) “is a nationally representative, multi-topic survey of 41,554 households in 1503 villages across India.” The survey includes 1116 variables on a wide variety of development indicators. While the data includes a lot of detail on each household member, I restrict my sample to household heads (usually adult males) because these are the survey respondents for household level variables. Therefore, the few individual-level variables that I use (for example, literacy) are for the household heads. I further restrict my sample to rural districts as defined by the Indian census because Gram Panchayats are local governments in rural areas. After these adjustments, I am working with 26,743 observations. The survey question from which I take my independent variables of interest is a yes or no question for attendance at local “village panchayat/nagarpalika/ward committees in the last year.” The key dependent variable is holding a below poverty line (BPL) card. This is constructed from the survey question asking the following: “Does the household have a ration card? If yes, what type of card is this? BPL=1.”

The data contains village-level, district-level, and state-level, and household level IDs. Because part of the interest of this paper lies at the village level and another part at the household level, I will elaborate on summary statistics for both below. At the village level there are a number of interesting measures of the quality of institutions, such as whether a given village has a Panchayat Bhavan (local government building), community centre, or a police station and if not, the distance to the closest. The particular strength of this data is that it includes abundant household-level information on a wide variety of covariates of relevance which have been discussed in the literature review and model sections. These are variables which have shown to be relevant for attendance at Gram Sabha meetings, holding a BPL card, or both. These include measures of being below the federally-defined poverty line, income, landlessness, caste status, religious affiliation, years of education, literacy, gender, as well as something I have yet to see tested in the literature- whether a given household has a family member or knows someone who is a member of the GP council. An important note for the remainder of the paper is that the BPL card refers to those households who have a below the poverty line ration card while the dummy variable Poor refers to those households who are below the federally-defined poverty line, as described in Section 1.

6ihds.umd.edu
5 General Empirical Strategy

In order to empirically test my hypotheses, I make some adjustments to a simple equation taken from Besley, Pande, and Rao (2005). Below I highlight three forms of this empirical specification using the three different independent variables specified in the model above. These equations are as follows:

Number of villagers attending

\[ b_{vds} = \gamma_d + \delta_v + \beta V_{vds} + \xi c_{vds} + \alpha N_{vds} + \theta (c_{vds} * N_{vds}) + \epsilon_{vds} \]  

(1)

Percentage of villagers attending

\[ b_{vds} = \gamma_d + \delta_v + \beta V_{vds} + \xi c_{vds} + \alpha (P_{vds}/N_{vds}) + \theta (c_{vds} * (P_{vds}/N_{vds})) + \epsilon_{vds} \]  

(2)

Household attendance

\[ b_{vds} = \gamma_d + \delta_v + \beta V_{vds} + \xi c_{vds} + \alpha H_{vds} + \theta (c_{vds} * H_{vds}) + \epsilon_{vds} \]  

(3)

where \( b_{vds} \) is a dummy for holding a BPL card for household \( i \) in village \( v \); \( \gamma_d \), and \( \delta_v \) represent district, and state level fixed effects respectively, and \( \epsilon_{vds} \) represents relevant household level characteristics. The independent variables of interest are \( N_{vds} \), \( (P_{vds}/N_{vds}) \), and \( H_{vds} \), which represent the number of villagers attending, percentage of villagers attending, and household attendance respectively. Therefore, the \( \alpha \) and \( \theta \) coefficients represent the representation effect of attendance while the coefficients are the coefficients of interest. Lastly I include a vector of village level characteristics in the term \( V_{vds} \), which includes measures of the quality of institutions as well as village population. I cluster standard errors at the village level in all regressions. I also try to control for village population by including a population category variable. This is the best estimate of a village’s population given in IHDS and includes three categories coded as 1, 2, or 3: 1- “<1000”, 2- “1000 to 5000”, 3- “>5000.” These village-level control variables are included in the \( V_{vds} \) term above. This control for village population is included in the \( V_{vds} \) term.

5.1 Robustness Check: Probit Model

As a robustness check on my OLS model, I also test my hypotheses with a probit model of the form below using number attending as an example:

\[ Pr( b_{vds} = 1/X ) = \phi(\gamma_d + \delta_v + \beta V_{vds} + \xi c_{vds} + \alpha N_{vds} + \theta (c_{vds} * N_{vds}) + \epsilon_{vds}) \]  

(4)

where \( \phi \) represents the cumulative standard normal distribution function. There are a number of advantages of a probit model over the linear probability model when dealing with binary dependent variables, especially in the case of BPL card holdership.
In this case, for particularly well-off or disadvantaged individuals, the probability of receiving a BPL card may exceed one or be less than zero. In this case these values bias the regression estimates from the OLS results. Furthermore, because probabilities cannot exceed one, the effect on the probability that a given household has a BPL card given a change in attendance is likely nonlinear.

5.2 Issues and Limitations
There are problems with the endogeneity of attendance. This is primarily because there could be a reverse causality bias in that receiving BPL cards could incentivize political participation. However, these problems would likely overstate the relationship between number attending and percentage attendance and receipt of BPL cards and because I see no significant relationship to begin with, the issue is likely not crucial for the qualitative results. There are other limitations for the independent variables of attendance. Despite having knowledge of whether someone attended at all in the past year, the survey includes no data on how often these meetings are held or more nuanced details of respondents’ involvement at Gram Sabha Meetings. For instance, the survey data used does not include nuanced information about how many times a given household attended a Gram Sabha meeting, what percentage of the overall village attended the meetings, or specific monitoring actions taken during the Gram Sabha meeting.

6 Results
6.1 Results for Attendance
Table 1 reports the OLS regression results using number attending, percent attendance, and household attendance as independent variables of interest. As column 1 shows, there are no statistically significant effects for number attending in either a representation or a monitoring role (as the interaction term shows). Similarly, column 2 shows that there are no statistically significant results using percentage attendance as a measure of attendance. A test of joint significance for number attending and percentage attending, along with their interaction terms with being poor, also shows no significance so there is little evidence that there is a representation or monitoring role of attendance when measured by number attending or percent attending. However, column 3 reveals a slightly different story. Although there is still no statistical significance on the coefficient for attendance, the interaction term between household attendance and being poor is significant at the 10% level. This signals that household attendance at a meeting can have some monitoring role by guaranteeing that they are on the BPL card list. All of the coefficients on relevant household characteristics of disadvantage are significant with the expected signs.

In this case, for particularly well-off or disadvantaged individuals, the probability of receiving a BPL card may exceed one or be less than zero. In this case these values bias the regression estimates from the OLS results. Furthermore, because probabilities cannot exceed one, the effect on the probability that a given household has a BPL card given a change in attendance is likely nonlinear.

5.2 Issues and Limitations
There are problems with the endogeneity of attendance. This is primarily because there could be a reverse causality bias in that receiving BPL cards could incentivize political participation. However, these problems would likely overstate the relationship between number attending and percentage attendance and receipt of BPL cards and because I see no significant relationship to begin with, the issue is likely not crucial for the qualitative results. There are other limitations for the independent variables of attendance. Despite having knowledge of whether someone attended at all in the past year, the survey includes no data on how often these meetings are held or more nuanced details of respondents’ involvement at Gram Sabha Meetings. For instance, the survey data used does not include nuanced information about how many times a given household attended a Gram Sabha meeting, what percentage of the overall village attended the meetings, or specific monitoring actions taken during the Gram Sabha meeting.

6 Results
6.1 Results for Attendance
Table 1 reports the OLS regression results using number attending, percent attendance, and household attendance as independent variables of interest. As column 1 shows, there are no statistically significant effects for number attending in either a representation or a monitoring role (as the interaction term shows). Similarly, column 2 shows that there are no statistically significant results using percentage attendance as a measure of attendance. A test of joint significance for number attending and percentage attending, along with their interaction terms with being poor, also shows no significance so there is little evidence that there is a representation or monitoring role of attendance when measured by number attending or percent attending. However, column 3 reveals a slightly different story. Although there is still no statistical significance on the coefficient for attendance, the interaction term between household attendance and being poor is significant at the 10% level. This signals that household attendance at a meeting can have some monitoring role by guaranteeing that they are on the BPL card list. All of the coefficients on relevant household characteristics of disadvantage are significant with the expected signs.
6.2 Results for Group Connections

The results above show little to no evidence that political participation by neighbors has any effect on the receipt of BPL cards but a little more evidence that attendance by a household itself has some monitoring effect. While this is an interesting finding and shows that there could be some instrumental role by Gram Panchayats for the poor, it is useful to compare this effect to other hypotheses in the literature. For example, some authors have suggested that connections with government that can give individuals a leg up in receiving BPL ration cards. Perhaps then the most important relationship with government is not simply to let one’s voice be heard by political participation, but to be in the in crowd with the government.

In Table 2, I test this hypothesis using proxies for connections with government officials. Unfortunately the IHDS survey does not have much further information about Gram Panchayat politics than what I use in Table 2. For example, Besley, Pande, and Rao (2005) have information on the specific makeup of GP boards, whether they have reservations, the electoral process for each village, etc. However, I do have information on whether a household has someone or knows someone on a GP council (variable called KnowOfficial). Columns 1 and 4 show that this variable is statistically significant for households which are below the poverty line using both OLS and probit regressions, suggesting that knowing someone on a GP board increases the likelihood of having a BPL card by about 15% (column 4). Furthermore, columns 2 and 5 show that having an acquaintance or relative in the government civil service is statistically significant for poor households, increasing the likelihood of having a BPL card by about 24% (column 5). Lastly, columns 3 and 6 show that being a part of a caste association has a statistically significant positive differential effect for poor households. Caste associations are formal organizations which lobby politicians on behalf of a common caste group. In general, these results show that the effects of having government connections or being in a group lobbying for patronage benefits is much more effective for poor households than attendance at local meetings. These results are consistent with previous studies of connections with government officials [Bardhan and Mookherjee (2010) or Chandra (2004)], which show that the importance of legal reservations for spots for scheduled castes and tribes (SC and ST) on GP councils or the extensive patronage system in India.

7 Conclusion

Taking into account all the evidence from the empirical results, I do not draw an encouraging picture about the functional relevance of political participation in local democratic institutions in India. While there is no apparent relationship between number of villagers or percentage of village sample attending, household attendance for the poor is a statistically significant predictor of receipt of BPL cards in OLS and probit regressions. While this is a decent sign for the functionality of local governance in India, these results pale in comparison to the importance of personal connections with government.
the government, either in the form of holding a government job, personally knowing a GP council member, or knowing someone in the government service. While there is certainly a degree of uncertainty with my results, the biggest lesson from these results is that the often-suspected channels of government connections, such as clientilistic relationships between politicians and their friends, play a more dominant role in the distribution of BPL cards than political participation.

One important caveat for these results is that the quality of my data on attendance at Gram Sabha meetings is dubious. Although I have data on whether a household attended a Gram Sabha meeting anytime in the past year, I do not have any information on how active attendees were at the meeting, on how much of the village actually attended, or how many times someone attended. On the other hand, Besley, Pande, and Rao (2005) have more detailed survey data from their more localized paper, and find that holding a Gram Sabha meeting is very important for the targeting of BPL cards to disadvantaged groups. One particular advantage of the IHDS data in my paper is the large amount of covariates available to explore alternative hypotheses to the role of participation. I use some of these additional variables to compare the results of attendance at Gram Sabha meetings with the alternatives of connections with government officials.

However, future avenues of research include more accurately testing this potential relationship with more detailed measures of participation, activeness at Gram Sabha meetings, or with better measures of controls such as frequency of meetings offered. Better data like this exists with localized survey experiments such as Besley, Pande, and Rao (2005). Therefore, similar surveys could be done in the future centered around the topics of political participation. Lastly, my discovery of positive relationships between having government connections and receiving BPL cards suggests future research into the importance of not only social networks in India. Therefore, future research on the relevance of participation in civil society and access to government benefits is warranted.
### Table 1: Attendance OLS Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nattending</td>
<td>0.00124</td>
<td>(0.00108)</td>
<td></td>
</tr>
<tr>
<td>NattendingXPoor</td>
<td>0.00128</td>
<td>(0.00162)</td>
<td></td>
</tr>
<tr>
<td>Perc. Attendance</td>
<td>0.00706</td>
<td>(0.0240)</td>
<td></td>
</tr>
<tr>
<td>Perc. AttendanceXPoor</td>
<td>0.0482</td>
<td>(0.0382)</td>
<td></td>
</tr>
<tr>
<td>HouseholdAttendance</td>
<td>0.00573</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HouseholdAttendanceXPoor</td>
<td>0.0334*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>0.0900***</td>
<td>0.0824***</td>
<td>0.0882***</td>
</tr>
<tr>
<td></td>
<td>(0.0142)</td>
<td>(0.0152)</td>
<td>(0.0195)</td>
</tr>
<tr>
<td>OBC</td>
<td>0.0290***</td>
<td>0.0292***</td>
<td>0.0291***</td>
</tr>
<tr>
<td></td>
<td>(0.0102)</td>
<td>(0.0102)</td>
<td>(0.0102)</td>
</tr>
<tr>
<td>SC</td>
<td>0.116***</td>
<td>0.116***</td>
<td>0.115***</td>
</tr>
<tr>
<td></td>
<td>(0.0113)</td>
<td>(0.0113)</td>
<td>(0.0113)</td>
</tr>
<tr>
<td>ST</td>
<td>0.133***</td>
<td>0.132***</td>
<td>0.132***</td>
</tr>
<tr>
<td></td>
<td>(0.0187)</td>
<td>(0.0188)</td>
<td>(0.0188)</td>
</tr>
<tr>
<td>Years Schooling</td>
<td>-0.0156***</td>
<td>-0.0157***</td>
<td>-0.0156***</td>
</tr>
<tr>
<td></td>
<td>(0.0017)</td>
<td>(0.00117)</td>
<td>(0.00118)</td>
</tr>
<tr>
<td>Literate</td>
<td>0.0272***</td>
<td>0.0276***</td>
<td>0.0271***</td>
</tr>
<tr>
<td></td>
<td>(0.0100)</td>
<td>(0.0100)</td>
<td>(0.0100)</td>
</tr>
<tr>
<td>Own Land</td>
<td>-0.0469***</td>
<td>-0.0467***</td>
<td>-0.0474***</td>
</tr>
<tr>
<td></td>
<td>(0.00812)</td>
<td>(0.00816)</td>
<td>(0.00816)</td>
</tr>
<tr>
<td>UpperCaste</td>
<td>-0.0117</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0173)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender HOH</td>
<td>-0.0226**</td>
<td>-0.0228**</td>
<td>-0.0219**</td>
</tr>
<tr>
<td></td>
<td>(0.00995)</td>
<td>(0.00994)</td>
<td>(0.00994)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.436***</td>
<td>0.441***</td>
<td>0.446***</td>
</tr>
<tr>
<td></td>
<td>(0.0429)</td>
<td>(0.0427)</td>
<td>(0.0428)</td>
</tr>
</tbody>
</table>

| Joint Significance          | No                | No                | Yes               |
|                            |                   |                   |                   |
| (1.63)                     | (1.65)            | (3.72)            |                   |
| Observations               | 25.741            | 25.741            | 25.723            |
| R-squared                  | 0.193             | 0.193             | 0.193             |

1. All columns include state and district dummies, as well as clustering SE at the village level
<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>OLS (1)</th>
<th>OLS (2)</th>
<th>OLS (3)</th>
<th>Probit (4)</th>
<th>Probit (5)</th>
<th>Probit (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KnowOfficial</td>
<td>-0.0168</td>
<td>-0.0448</td>
<td>-0.0448</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0117)</td>
<td>(0.0386)</td>
<td>(0.0386)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KnowOfficialXPoor</td>
<td>0.0534*</td>
<td>0.153***</td>
<td>0.153***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0273)</td>
<td>(0.0766)</td>
<td>(0.0766)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KnowGovtEmployee</td>
<td>-0.0928***</td>
<td>-0.296***</td>
<td>-0.296***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00834)</td>
<td>(0.0270)</td>
<td>(0.0270)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KnowGovtEmployeeXPoor</td>
<td>0.0685***</td>
<td>0.240***</td>
<td>0.240***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0231)</td>
<td>(0.0869)</td>
<td>(0.0869)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CasteAssociation</td>
<td>-0.00573</td>
<td>-0.0118</td>
<td>-0.0118</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0128)</td>
<td>(0.0368)</td>
<td>(0.0368)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CasteAssociationXPoor</td>
<td>0.109***</td>
<td>0.290***</td>
<td>0.290***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0288)</td>
<td>(0.0804)</td>
<td>(0.0804)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>0.147***</td>
<td>0.129***</td>
<td>0.129***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00976)</td>
<td>(0.0101)</td>
<td>(0.0101)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.350***</td>
<td>0.352***</td>
<td>0.352***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0370)</td>
<td>(0.0386)</td>
<td>(0.0386)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>26,488</td>
<td>26,149</td>
<td>26,567</td>
<td>26,428</td>
<td>26,089</td>
<td>26,507</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.160</td>
<td>0.167</td>
<td>0.160</td>
<td>0.127</td>
<td>0.133</td>
<td>0.127</td>
</tr>
</tbody>
</table>

1. All columns include state and district dummies as well as clustering SE at the village level.
2. All columns include controls from Tables 2 and 3.
3. R Squared for Probit Regressions is the Pseudo R Squared.
4. Robust standard errors in parentheses

---

### Table 2: Results of Group Connections

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>OLS (1)</th>
<th>OLS (2)</th>
<th>OLS (3)</th>
<th>Probit (4)</th>
<th>Probit (5)</th>
<th>Probit (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KnowOfficial</td>
<td>-0.0168</td>
<td>-0.0448</td>
<td>-0.0448</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0117)</td>
<td>(0.0386)</td>
<td>(0.0386)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KnowOfficialXPoor</td>
<td>0.0534*</td>
<td>0.153***</td>
<td>0.153***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0273)</td>
<td>(0.0766)</td>
<td>(0.0766)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KnowGovtEmployee</td>
<td>-0.0928***</td>
<td>-0.296***</td>
<td>-0.296***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00834)</td>
<td>(0.0270)</td>
<td>(0.0270)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KnowGovtEmployeeXPoor</td>
<td>0.0685***</td>
<td>0.240***</td>
<td>0.240***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0231)</td>
<td>(0.0869)</td>
<td>(0.0869)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CasteAssociation</td>
<td>-0.00573</td>
<td>-0.0118</td>
<td>-0.0118</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0128)</td>
<td>(0.0368)</td>
<td>(0.0368)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CasteAssociationXPoor</td>
<td>0.109***</td>
<td>0.290***</td>
<td>0.290***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0288)</td>
<td>(0.0804)</td>
<td>(0.0804)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>0.147***</td>
<td>0.129***</td>
<td>0.129***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00976)</td>
<td>(0.0101)</td>
<td>(0.0101)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.350***</td>
<td>0.352***</td>
<td>0.352***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0370)</td>
<td>(0.0386)</td>
<td>(0.0386)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>26,488</td>
<td>26,149</td>
<td>26,567</td>
<td>26,428</td>
<td>26,089</td>
<td>26,507</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.160</td>
<td>0.167</td>
<td>0.160</td>
<td>0.127</td>
<td>0.133</td>
<td>0.127</td>
</tr>
</tbody>
</table>

1. All columns include state and district dummies as well as clustering SE at the village level.
2. All columns include controls from Tables 2 and 3.
3. R Squared for Probit Regressions is the Pseudo R Squared.
4. Robust standard errors in parentheses

---
References

Department of Planning, Govt. of Punjab. Survey for BPL Families.
Government of Uttarakhand, Department of Food and Civil Supplies. Services, PDS Services. Issuance of a New Ration Card.

7Potential future source

References

Department of Planning, Govt. of Punjab. Survey for BPL Families.
Government of Uttarakhand, Department of Food and Civil Supplies. Services, PDS Services. Issuance of a New Ration Card.

7Potential future source


Disruptive innovations begin in a niche but soon mature to compete with established technologies in mainstream markets, thereby changing industry boundaries. Theories of disruption fundamentally concern the question of how a new technology competes with an established one. This paper builds on previous research by considering a Stackelberg model of quantity competition with both horizontal and vertical differentiation. While supporting many results of other models considering Cournot competition, I demonstrate that the impact of market segment size and number of firms using each technology depends on which type of firm is the Stackelberg leader, and whether the Stackelberg leader has a first mover advantage. This, in turn, depends on the relative value that technology creates. A firm may have an advantage as a second mover when it can create more value than the first in the mainstream market. I show how changes in parameter values impact an industry’s susceptibility to disruption to try to understand why some markets are more prone to disruptive innovation, why some companies successfully beat back disruption and why some technologies never break out of specialized niche markets.

1 Introduction

Our ability to innovate is growing exponentially. New technologies appear at an alarming rate, yet surprisingly few technologies break into the mainstream market. Instead, many innovations – especially “hi-tech” innovations – only serve the needs of a small, niche group of consumers. Given technology’s ability to significantly change consumer behavior and facilitate the sort of significant revolutions discussed by authors such as Khun (1962), the study of technological adoption is extremely important. The struggle of selling
new technologies and innovations to mainstream consumers has received significant at-
tention from business and marketing strategists, such as Moore (1991). Moore was
amongst the first authors to formulate a theory about how to market highly technologi-
cal products and address the gap between early adopters and other eventual consumers
of a product. Theories from a variety of different fields about how technologies compete
began to receive more attention thanks to the work of Clayton Christensen (1997), who
coined the term “disruptive innovation.”

Disruptive innovations drastically change the competitive landscape. They force es-
tablished technologies to change; create new regulatory problems; and alter consumer
behavior. In today’s age, “disruption” is a hot topic, especially in the start-up world.
The call has been made to disrupt higher education, public schools, and the healthcare
system. Is such a disruption even possible? The theory of disruption, as Jill Lepore
writes, is a theory about why businesses fail. She questions the supposed predictive
power of disruption and its relevance to industries such as education and journalism.1
Her criticism has merit, as the theory focuses more on the process of innovation than
competition between technologies. Perhaps the application of formal economic theory
can help bridge some of these gaps.

The ability to predict when an innovation is likely to be disruptive – or when an indus-
try is vulnerable to disruption – allows businesses, consumers, regulators, and others
to anticipate and prepare for inevitable shifts. Fundamentally, an innovation’s poten-
tial for disruption is a question about how it may compete with what already exists.
The question is that of consumer choice: when will an average, mainstream consumer
choose to purchase (and use) the new technology instead of what he already knows.
It is this question that I seek to address in this paper. In the end, hopefully we can
establish a framework to examine why disruption in the way Christensen speaks about
it has been so difficult to attain in industries like education, health care, and journalism.

In this paper, I build on the model established by Adner and Zemsky and the pri-
mary research question that they ask. I am interested in the structure of industries and
the attributes of technologies in these industries that facilitate technological disruption.
In short, when is an industry faced with a choice between two technologies likely to
be disrupted? I seek to provide a framework through which industries and technolo-
gies can be compared to identify those that are particularly vulnerable to technological
disruption. As Adner and Zemsky state, while economics literature on product differ-
tentiation can be interpreted as being about technological competition, that literature
assumes away the question of whether or not technologies compete. By assuming that
the market is covered, spatial models take as a given that technologies are always in
direct competition. This, however, is not always the case – especially when the tech-
nologies are demanded by consumers in two distinct market segments. Fundamental to

1Jill Lepore, “The Disruption Machine,” The New Yorker, 06/23/2014

new technologies and innovations to mainstream consumers has received significant at-
tention from business and marketing strategists, such as Moore (1991). Moore was
amongst the first authors to formulate a theory about how to market highly technologi-
cal products and address the gap between early adopters and other eventual consumers
of a product. Theories from a variety of different fields about how technologies compete
began to receive more attention thanks to the work of Clayton Christensen (1997), who
coined the term “disruptive innovation.”

Disruptive innovations drastically change the competitive landscape. They force es-
tablished technologies to change; create new regulatory problems; and alter consumer
behavior. In today’s age, “disruption” is a hot topic, especially in the start-up world.
The call has been made to disrupt higher education, public schools, and the healthcare
system. Is such a disruption even possible? The theory of disruption, as Jill Lepore
writes, is a theory about why businesses fail. She questions the supposed predictive
power of disruption and its relevance to industries such as education and journalism.1
Her criticism has merit, as the theory focuses more on the process of innovation than
competition between technologies. Perhaps the application of formal economic theory
can help bridge some of these gaps.

The ability to predict when an innovation is likely to be disruptive – or when an indus-
try is vulnerable to disruption – allows businesses, consumers, regulators, and others
to anticipate and prepare for inevitable shifts. Fundamentally, an innovation’s poten-
tial for disruption is a question about how it may compete with what already exists.
The question is that of consumer choice: when will an average, mainstream consumer
choose to purchase (and use) the new technology instead of what he already knows.
It is this question that I seek to address in this paper. In the end, hopefully we can
establish a framework to examine why disruption in the way Christensen speaks about
it has been so difficult to attain in industries like education, health care, and journalism.

In this paper, I build on the model established by Adner and Zemsky and the pri-
mary research question that they ask. I am interested in the structure of industries and
the attributes of technologies in these industries that facilitate technological disruption.
In short, when is an industry faced with a choice between two technologies likely to
be disrupted? I seek to provide a framework through which industries and technolo-
gies can be compared to identify those that are particularly vulnerable to technological
disruption. As Adner and Zemsky state, while economics literature on product differ-
tentiation can be interpreted as being about technological competition, that literature
assumes away the question of whether or not technologies compete. By assuming that
the market is covered, spatial models take as a given that technologies are always in
direct competition. This, however, is not always the case – especially when the tech-
nologies are demanded by consumers in two distinct market segments. Fundamental to

1Jill Lepore, “The Disruption Machine,” The New Yorker, 06/23/2014
the dilemma of how new technologies enter the mainstream market is the question of when does the new technology actually compete with an established technology. I use many aspects of Adner and Zemsky’s model of horizontal and vertical differentiation. Consumers belong to one of two market segments: a mainstream segment and a niche market segment. Consumers evaluate products differently between segments, and the industry is served by two product technologies: a new technology and an established technology.

In considering whether there is competition between two technologies, I show that many of Adner and Zemsky’s results from Cournot competition hold: the threat of disruption is increasing in the mainstream market segment’s utility from the new technology and the marginal costs of the established firms. The threat of disruption is decreasing in the niche market segment’s utility from the new technology, mainstream consumer’s utility from the established technology, the marginal cost of the new technology, and the size of the niche market segment. However, the effect of the number of new technology firms, the number of established technology firms, and the size of the mainstream segment on the threat of disruption are not the same. This effect depends on who moves first and whether the first mover has an advantage. The first mover has an advantage only if it can create more value relative to firms producing the other type of technology. In many cases the threat of disruption is increasing in the number of new technology firms and the size of the mainstream market segment while decreasing in the number of established technology firms. However, if established technology firms have an advantage as first movers it is possible that the threat of disruption is decreasing in the number of new technology firms and the size of the mainstream segment. Conversely, if new technology firms move first and have an advantage, it is possible that the threat of disruption is actually increasing in the number of established technology firms. If the second mover creates more relative value in the mainstream segment than the first, the second mover has an advantage. When new technology firms have an advantage as second movers, it is possible for the threat of disruption to be decreasing in the number of established technology firms. Similarly, when established technology firms have an advantage as second movers, it is possible for the threat of disruption to be decreasing in the number of new technology firms.

I seek to contribute to literature on market definition. As technologies compete, market boundaries may shift, yielding results relevant to evaluating mergers and acquisitions and the details of many anti-trust cases. The typical approach focuses on the substitutability of products – an approach that is unable to account for the eventual emergence of future competition (Teece and Coleman, 1998). My approach accounts for shifts in competitive boundaries and how these shifts may impact competitive outcomes. By identifying the drivers of disruption, I provide a framework through which we can evaluate industries faced by a potentially disruptive technology. This framework allows us to compare technological attributes and details about market structure to previous examples. Through such a process, we can better predict whether or not
the industry in question will be particularly susceptible to technological disruption. Furthermore, the concept of value creation acts as a powerful tool. By refining the concept of value creation through the addition of industry relevant components, we can analyze “unconventional” industries, such as education, healthcare, and journalism.

2 Model

This paper builds on the model established by Adner and Zemsky (2005). Let there be a new technology and an established technology denoted by $\lambda = N,E$. There are $n_\lambda$ firms using each technology. The cost of production for firm $i$ using technology $\lambda$ is $c_\lambda_i = \frac{1}{n_\lambda} \sum_{i=1}^{n_\lambda} c_i$. Let us assume that the cost of production for firms using the same type of technology is the same. That is, $c_\lambda_i = c_\lambda$ for all $i$ firms.

Let there be two discrete market segments given by $m = 1, 2$. The two market segments represent the mainstream and the niche market segments respectively. Consumers in each market segment have a willingness to pay for the new technology of $\mu_{Nm} + d$ and a willingness to pay for the established technology of $\mu_E + d$. Consumers in a given market segment all share the same value of $\mu_{Nm}$. The ordering of $\mu_{Nm}$ and $\mu_E$ captures vertical differences within a segment and illustrates product differentiation. That is, the difference between $\mu_{Nm}$ and $\mu_E$ illustrates how consumers in a given market segment value the mix of attributes offered by one technology relative to the other. Horizontal differentiation (how consumers in different market segments value the same product) is reflected in the ordering of $\mu_{12}$ and $\mu_{21}$. The established technology is only demanded by consumers in the mainstream market, implying that $\mu_{21} \geq c_E$ while $\mu_{12} \leq 0$ and $\mu_{E2} \leq c_E$ since consumers in the niche market segment get no utility from consuming the established technology. The new technology can be demanded by consumers in both segments: that is $\mu_{11} > c_N$ and $\nu_{12} > 0$. The second component of willingness to pay, parameter $d$, is independent of the product, varies across consumers within a segment and can be assumed to be uniformly distributed between $-\infty$ and $D$ with a density of $S_m > 0$. We assume that $D = 0$. This parameter differentiates consumers in each market segment, allows for some consumer surplus, and lets $S_m$ to parameterize the size of the market segment. Additionally, $|d| \leq \nu_{Nm}$.4

I consider Stackelberg competition. The output of the $r$th firm using technology $\lambda$
is \( q_u \) where the total output of all firms using one technology is \( Q_u = \sum_{i=1}^{n} q_u \). Since the marginal cost of production for firms using the same type of technology is uniform, \( Q_u = n q_u \). Firm profits are then given by \( \pi_{ni} = q_n (P_\lambda(Q_u, Q_u) - c_n) \) and \( \pi_{Ei} = q_E (P_\lambda(Q_u, Q_u) - c_E) \). The price is determined by the inverse demand functions, ensuring that each product is sold at a single price. I am interested in sub-game perfect Nash equilibria (SPNE).

There are two scenarios of interest: when the new technology is sold only in the niche segment; and when the new technology is sold in both segments, competing with the established technology for consumers in the mainstream market segment. Distinguishing between the two scenarios leads us to the following definition:

**Definition 1:** The new technology is considered disruptive if when producing at their profit maximizing quantity considering the possible actions of other firms, the new technology will be demanded by consumers in both the mainstream and niche market segments. The new technology remains isolated if demand for the new technology when firms produce at their profit maximizing quantity is confined to the niche market segment.

### 3 The Drivers of Disruption

This paper defines the threat of disruption differently than Adner and Zemsky do. The goal of this paper is to establish a framework to compare industries facing potential disruption by a new technology with other industries in similar positions, as well as with industries that have already been disrupted. To do so, when discussing the “threat of disruption,” I aim to identify how different parameter values might indicate that a specific industry, faced by a specific set of technologies may be more (or less) vulnerable than another. The vulnerability of an industry to disruption is defined by the quantity at which disruption will be triggered (or the quantity at which the isolated scenario could be forced). The lower the quantity at which disruption can be triggered, the more vulnerable the industry is. Likewise, the higher the quantity that must be produced to to force the isolated scenario, the more vulnerable the industry is. When considered the equilibrium \( \phi = F \) and the drivers of disruption, I choose to focus on how \( Q_u \) changes instead of how \( \pi_{Ei} \) changes because changes in the profit of the firm will depend on how costly it is for that firm to produce more. With this goal in mind, I propose the following definition regarding the threat of disruption: Let \( \frac{Q_u}{S_n} = \beta \) and \( \frac{Q_u}{S_n} - (w_{N1} - w_{N2}) = \theta \).

**Definition 2:** The threat of disruption is increasing in some parameter \( x \) if an increase in that parameter increases the value of \( \beta \) relative to \( \theta \) in both the isolated and disruptive equilibrium, and/or if an increase in that parameter decreases \( Q_u \). The threat of disruption is decreasing in some parameter \( x \) if an increase in that parameter
decreases the value of $\beta$ relative to $\theta$ in both the disruptive and isolated equilibrium, and/or if an increase in that parameter decreases $Q_E^F$.

Figure 1: The event that changing some parameter $x$ increases the value of $\beta$ relative to $\theta$. The location of the kink shifts to the right, changing the outcome of the game from an isolated scenario to a disruptive one.

Figure 3 demonstrates graphically what it means for the threat of disruption to be increasing. Notice that even if equilibrium outputs did not shift, the location of the kink in the demand curve moved to the right, illustrating that the maximum price that consumers in the mainstream segment will demand the new technology at has increased. Intuitively, then the threat of disruption increasing means that if a parameter increases, increasing that parameter makes disruption more attractive to new technology firms as profits from the disruptive equilibrium will increase. Defining the threat of disruption in this fashion allows us to establish the fundamental mechanics of this model.

Proposition 4: The threat of disruption is increasing in $u_{N1}$ and both $c_E$, $c_N$. It is decreasing in $u_{E1}$, $u_{N2}$, $S_x$, and both $c_{N1}$, $c_{N2}$.

These results are in accordance with Adner and Zemsky’s results, and follow the same intuition. In the case where the threat of disruption is increasing, it is because a change in that parameter means that the price need not fall as much for disruption to occur, implying that following the disruptive strategy will be relatively more profitable. The effect and intuition of changing cost is trivial. The higher $u_{N1}$ is, the more consumers in the first segment will be willing to pay for the new technology. However, the higher $u_{E1}$, the more that the price of the new technology will have to fall in order for consumers in the mainstream segment to demand it. The higher $u_{N2}$ is, the more attractive it is to remain in the niche segment. The same logic applies to $S_x$.

$S_x$ is slightly different than the other parameters in this proposition because it is possible for the threat of disruption to be increasing in $S_x$. While increases in $S_x$ decrease the quantity required for established technology firms to force the isolated scenario, when new technology firms create significantly more value than established technology firms, increases in $S_x$ will decrease the feasibility of the isolated equilibrium. This will occur when the new technology is able to take advantage of demand in the niche market segment to further reduce the price of its already superior product, thereby making consumers in the mainstream market segment more likely to buy the new technology. The ease with which new technology firms can do this depends on the number of firms decreases the value of $\beta$ relative to $\theta$ in both the disruptive and isolated equilibrium, and/or if an increase in that parameter decreases $Q_E^F$.

These results are in accordance with Adner and Zemsky’s results, and follow the same intuition. In the case where the threat of disruption is increasing, it is because a change in that parameter means that the price need not fall as much for disruption to occur, implying that following the disruptive strategy will be relatively more profitable. The effect and intuition of changing cost is trivial. The higher $u_{N1}$ is, the more attractive it is to remain in the niche segment. The same logic applies to $S_x$.

$S_x$ is slightly different than the other parameters in this proposition because it is possible for the threat of disruption to be increasing in $S_x$. While increases in $S_x$ decrease the quantity required for established technology firms to force the isolated scenario, when new technology firms create significantly more value than established technology firms, increases in $S_x$ will decrease the feasibility of the isolated equilibrium. This will occur when the new technology is able to take advantage of demand in the niche market segment to further reduce the price of its already superior product, thereby making consumers in the mainstream market segment more likely to buy the new technology. The ease with which new technology firms can do this depends on the number of firms.
using each type of technology, how much more relative value new technology firms can create, and the size of the respective market segments.

Figure 2: If new technology costs are lower than established technology costs, the new technology can drive prices down lower than the established technology can.

Proposition 5: The threat of disruption is decreasing in \( n_E \) if established technology firms create more relative value in the mainstream segment than new technology firms. However, if new technology firms create more relative value in the mainstream segment, it is possible for the threat of disruption to be increasing in \( n_E \).

The fact that the threat of disruption will be decreasing in \( n_E \) reflects established firms advantage as first movers, and follows the same logic as in Adner and Zemsky. With more established technology firms, total quantities of production will increase. Higher values of \( Q_N \) will be required to trigger disruption, and the threat of disruption is decreasing. However, when new technology firms can create far more value than established technology firms, established technology firms are actually at a disadvantage as first movers. Exactly how much more value new technology firms must create depends on the number of new technology firms and new technology firms commitment to the mainstream segment, \( S_1 + S_2 (u_{N1} - u_{N2}) \). The more new technology firms there are, the less the difference in relative value has to be. Similarly, the more committed to the mainstream segment new technology firms are, the less relative value they have to create for the threat of disruption to be decreasing in \( n_E \).

Despite moving second, when new technology firms offer a superior product (in the eyes of both mainstream and niche segment consumers), they can exploit their advantage. After observing what quantity the established technology firms produce at, new technology firms can undercut the price. Consider figure 4. In this example, the new technology creates more relative value because it faces lower costs. As more firms enter, the established technology firms can only lower their prices so much before it would no longer be profitable for them to be in the industry. New technology firms, however, can lower their price far further, thereby facilitating the disruptive outcome. After observing what the established technology firms play, and knowing the limitations imposed on established technology firms by their own cost structure, new technology firms can produce a quantity corresponding to such a low price that they can achieve the desired disruptive outcome. When new technology firms are more committed to the mainstream market, increases in established technology output are not able to drive new technology production back into the niche market as easily.
Proposition 6: The threat of disruption is increasing in $n_N$ if new technology firms create more value relative to established technology firms in the mainstream segment. If established technology firms create more relative value, the threat of disruption depends on the size of the difference between the absolute value that new technology firms can create across both market segments and the relative value that they can create and how large the mainstream market segment is compared to the size of the industry.

If new technology firms create more value than established technology firms in the mainstream segment, they can still rely on demand in the niche segment to make up for their relative inferiority in the mainstream segment. The extent to which new technology firms must rely on demand in the niche segment to compensate for their deficiency depends on the size of the mainstream market segment relative to the size of the industry as a whole. The larger the mainstream market segment is compared to the size of the industry, the easier it will be for the threat of disruption to be increasing in $n_N$. The smaller the mainstream market segment is – or, perhaps more logically, the larger the niche market segment is as a part of the whole – the easier it will be for established technologies to push new technology production back into the niche market, thereby making disruption more difficult. The possibility that the threat of disruption is decreasing in $n_N$ reflects the advantage that established technology firms have as first movers provided that their technology is better than the new technology. My analysis about how $n_N$ affects the threat of disruption elucidates an interesting fact: that established technology firms may only have an advantage as first movers provided they can create more value in the mainstream market segment. We will continue to return to this fact as we explore the last two remaining parameters.

Proposition 7: The threat of disruption is increasing in $S_1$ unless established technology firms create significantly more value than new technology firms in the mainstream segment, and $u_{N2}$ is very large compared to $u_{N1}$.

If new technology firms create more value than established technology firms in the mainstream segment, then, as discussed, new technology firms effectively have a second mover advantage. However, in the other case, where established technology firms create more value in the mainstream segment, things grow more complicated. Of importance is the difference between the absolute value that new technology firms can create and the relative value that they create. Let us create a constant $\alpha$, where $\alpha = (u_{N1} - u_{S1})/(u_{N1} - c_{N1})$. Thus, the size of $\alpha$ tells us about the difference between absolute value created and relative value created. If $\alpha \geq 1$, $\partial \beta / \partial n_N \leq 0$. Under these conditions, the threat of disruption will be decreasing in $n_N$. If $\alpha < 1$, then $\beta > 0$, and we are in scenario 8. The threat of disruption will be increasing if $\beta > 0$ and decreasing otherwise.

If new technology firms create more value than established technology firms in the mainstream segment, things grow more complicated. Of importance is the difference between the absolute value that new technology firms can create and the relative value that they create. Let us create a constant $\alpha$, where $\alpha = (u_{N1} - u_{S1})/(u_{N1} - c_{N1})$. Thus, the size of $\alpha$ tells us about the difference between absolute value created and relative value created. If $\alpha \geq 1$, $\partial \beta / \partial n_N \leq 0$. Under these conditions, the threat of disruption will be decreasing in $n_N$. If $\alpha < 1$, then $\beta > 0$, and we are in scenario 8. The threat of disruption will be increasing if $\beta > 0$ and decreasing otherwise.

If new technology firms are disadvantaged compared to their established counterparts in the mainstream segment, they can still rely on demand in the niche segment to make up for their relative inferiority in the mainstream segment. The extent to which new technology firms must rely on demand in the niche segment to compensate for their deficiency depends on the size of the mainstream market segment relative to the size of the industry as a whole. The larger the mainstream market segment is compared to the size of the industry, the easier it will be for the threat of disruption to be increasing in $n_N$. The smaller the mainstream market segment is – or, perhaps more logically, the larger the niche market segment is as a part of the whole – the easier it will be for established technologies to push new technology production back into the niche market, thereby making disruption more difficult. The possibility that the threat of disruption is decreasing in $n_N$ reflects the advantage that established technology firms have as first movers provided that their technology is better than the new technology. My analysis about how $n_N$ affects the threat of disruption elucidates an interesting fact: that established technology firms may only have an advantage as first movers provided they can create more value in the mainstream market segment. We will continue to return to this fact as we explore the last two remaining parameters.

Proposition 7: The threat of disruption is increasing in $S_1$ unless established technology firms can create significantly more value than new technology firms in the mainstream segment, and $u_{N2}$ is very large compared to $u_{N1}$.
In most cases, the bigger the size of the mainstream segment, the new technology firms can increase output through disruption. Thus, the greater the incentive for disruption. However, in the event that established technology firms can create significantly more value than new technology firms in the mainstream segment, as defined earlier, new technology firms will be disadvantaged as second movers. Since the threat of disruption is decreasing in $n_E$ from proposition 1, it follows that for large values of $n_E$, it is possible that there exists circumstances under which these firms are so disadvantaged in the mainstream segment as second movers that it is more profitable to remain isolated even if the mainstream segment is very large.

When considering propositions 4 through 7, it is important to note how they can be used to examine the potential for technological disruption between different industries. There are a plethora of technologies that have proved disruptive in their respective industries. Thus, we have a baseline to which we can potentially relate other industries in which a new technology is emerging to determine whether or not that technology will be disruptive. By considering aspects such as firm value creation, the size of the different market segments, and the number of firms using each type of technology in our baseline case, we can then examine the attributes of the industry we are interested in and use the guidelines outlined by propositions 4 through 7 to determine whether or not disruption will be more, or less, likely in the industry we seek to analyze.

However, as with Adner and Zemsky, it is important to note that I assume all parameters are exogenous. If some parameters affect the value of other parameters, such as a connection between segment size and number of firms, then my results on the drivers of disruption may change. This offers a possibility for future work.

**First Mover Advantage**

For the most part, my conclusions about the drivers of disruption (even despite the different definition that I use), are in accordance with Adner and Zemsky’s. However, I find different results for $n_S$, $n_E$, and $S_1$. Because of the Stackelberg nature of the game, there exists situations in which the threat of disruption is decreasing in $S_1$, decreasing in $n_S$, and increasing in $n_E$. The differences in results stem from the fact that established technology firms move first and whether or not established technology firms have a first mover advantage – an advantage that established technology firms can exploit to deter the disruptive outcome – or if the new technology firms have an advantage as second movers. If we were to allow new technology firms to move first it is easy to see that the threat of disruption would be increasing in $n_S$ unless established technology firms create significantly more value than new technology firms. In that scenario, it is possible for the threat of disruption to be decreasing in $n_S$. As long as

In most cases, the bigger the size of the mainstream segment, the new technology firms can increase output through disruption. Thus, the greater the incentive for disruption. However, in the event that established technology firms can create significantly more value than new technology firms in the mainstream segment, as defined earlier, new technology firms will be disadvantaged as second movers. Since the threat of disruption is decreasing in $n_E$ from proposition 1, it follows that for large values of $n_E$ it is possible that there exists circumstances under which these firms are so disadvantaged in the mainstream segment as second movers that it is more profitable to remain isolated even if the mainstream segment is very large.

When considering propositions 4 through 7, it is important to note how they can be used to examine the potential for technological disruption between different industries. There are a plethora of technologies that have proved disruptive in their respective industries. Thus, we have a baseline to which we can potentially relate other industries in which a new technology is emerging to determine whether or not that technology will be disruptive. By considering aspects such as firm value creation, the size of the different market segments, and the number of firms using each type of technology in our baseline case, we can then examine the attributes of the industry we are interested in and use the guidelines outlined by propositions 4 through 7 to determine whether or not disruption will be more, or less, likely in the industry we seek to analyze.

However, as with Adner and Zemsky, it is important to note that I assume all parameters are exogenous. If some parameters affect the value of other parameters, such as a connection between segment size and number of firms, then my results on the drivers of disruption may change. This offers a possibility for future work.

**First Mover Advantage**

For the most part, my conclusions about the drivers of disruption (even despite the different definition that I use), are in accordance with Adner and Zemsky’s. However, I find different results for $n_S$, $n_E$, and $S_1$. Because of the Stackelberg nature of the game, there exists situations in which the threat of disruption is decreasing in $S_1$, decreasing in $n_S$, and increasing in $n_E$. The differences in results stem from the fact that established technology firms move first and whether or not established technology firms have a first mover advantage – an advantage that established technology firms can exploit to deter the disruptive outcome – or if the new technology firms have an advantage as second movers. If we were to allow new technology firms to move first it is easy to see that the threat of disruption would be increasing in $n_S$ unless established technology firms create significantly more value than new technology firms. In that scenario, it is possible for the threat of disruption to be decreasing in $n_S$. As long as
new technology firms create more relative value than established technology firms, the threat of disruption would be increasing in $nE$. While this paper will not go into the details of new technology firms as first movers, we can consider what outputs in the disruptive equilibrium would look like, and what output new technology firms would have to play in order to force the disruptive equilibrium.

Lemma 5: i) Allowing new technology firms to act as first movers, new and established technology outputs in the disruptive equilibrium are given by:

$$Q_N^D = (S_1 + S_2)n_N\left(\frac{S_1 + S_2}{S_1 + S_2} - c_E + n_E\left[(u_{N1} - c_N) - (u_{E1} - c_E)\right]\right)$$

$$Q_E^D = (S_1 + S_2)n_E\left(\frac{u_{E1} - c_E - \frac{S_1(u_{N1} - u_{N2})}{S_1 + S_2} - n_N(1 + n_E)\left[(u_{N1} - c_N) - (u_{E1} - c_E)\right]}{(1 + n_N)(1 + n_E)}\right)$$

ii) Allowing new technology firms to act as first movers, the output required to force the disruptive outcome is given by:

$$Q_N^E = (S_1 + S_2)n_N\left(u_{E1} - c_E - \frac{S_2(n_{N1} - u_{N2})}{S_1 + S_2} - \frac{S_2}{S_1 + S_2}(u_{E1} - c_E)\right)$$

As can be seen, the ordering of the relative value created term, namely the ordering of $n_{N1} - c_N$ and $u_{E1} - c_E$ captures the “advantage” a firm has as a first mover. When established technology firms move first, the relative value term manifests as $[(u_{N1} - c_N) - (u_{E1} - c_E)]$. However, when new technology firms move first, this term manifests as $[(u_{N1} - c_N) - (u_{E1} - c_E)]$. Thus, one might draw the conclusion that the game, in a sense of the phrase, makes an assumption that the first mover creates more value than the second mover. And when this assumption is violated, it is possible for the second mover to have an advantage. As discussed, this is because the second mover may be able to exploit the relative weakness of the first movers product in the mainstream market segment.

Thus, while being able to move first is a powerful asset, in cases where the first mover creates less value than the second mover in the mainstream segment, moving first can actually be a disadvantage.

4 Conclusion

The importance of this paper is closely connected to the importance of Adner and Zemsky’s work. By deviating from the standard approach in industrial organization of new technology firms create more relative value than established technology firms, the threat of disruption would be increasing in $nE$. While this paper will not go into the details of new technology firms as first movers, we can consider what outputs in the disruptive equilibrium would look like, and what output new technology firms would have to play in order to force the disruptive equilibrium.

Lemma 5: i) Allowing new technology firms to act as first movers, new and established technology outputs in the disruptive equilibrium are given by:

$$Q_N^D = (S_1 + S_2)n_N\left(\frac{S_1 + S_2}{S_1 + S_2} - c_N + n_E\left[(u_{N1} - c_N) - (u_{E1} - c_E)\right]\right)$$

$$Q_E^D = (S_1 + S_2)n_E\left(\frac{u_{E1} - c_E - \frac{S_1(u_{N1} - u_{N2})}{S_1 + S_2} - n_N(1 + n_E)\left[(u_{N1} - c_N) - (u_{E1} - c_E)\right]}{(1 + n_N)(1 + n_E)}\right)$$

ii) Allowing new technology firms to act as first movers, the output required to force the disruptive outcome is given by:

$$Q_N^E = (S_1 + S_2)n_N\left(u_{E1} - c_E - \frac{S_2(n_{N1} - u_{N2})}{S_1 + S_2} - \frac{S_2}{S_1 + S_2}(u_{E1} - c_E)\right)$$

As can be seen, the ordering of the relative value created term, namely the ordering of $n_{N1} - c_N$ and $u_{E1} - c_E$ captures the “advantage” a firm has as a first mover. When established technology firms move first, the relative value term manifests as $[(u_{N1} - c_N) - (u_{E1} - c_E)]$. However, when new technology firms move first, this term manifests as $[(u_{N1} - c_N) - (u_{E1} - c_E)]$. Thus, one might draw the conclusion that the game, in a sense of the phrase, makes an assumption that the first mover creates more value than the second mover. And when this assumption is violated, it is possible for the second mover to have an advantage. As discussed, this is because the second mover may be able to exploit the relative weakness of the first movers product in the mainstream market segment.

Thus, while being able to move first is a powerful asset, in cases where the first mover creates less value than the second mover in the mainstream segment, moving first can actually be a disadvantage.

4 Conclusion

The importance of this paper is closely connected to the importance of Adner and Zemsky’s work. By deviating from the standard approach in industrial organization of
taking the set of competing firms and relevant consumers as a given, this paper seeks to address key aspects of market definition. As demonstrated by recent empirical observations by authors such as Clayton Christensen, competition between technologies often shifts competitive boundaries. Thus, I expand on Adner and Zemsky’s model of horizontal and vertical differentiation to understand how market boundaries are shaped by the rational behavior of firms, and what industries are particularly vulnerable to disruption. However, rather than assume that all firms make decisions simultaneously, I let established technology firms move first, although it is easy to extend the analysis to situations where new technology firms move first.

My findings are important because they demonstrate that the drivers of disruption change based on whether or not the first mover or the second mover has the advantage. Of crucial importance is the difference between the absolute value that a firm can create and the relative value that it can create. Moving second is not necessarily a disadvantage, especially if the second mover can create more relative value than the first. I show that the drivers of disruption found under Cournot competition may change based on who moves first. The effect of parameters such as number of firms using each technology and the size of the mainstream market segment on the threat of disruption may switch signs depending on who moves first.
5 Bibliography

References

The effect of education on wages has been a widely explored topic. This paper will contribute to the existing literature by studying the causal effect of education on wages in China, a context that has been far less studied. China’s compulsory education laws and minimum age labor laws provide potentially exogenous changes in educational attainment. This paper estimates the private return to education in China (the effect of an extra year of individual educational attainment on wages). Two Chinese institutional education/labor laws provide potential sources of exogenous variation in individual educational attainment. Using these two laws as instruments, we find weakly statistically significant evidence that raising an individual’s educational attainment by one year raises earnings by about 6-9%. This point estimate is relatively consistent with literature from studies done in developed nations, but is significantly different from estimates proposed by other literature done in China.
1 Introduction

It is a certain "fact" about the labor market that education is an important influence on wages. Generally, better educated workers earn more than less educated workers. In a seminal contribution, Mincer (1974) introduces a human capital framework which provided a starting point from which a large literature studying the effect of education on wages spawned. While it is certainly apparent that there is generally a positive correlation between education and wages, one must be careful in interpreting the true effect of education. Education is not randomly assigned across society—indeed, it is certainly plausible that the individuals choosing to achieve more or less education will have other characteristics that can affect their performance in the labor market.

To isolate the causal effect of education requires an exogenous source of variation in educational attainment. Many existing studies have utilized different approaches to address the endogeneity. For example, Ashenfelter and Rouse (1998) among several other studies have employed twins/siblings data to control for family background and/or inherent ability. Other papers, use other exogenous factors to instrument for education. Card (1995) uses proximity to a nearby college to instrument for individual educational attainment.

However, much of the existing literature has focused mostly in the US and Europe. A much smaller subset of the literature has examined the relationship between education and earnings using data from China. At the same time, China is an important country to study this relationship. The Chinese Economic Reform policies in the late 1970’s has paved the way for China’s rapid economic growth. During the same time period, China’s education system was also massively overhauled starting with the implementation of a compulsory education law in 1986.

Our study will contribute to the relatively nascent literature regarding the causal effect of education on wages in China by using institutional features to instrument for individual educational attainment. While using China’s Compulsory Education Law as an instrument for educational attainment is free of reverse causality, individuals born later and thus affected by the education law would have also entered the labor market at a time when China was experiencing massive economic growth. This is a cause of econometric concern because not properly accounting for this effect will overestimate the “true” return to education.

In addition to using China’s Compulsory Education Law, we consider other pos-
sible instruments for educational attainment. One possibility is the enactment of the Provisions on the Prohibition of Using Child Labor in 1991. The Provisions, among many other objectives, establishes a minimum working age of 16 which may have induced greater educational attainment.

We find that OLS estimates imply that an extra year of education is associated with about an 8.4% increase in wages. Furthermore, we find that China’s compulsory schooling law and the Provisions on the Prohibition of Using Child Labor has increased individual educational attainment by 0.82 and 1.15 years, respectively. Using the compulsory schooling law as an instrument, we fail to find statistically significant estimate of the return to education. However, using the Provisions as an instrument, we find that an extra year of education is associated with about a 9% increase in wages. Results differ by sub-populations.

2 Literature Review

From the foundation of the People’s Republic of China in 1949 to the late 1970’s when Deng Xiaoping finally rose to power, the former Ministry of Personnel and Ministry of Labour and Social Security centrally determined the wages of all workers. Empirical research done on the return to education during the Maoist era in China has substantiated claims that these Maoist era policies suppressed returns to education.

Empirical research studying the returns to education during the 1990’s and 2000’s have suggested that China experienced an increasing return to education following reforms. Zhang et. al (2005) use data from fourteen consecutive annual surveys of urban households conducted by China’s National Bureau of Statistics from 1988 through 2001. They estimate that the return to education in urban China was only 4.0% in 1988, but by 2001, it had risen to 10.2% – very similar to estimates in other industrialized countries.

This paper will most closely resemble that of Fang et al. (2012). They estimate the returns to an additional year of schooling in 1997-2006, and they account for the endogeneity problem by using the Compulsory Education Law as an instrument. More specifically, using data from the China Health and Nutrition Survey (CHNS), they first create a binary variable indicating whether an individual was less than 15
years old on the law’s effective date in that individual’s province. Thus, they define their “treatment group” as those who were younger than 15 years old on the law’s effective date and their “control group” as those who were older than 15 years old and thus not affected by the law. They use this binary variable as an instrument for education, and they show that the Compulsory Education Law raised overall individual educational attainment by about 0.8 years.

2.1 Institutional Background

China’s Compulsory Education Law was first established on July 1st, 1986 and is the first Chinese law to stipulate educational policies for the entire country. Among other things, the Law established (China Education Research Network): nine years of education became compulsory, children would begin compulsory education at the age of 6, compulsory education would be free for all students, and employment of children still undergoing compulsory education was outlawed. This means that the typical student who begins compulsory schooling at age 6 would be eligible to leave by age 15. Those under 15 who had already left school by the law’s effective date were required to return to school until they turned 15 (Fang et al 2012).

The other institutional feature which may serve as a potential instrumental variable is The Provisions on the Prohibition of Using Child Labor. First instituted on April 15, 1991, and later amended in 2002, the Provisions establishes guidelines for hiring young adults. Among other objectives, it established the minimum working age to be 16, one year older than the average age of completion of compulsory education, 15. In addition, enforcement of the Provisions was much stricter than enforcement of the Compulsory Education Law (Child Labour Bulletin 2007). At a minimum, employers found to be employing child workers were forced to dismiss the workers at once. Furthermore, employers were fined 5,000 Yuan per child worker per month employed. Repeat offenders faced a much heavier fine and a possibility of getting their business license revoked. Lastly, in 2002, the Provisions were slightly modified to increase the severity of the penalties levied on employers of child laborers. However, the main content remained the same.

2.1 Institutional Background

China’s Compulsory Education Law was first established on July 1st, 1986 and is the first Chinese law to stipulate educational policies for the entire country. Among other things, the Law established (China Education Research Network): nine years of education became compulsory, children would begin compulsory education at the age of 6, compulsory education would be free for all students, and employment of children still undergoing compulsory education was outlawed. This means that the typical student who begins compulsory schooling at age 6 would be eligible to leave by age 15. Those under 15 who had already left school by the law’s effective date were required to return to school until they turned 15 (Fang et al 2012).

The other institutional feature which may serve as a potential instrumental variable is The Provisions on the Prohibition of Using Child Labor. First instituted on April 15, 1991, and later amended in 2002, the Provisions establishes guidelines for hiring young adults. Among other objectives, it established the minimum working age to be 16, one year older than the average age of completion of compulsory education, 15. In addition, enforcement of the Provisions was much stricter than enforcement of the Compulsory Education Law (Child Labour Bulletin 2007). At a minimum, employers found to be employing child workers were forced to dismiss the workers at once. Furthermore, employers were fined 5,000 Yuan per child worker per month employed. Repeat offenders faced a much heavier fine and a possibility of getting their business license revoked. Lastly, in 2002, the Provisions were slightly modified to increase the severity of the penalties levied on employers of child laborers. However, the main content remained the same.
3 Methodology

As mentioned previously, to estimate the private returns to education, we will use China’s Compulsory Education Law and The Provisions on the Prohibition of using Child Labor as instruments for individual educational attainment. Our equation of interest is a standard Mincerian model:

\[
\ln(Wages) = \beta_0 + \beta_1(Schooling) + \beta_2X + \mu, \tag{1}
\]

- \(\ln(Wages)\) is the natural log of annual wages (Renminbi 2006 Yuan)
- \(Schooling\) is years of education completed
- \(X\) represents other controls (e.g., gender, marital status, etc.)
- \(\mu\) is the error term.

Using instrumental variables may solve the potential problem that education is endogenous. A valid instrument should significantly correlated with the main independent variable, but should not be correlated with other variables that affect the main dependent variable. We believe that both the Compulsory Education Law and The Provisions on the Prohibition of using Child Labor should be statistically significantly correlated with education, and we can explicitly test this in a first stage regression.

However, institutional features may potentially be correlated with omitted variables that also affect individual educational attainment and wages. While this has generally not been found to be a problem in industrialized countries (such as US or other countries in Europe), this may be a problem in China. The Compulsory Education Law, first enacted in 1986, and the Provisions on the Prohibition of using Child Labor, first enacted in 1991, were both instituted at a time when China was beginning to experience economic growth. Thus, IV estimates must be interpreted with caution. Without additional care, IV estimates of the impact of individual educational attainment on wages may simply just reflect time trends. As mentioned previously in Section I, this is a cause of concern for the seemingly high estimates provided in Fang et. al (2012). Therefore, this paper will also conduct further identification tests to examine the possibility of potential bias in these estimates.
This study will use longitudinal data from the China Health and Nutrition Survey (CHNS), which is collaboratively maintained by both the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention. The survey is conducted every 2-4 years since 1989.

Consistent with the existing literature we will keep only those individuals born in 1961 or later. As we mentioned earlier, Maoist-era policies potentially suppressed returns to education. In addition, events such as the Great Leap Forward (1958-1961) and the Chinese Cultural Revolution (1966-1976) marked a time of extreme political turmoil in China. Individuals born in these periods may not be appropriate comparisons due to the very different socioeconomic conditions of their childhoods. Individuals born in 1961 or later would have been younger than 15 by the end of the Cultural Revolution. Furthermore, we will exclude students currently studying in school since they do not currently have any earnings.

To estimate the private returns to schooling, the dependent variable of interest is the natural logarithm of an individual’s earnings. This will include: annual wages, income from businesses, faring, gardening, fishing, livestock, and retirement benefits. The main independent variable will be completed years of education. Other controls may include: age, gender, ethnic status, marital status, urban residency, province, and self-reported health status.

5 Results

5.1 OLS estimates of the impact of individual educational attainment on earnings

Table 1 displays OLS estimates of the effect of years of education on individual wages. These estimates imply that an extra year of individual educational attainment is on average associated with an 8.42% increase in individual wages. These OLS estimates seem to be just a little bit higher than OLS estimates presented in other Western studies (Card 1999); however, they are relatively consistent to OLS estimates presented in Fang et. al (2012).
Table 1: OLS Results. Dependent Variable: Natural Logarithm of Individual Wages (2006 Yuan)

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of Education</td>
<td>0.119***</td>
<td>0.118***</td>
<td>0.103***</td>
<td>0.101***</td>
<td>0.0842***</td>
</tr>
<tr>
<td></td>
<td>(0.00363)</td>
<td>(0.00359)</td>
<td>(0.00377)</td>
<td>(0.00378)</td>
<td>(0.00369)</td>
</tr>
<tr>
<td>Age</td>
<td>0.0241***</td>
<td>-0.00725</td>
<td>-0.00712</td>
<td>0.0157</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00174)</td>
<td>(0.0160)</td>
<td>(0.0161)</td>
<td>(0.0164)</td>
<td></td>
</tr>
<tr>
<td>Age(^2)</td>
<td>0.000392</td>
<td>0.000416*</td>
<td>-0.000139</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000247)</td>
<td>(0.000247)</td>
<td>(0.000254)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.143***</td>
<td>0.181***</td>
<td>0.169***</td>
<td>0.177***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0259)</td>
<td>(0.0262)</td>
<td>(0.0262)</td>
<td>(0.0253)</td>
<td></td>
</tr>
<tr>
<td>Race Minority</td>
<td>-0.229***</td>
<td>-0.230***</td>
<td>-0.0266</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0394)</td>
<td>(0.0392)</td>
<td>(0.0449)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.167***</td>
<td>0.159***</td>
<td>-0.00271</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0420)</td>
<td>(0.0419)</td>
<td>(0.0415)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>0.328***</td>
<td>0.343***</td>
<td>0.384***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0290)</td>
<td>(0.0288)</td>
<td>(0.0279)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0373)</td>
<td>(0.0683)</td>
<td>(0.237)</td>
<td>(0.238)</td>
<td>(0.244)</td>
</tr>
<tr>
<td>Observations</td>
<td>10,958</td>
<td>10,958</td>
<td>10,958</td>
<td>10,958</td>
<td>10,958</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.096</td>
<td>0.121</td>
<td>0.140</td>
<td>0.143</td>
<td>0.212</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As mentioned before, OLS estimates above may be potentially plagued by omitted variables bias. Another source of concern is that measurement error could attenuate the “true” effect of education on wages. If the measurement error is nonclassical, the direction of bias is unknown. In the following sections, we present results from instrumental variables regressions which have the potential to eliminate said biases.

5.2 Instrumental Variable estimates for the impact of individual educational attainment on earnings (Compulsory Schooling Law)

Our first approach is very similar to that of Fang et. al (2012). We first proceed by using China’s CSL as an instrument for individual educational attainment. The instrument is a dummy variable which equals 1 if the individual was less than 15 years old.
old on the law’s effective date in the individual’s province and equal to 0 otherwise.

Next, Table 2 presents 2SLS results, estimating the effect of individual educational attainment on earnings using the effect of the CSL as an instrument for individual educational attainment.

Using China’s CSL as an instrument for individual educational attainment, we fail to find overall statistically significant evidence for the effect of education on earnings. As Column 1 of Table 2 suggests, the overall impact of an increase in individual educational attainment by 1 year was, on average, associated with about 5.8% increase in wages. However, this estimate is not significant at any conventional levels of statistical significance.

We acknowledge that column 7 of Table 2 suggests that an increase in individual educational attainment of 1 year in a coastal province is on average associated with 32.6% increase in wages. This estimate should also be viewed with caution as it is much higher than estimates found in the existing literature. Certainly, since the standard errors are quite large, the 95% confidence interval for the estimate on years of education in a Coastal province is (0.026, 0.626); thus, we cannot reject an estimate of a more reasonable estimate such as 11% found for the rural sub-population.

Lastly, it should be noted that these 2SLS results presented in Table 2 are very different to the results presented in Fang et. al (2012). As stated in Section I, Fang et. al (2012) report an IV estimate suggesting an increase in individual educational attainment by one year is, on average, associated with 20% increase in earnings. Furthermore, Fang et. al (2012) report much more statistically significant estimates by each sub-population as well. This stands in contrast to our first stage results which seem to be very consistent with that of Fang et. al (2012)’s. We have not yet reconciled these stark differences in the 2SLS estimates and attempts have been made to reach out to the authors of Fang et. al (2012) to further investigate. The Appendix at the end of this paper provides a side-by-side comparison of this paper’s results to that of Fang et. al (2012)’s. As Table ?? shows, this paper’s first stage results are consistent with that of Fang et. al (2012)’s. Furthermore, OLS estimates of returns to education are also consistent as shown in Table ?? . However, 2SLS estimates are statistically significantly different in aggregate. Results vary by subpopulation.
Table 2: Instrumental Variables (2SLS) Estimations.

Instrument: Compulsory Schooling Law. Dependent Variable: Natural Log of Earnings

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>All Female</th>
<th>Male</th>
<th>Rural</th>
<th>Urban</th>
<th>Inland</th>
<th>Coastal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of Education</td>
<td>0.0585</td>
<td>0.0367</td>
<td>0.0935</td>
<td>0.116</td>
<td>-0.139</td>
<td>-0.326***</td>
</tr>
<tr>
<td>Age</td>
<td>0.0225</td>
<td>0.0156</td>
<td>0.0488</td>
<td>-0.0124</td>
<td>0.134*</td>
<td>0.0347</td>
</tr>
<tr>
<td>Age^2</td>
<td>0.000262</td>
<td>0.000236</td>
<td>0.000638</td>
<td>0.00210</td>
<td>0.000507</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.190***</td>
<td>0.128</td>
<td>0.248**</td>
<td>0.162***</td>
<td>0.152</td>
<td>0.162***</td>
</tr>
<tr>
<td>Married</td>
<td>-0.0166</td>
<td>0.0712</td>
<td>0.0450</td>
<td>-0.277</td>
<td>-0.0617</td>
<td>0.0694***</td>
</tr>
<tr>
<td>Race Minority</td>
<td>-0.0333</td>
<td>-0.0301</td>
<td>-0.0570</td>
<td>0.0548</td>
<td>-0.0489</td>
<td>0.0489</td>
</tr>
<tr>
<td>Urban</td>
<td>0.436***</td>
<td>0.428</td>
<td>0.516***</td>
<td>-0.133</td>
<td>0.125***</td>
<td></td>
</tr>
<tr>
<td>Health Status Dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Provincial Dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>CHINS Wave Dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Constant</td>
<td>6.904***</td>
<td>7.384***</td>
<td>6.287***</td>
<td>6.896***</td>
<td>8.025***</td>
<td>7.381***</td>
</tr>
<tr>
<td>Observations</td>
<td>10.958</td>
<td>5.410</td>
<td>5.548</td>
<td>7.819</td>
<td>3.139</td>
<td>7.749</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Next, we test the affect of the Provisions on individual educational attainment. Table 3 presents first stage results. As stated in Section 2, the Provisions estab-
lished a minimum age labor law of 16 which was enforced more strictly than the compulsory schooling laws. We believe that the Provisions may have increased individual educational attainment because the opportunity cost of staying in school at a younger age was lower after the implementation of the Provisions. Since the Provisions was a national law that was implemented uniformly in each of the province, the instrument is simply defined as a dummy variable equal to 1 if the individual was less than 16 at the time of the implementation of the Provisions, and 0 otherwise.

The effect of The Provisions on individual educational attainment, as shown in Table 3, seem to be greater than the effect of the CSL, as shown in Table ???. This seems reasonable; as mentioned in Section 2.1, the Provisions were enforced much more strictly than the CSL. In aggregate, our estimates imply that being affected by the Provisions, on average, increased individual educational attainment by about 1.1 years. Results vary by sub-population.
Table 3: First Stage Results.
Instrument: The Provisions. Dependent Variable: Completed Years of Schooling

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>All</th>
<th>Female</th>
<th>Male</th>
<th>Rural</th>
<th>Urban</th>
<th>Inland</th>
<th>Coastal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td>Affected by the Provisions</td>
<td>1.150***</td>
<td>1.364***</td>
<td>0.863***</td>
<td>1.179***</td>
<td>1.097***</td>
<td>1.248**</td>
<td>0.962*</td>
</tr>
<tr>
<td></td>
<td>(0.241)</td>
<td>(0.255)</td>
<td>(0.256)</td>
<td>(0.258)</td>
<td>(0.284)</td>
<td>(0.310)</td>
<td>(0.305)</td>
</tr>
<tr>
<td>Age</td>
<td>0.410***</td>
<td>0.459***</td>
<td>0.413***</td>
<td>0.374***</td>
<td>0.382***</td>
<td>0.406***</td>
<td>0.358**</td>
</tr>
<tr>
<td></td>
<td>(0.0389)</td>
<td>(0.0409)</td>
<td>(0.0625)</td>
<td>(0.0547)</td>
<td>(0.0895)</td>
<td>(0.0539)</td>
<td>(0.0595)</td>
</tr>
<tr>
<td>Age^2</td>
<td>0.006***</td>
<td>0.007***</td>
<td>0.006***</td>
<td>0.005***</td>
<td>0.006***</td>
<td>0.006***</td>
<td>-0.006*</td>
</tr>
<tr>
<td></td>
<td>(0.000739)</td>
<td>(0.000896)</td>
<td>(0.00103)</td>
<td>(0.00145)</td>
<td>(0.00168)</td>
<td>(0.001414)</td>
<td></td>
</tr>
<tr>
<td>Race Minority</td>
<td>-0.391</td>
<td>-0.873***</td>
<td>0.191</td>
<td>0.0181</td>
<td>0.565**</td>
<td>0.655</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
<td>(0.173)</td>
<td>(0.181)</td>
<td>(0.183)</td>
<td>(0.338)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.549***</td>
<td>0.765***</td>
<td>0.0181</td>
<td>0.505**</td>
<td>0.655</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
<td>(0.173)</td>
<td>(0.181)</td>
<td>(0.183)</td>
<td>(0.338)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1.995***</td>
<td>2.323***</td>
<td>1.619***</td>
<td>1.818***</td>
<td>2.445***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.231)</td>
<td>(0.283)</td>
<td>(0.201)</td>
<td>(0.289)</td>
<td>(0.196)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Province Dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>CHNS Wave Dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Constant</td>
<td>1.689**</td>
<td>1.501</td>
<td>1.847**</td>
<td>1.891**</td>
<td>4.760***</td>
<td>4.611**</td>
<td>3.190**</td>
</tr>
<tr>
<td></td>
<td>(0.699)</td>
<td>(1.066)</td>
<td>(0.905)</td>
<td>(0.816)</td>
<td>(1.262)</td>
<td>(0.824)</td>
<td>(1.258)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.153</td>
<td>0.189</td>
<td>0.127</td>
<td>0.072</td>
<td>0.154</td>
<td>0.140</td>
<td>0.170</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.00141</td>
<td>0.009684</td>
<td>0.00974</td>
<td>0.00182</td>
<td>0.00483</td>
<td>0.0100</td>
<td>0.0875</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses  
*** p<0.01, ** p<0.05, * p<0.1

Table 3: First Stage Results.  
Instrument: The Provisions. Dependent Variable: Completed Years of Schooling

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>All</th>
<th>Female</th>
<th>Male</th>
<th>Rural</th>
<th>Urban</th>
<th>Inland</th>
<th>Coastal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td>Affected by the Provisions</td>
<td>1.150***</td>
<td>1.364***</td>
<td>0.863***</td>
<td>1.179***</td>
<td>1.097***</td>
<td>1.248**</td>
<td>0.962*</td>
</tr>
<tr>
<td></td>
<td>(0.241)</td>
<td>(0.255)</td>
<td>(0.256)</td>
<td>(0.258)</td>
<td>(0.284)</td>
<td>(0.310)</td>
<td>(0.305)</td>
</tr>
<tr>
<td>Age</td>
<td>0.410***</td>
<td>0.459***</td>
<td>0.413***</td>
<td>0.374***</td>
<td>0.382***</td>
<td>0.406***</td>
<td>0.358**</td>
</tr>
<tr>
<td></td>
<td>(0.0389)</td>
<td>(0.0409)</td>
<td>(0.0625)</td>
<td>(0.0547)</td>
<td>(0.0895)</td>
<td>(0.0539)</td>
<td>(0.0595)</td>
</tr>
<tr>
<td>Age^2</td>
<td>0.006***</td>
<td>0.007***</td>
<td>0.006***</td>
<td>0.005***</td>
<td>0.006***</td>
<td>0.006***</td>
<td>-0.006*</td>
</tr>
<tr>
<td></td>
<td>(0.000739)</td>
<td>(0.000896)</td>
<td>(0.00103)</td>
<td>(0.00145)</td>
<td>(0.00168)</td>
<td>(0.001414)</td>
<td></td>
</tr>
<tr>
<td>Race Minority</td>
<td>-0.391</td>
<td>-0.873***</td>
<td>0.191</td>
<td>0.0181</td>
<td>0.565**</td>
<td>0.655</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
<td>(0.173)</td>
<td>(0.181)</td>
<td>(0.183)</td>
<td>(0.338)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.549***</td>
<td>0.765***</td>
<td>0.0181</td>
<td>0.505**</td>
<td>0.655</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
<td>(0.173)</td>
<td>(0.181)</td>
<td>(0.183)</td>
<td>(0.338)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1.995***</td>
<td>2.323***</td>
<td>1.619***</td>
<td>1.818***</td>
<td>2.445***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.231)</td>
<td>(0.283)</td>
<td>(0.201)</td>
<td>(0.289)</td>
<td>(0.196)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Province Dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>CHNS Wave Dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Constant</td>
<td>1.689**</td>
<td>1.501</td>
<td>1.847**</td>
<td>1.891**</td>
<td>4.760***</td>
<td>4.611**</td>
<td>3.190**</td>
</tr>
<tr>
<td></td>
<td>(0.699)</td>
<td>(1.066)</td>
<td>(0.905)</td>
<td>(0.816)</td>
<td>(1.262)</td>
<td>(0.824)</td>
<td>(1.258)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.153</td>
<td>0.189</td>
<td>0.127</td>
<td>0.072</td>
<td>0.154</td>
<td>0.140</td>
<td>0.170</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.00141</td>
<td>0.009684</td>
<td>0.00974</td>
<td>0.00182</td>
<td>0.00483</td>
<td>0.0100</td>
<td>0.0875</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses  
*** p<0.01, ** p<0.05, * p<0.1
Table 4: Instrumental Variables (2SLS) Estimations.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Errors Clustered by Province</td>
<td>---------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>YEARS OF EDUCATION</td>
<td>0.0898*</td>
<td>0.0477</td>
<td>0.161</td>
<td>0.0317</td>
<td>0.254***</td>
<td>0.0722*</td>
<td>0.159</td>
</tr>
<tr>
<td>(0.0549)</td>
<td>(0.0484)</td>
<td>(0.106)</td>
<td>(0.0642)</td>
<td>(0.0983)</td>
<td>(0.0406)</td>
<td>(0.143)</td>
<td></td>
</tr>
<tr>
<td>AGE</td>
<td>0.0142</td>
<td>0.0123</td>
<td>0.0288</td>
<td>0.0067</td>
<td>0.0451</td>
<td>0.0148</td>
<td>0.0249</td>
</tr>
<tr>
<td>(0.0283)</td>
<td>(0.0243)</td>
<td>(0.0561)</td>
<td>(0.0231)</td>
<td>(0.0585)</td>
<td>(0.0251)</td>
<td>(0.0533)</td>
<td></td>
</tr>
<tr>
<td>AGE$^2$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(0.000113)</td>
<td>(0.000173)</td>
<td>(0.000331)</td>
<td>0.05</td>
<td>0.000340</td>
<td>0.000163</td>
<td>0.000127</td>
<td></td>
</tr>
<tr>
<td>(0.000511)</td>
<td>(0.000434)</td>
<td>(0.000881)</td>
<td>(0.000439)</td>
<td>(0.00103)</td>
<td>(0.000435)</td>
<td>(0.00107)</td>
<td></td>
</tr>
<tr>
<td>MALE</td>
<td>0.174***</td>
<td>0.191**</td>
<td>0.246***</td>
<td>0.125**</td>
<td>0.260***</td>
<td>0.0722*</td>
<td>0.159</td>
</tr>
<tr>
<td>(0.0577)</td>
<td>(0.0830)</td>
<td>(0.0667)</td>
<td>(0.0823)</td>
<td>(0.0446)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MARITAL</td>
<td>0.000271 -</td>
<td>0.117</td>
<td>0.0245</td>
<td>0.102</td>
<td>-0.0614</td>
<td>0.0311</td>
<td></td>
</tr>
<tr>
<td>(0.0590)</td>
<td>(0.0878)</td>
<td>(0.127)</td>
<td>(0.0458)</td>
<td>(0.164)</td>
<td>(0.0699)</td>
<td>(0.0608)</td>
<td></td>
</tr>
<tr>
<td>RACE MINORITY</td>
<td>0.373***</td>
<td>0.414***</td>
<td>0.317***</td>
<td>0.371***</td>
<td>0.274</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.0677)</td>
<td>(0.0802)</td>
<td>(0.0592)</td>
<td>(0.0656)</td>
<td>(0.0572)</td>
<td>(0.049)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>URBAN</td>
<td>0.373***</td>
<td>0.414***</td>
<td>0.317***</td>
<td>0.371***</td>
<td>0.274</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.119)</td>
<td>(0.136)</td>
<td>(0.157)</td>
<td>(0.0774)</td>
<td>(0.395)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEALTH STATUS DUMMIES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>PROVINCIAL DUMMERS</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>CHNS WAVE DUMMERS</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Constant</td>
<td>6.750***</td>
<td>7.328***</td>
<td>5.993***</td>
<td>7.326***</td>
<td>4.861***</td>
<td>6.966***</td>
<td>5.678***</td>
</tr>
<tr>
<td>(0.264)</td>
<td>(0.404)</td>
<td>(0.402)</td>
<td>(0.391)</td>
<td>(0.654)</td>
<td>(0.237)</td>
<td>(0.630)</td>
<td></td>
</tr>
<tr>
<td>OBSERVATIONS</td>
<td>10,958</td>
<td>5,410</td>
<td>5,548</td>
<td>7,819</td>
<td>3,139</td>
<td>7,749</td>
<td>3,209</td>
</tr>
</tbody>
</table>

As Table 4 shows, using The Provisions as an instrument for individual educational attainment yields different results than using China’s CSL as an instrument. As column 1 shows, in aggregate, the 2SLS estimate implies that an extra year of education is, on average, associated with about a 9% increase in wages. In it of itself,
this estimate seems to be very reasonable – it is very similar to estimates produced in Western studies (Card 1999). However, it is only significant at the 10% level.

Examining the estimates in each sub-population, we do not observe similar trends as with the 2SLS estimates given by using the CSL as an instrument. The estimate on the rural sub-population is no longer significant. Furthermore, the estimate on the urban sub-population is significant at the 1% level and suggests that the return on education is almost 25%. At the 10% significance level, we find that the effect of individual educational attainment on earnings is significant with the inland provinces sub-population.

We note that these estimates should be viewed with caution as these regressions do not necessarily account for overall time trends adequately. Robustness checks on these estimates and instruments are still being considered.

5.3 Instrumental Variable estimates of the impact of individual educational attainment on earnings (Quarter of Birth)

In this section, we explore the possibility of using quarter of birth as an instrument for individual educational attainment. Table ?? depicts first stage results of quarter of birth as an instrument.

A priori, we believe that the Provisions may have generated some variation in individual educational attainment by quarter of birth. If the effect of the Provisions is significant, then we expect that students will stay in school until age 16. If this is the case, then the Provisions establish a minimum dropout age similar to the compulsory schooling laws that Angrist and Krueger (1991) examine in the United States. Thus, the regressions shown in Table ?? only includes individuals who are affected by the Provisions. This is because we have no reason to believe that individuals not affected by the Provisions would have any variation in individual educational attainment by quarter of birth because there was no previous minimum dropout age.
6 Conclusion

In this paper, we examine two policies implemented at the vanguard of China’s socioeconomic reformation: the Compulsory Schooling Law of 1986 and the Provisions on the Prohibition of using Child Labor of 1991. We first test whether each of these policies affected individual educational attainment. Second, we use the implementation of these policies as instruments for individual educational attainment to estimate the return to education in China.

Our first finding is that it seems that each of these policies did statistically significantly raise individual educational attainment. In aggregate, China’s compulsory schooling law raised educational attainment by about 0.82 years. Similarly, The Provisions raised educational attainment by over 1 year.

Second, we fail to find statistically significant evidence that years of education affected earnings when using the compulsory schooling law as an instrument. This is in stark contrast to Fang et al (2012) who find that the return to education was about 20% in aggregate. In fact, using the compulsory schooling law as an instrument, we only find an estimate of 5.84% returns to education with a 95% confidence interval of (–7.56%, 19.25%). Thus, we can reject a return of 20% at the 5% level. A side-by-side comparison to Fang et al (2012)’s results can be found in the Appendix.

Third, we do find very marginal statistically significant evidence that years of education affected earnings when using the Provisions as an instrument for individual educational attainment. In aggregate, we find that an extra year of education is associated with about a 9% increase in wages – an estimate that is highly consistent with those presented in other Western studies (Card 1999). Estimates vary considerably across sub-populations. Note that the two estimates obtained using the compulsory schooling law and the Provisions are not statistically different from each other.

Finally, we conclude that estimates presented in this paper thus far must be viewed with caution. We have several causes for concern and we hope to build on this literature in the following manner. We aim to reconcile the stark differences between our results and those of Fang et al (2012)’s. Although it seems that we have used the same data set, created the instrument equivalently, used the same set of controls, achieved similar first stage results, 2SLS results when using the compulsory schooling law as an instrument are extremely different.
This study ultimately aims to provide updated results for the causal impact of individual educational attainment on earnings in modern China. Demonstrating the returns to education in China is important from a policy standpoint not only in China but also in other developing countries. More accurate estimates may even ultimately contribute to the literature of cross-country income disparities and how much are due to differences in human capital.

References


DRIVING STYLE AS A GAME: BEHAVIORAL RESPONSES TO ROAD SAFETY FACTORS

RACHEL SZYMANSKI
Georgetown University

ABSTRACT
Road safety factors influence how carefully people drive. This paper presents a two-driver model that regards the driving decision as a game, where the probability of an accident depends on both drivers’ decisions. The model accounts for accident probability, cost, and heterogeneous benefit to reckless driving. It predicts how various parameters affect the fraction of careful drivers and driver welfare. I then use data from the National Highway Traffic Safety Administration’s Fatality Analysis Reporting System (FARS) to test the model’s predictions on what specific road safety variables increase careful driving. These include speed limits, traffic signs, weather, light conditions, and other road conditions. I find evidence that drivers change their behavior in response to bad weather, dark lighting, ramps, traffic signs, and congestion.

1 Introduction
From drunk driving to inattention, reckless driving causes many accidents on the road each year. Transportation policy makers often assume that reducing accident cost or probability will increase safety and welfare. However, they must consider changes in driver behavior due to such policies. A policy designed to increase safety may induce reckless behavior, decreasing welfare. Similarly, a policy that initially appears to decrease safety may increase careful behavior and increase welfare. Raising speed limits, for example, increases accident probability, but that higher speed limits could also increase careful driving. This develops a model for driver behavior and determines the effect of various traffic variables on driver behavior in individual accidents.

I’ve developed a two-driver model for how people choose how safely to drive. The model will take into account accident cost, accident probability,
heterogeneous driver benefit, and the strategic interaction present among drivers. The intuition is as follows: Everyone would drive recklessly if they knew that they wouldn’t get into an accident. However, there is accident risk and drivers must take this into account. Each driver must decide if the benefit to reckless driving outweighs the expected accident cost, considering others’ driving styles. Further, drivers must take into account that accident risk depends on other drivers’ driving style. If a given driver encounters a careful driver, the probability of an accident will be lower than if the given driver encounters a reckless driver. Since each driver’s driving style decision depends on others’ decisions, choosing a driving style can be thought of as a game. If there are more careful drivers on the road, a given driver has incentive to drive recklessly knowing that other careful drivers will prevent an accident. If there are more reckless drivers, a given driver has incentive to drive carefully to prevent an accident.

The model provides insights on the following two questions: 1. How does changing incentives affect the driving style decision? 2. How does changing incentives and driving style affect welfare? I find increases in accident cost or equal proportion increases in accident probabilities given others’ driving styles increase careful driving. However, when accident probabilities do not change by the same proportion, there may be a decrease in careful driving. The effect of accident probability on careful driving depends on relative changes in accident probabilities given others’ driving styles, not overall accident probability. I also find that, all else equal, an increase in the fraction of careful drivers increases welfare. However, an increase in accident cost or a proportional increase in all accident probabilities has an ambiguous effect on welfare.

**Literature Review**

Peltzman (1975) initiated controversial discussion on how traffic regulations influence driver behavior. He puts forth an idea now known as the compensating behavior theory arguing that car safety devices increase reckless driving. He explains that safety devices essentially make the price of reckless driving less expensive. Assuming that recklessness is a normal good, a decrease in the price of reckless driving will increase reckless driving. Peltzman then empirically analyzes the safety and behavior effects of the Vehicle Safety Act, which mandated certain vehicle safety standards.

Since then, the empirical research testing the connection between seat belts and driving behavior has not consistently supported either theory. The standard analysis is to determine how seat belt laws affect non-occupant accidents and fatalities. Non-occupants are generally defined as bicyclists and pedestrians. If seat belt laws increase non-occupant accidents or fatalities, researchers conclude that seat-belt wearing drivers are driving more recklessly. Asch et al. (1991) and others find that mandatory seat belt laws led to an increase in accidents among non-occupants. Many of such studies, however, have been criticized for lack of sufficient controls. Cohen and Einav (2003) most notably criticized these previous papers, which use time series data and changes in seat belt laws. Cohen and Einav argue that these papers could not account
for simultaneous law changes, various macroeconomic effects, publicity campaigns, and other factors. They conduct a study that controls for the issues they raised. First, they use panel data to control for state and year fixed effects. Second, they recognize that seat belt laws affect seat belt usage and seat belt usage is ultimately the variable that affects fatalities. They argue that seat belt usage is likely endogenous due to various environmental reasons affecting the probability of an accident. To correct for this endogeneity, they use seat belt law intensity as an instrument. They find that seat belt use had no significant effect on fatalities among non-occupants and claim that this provides evidence that seat belt laws and usage do not encourage reckless driving.

One variable I wish to test for an effect on driving behavior is speed limit. The speed limit literature mainly analyzes speed limits’ effects on accidents and fatalities. In 1985, the federal government allowed state governments to relax the speed limit on rural interstate roads from 55 mph to 65 mph. Many papers examine the effect of this speed limit change on road safety. Most studies found evidence of safety reductions. Brown et. al (1990) finds that a speed limit increase increased accidents by 19 percent in Alabama. Wagenaar et. al (1990) find that a speed limit increase increased injuries and fatalities in Michigan. Other studies, however, provide conflicting results. Chang and Paniati (1990) and Sidhu (1990) find little to no effect on traffic safety. McCarthy (1991) even find evidence of safety improvements. In 1995, the federal government allowed states to freely set their own speed limits on interstates.

Model

Driving Style Decision

The model presented in this paper will make predictions about the change in the fraction of careful drivers given various parameter changes and will account for the strategic interaction among drivers in choosing a driving style. Assume a population of drivers heterogeneous only in the benefit they receive from reckless driving. Assume that the benefit to reckless driving is exogenous and is greater than the benefit to careful driving. Note that the benefit to reckless driving is gross benefit, not net benefit, so does not include any accident risk or cost. I normalize the benefit to careful driving to 0. I denote the benefit to reckless driving by \( b \) and assume that \( b \) is distributed over \([0, \tilde{b}]\) according to a probability density function \( f(b) \). Each driver randomly encounters another driver and the two play a symmetric normal form game where each chooses a driving style, \( R \) (Reckless) or \( C \) (Careful). I assume that each driver knows his own benefit, but not the other driver’s benefit, and that both know the benefit density function \( f(b) \).

The expected payoff from choosing strategy \( R \) for a type \( b \) driver matched with an unknown \( b \) driver is given by the following utility function:

\[
U(R, b) = b - A[\Pi(R, R)Pr(R) + \Pi(R, C)Pr(C)]
\]

(1)

where \( \Pi(i, j) \) is the probability of an accident given a driver pair playing \((i, j) \) \( \in \{R, C\} \), \( Pr(j) \) is the probability that the other driver plays strategy \( j \), and \( A \) is the for simultaneous law changes, various macroeconomic effects, publicity campaigns, and other factors. They conduct a study that controls for the issues they raised. First, they use panel data to control for state and year fixed effects. Second, they recognize that seat belt laws affect seat belt usage and seat belt usage is ultimately the variable that affects fatalities. They argue that seat belt usage is likely endogenous due to various environmental reasons affecting the probability of an accident. To correct for this endogeneity, they use seat belt law intensity as an instrument. They find that seat belt use had no significant effect on fatalities among non-occupants and claim that this provides evidence that seat belt laws and usage do not encourage reckless driving.

One variable I wish to test for an effect on driving behavior is speed limit. The speed limit literature mainly analyzes speed limits’ effects on accidents and fatalities. In 1985, the federal government allowed state governments to relax the speed limit on rural interstate roads from 55 mph to 65 mph. Many papers examine the effect of this speed limit change on road safety. Most studies found evidence of safety reductions. Brown et. al (1990) finds that a speed limit increase increased accidents by 19 percent in Alabama. Wagenaar et. al (1990) find that a speed limit increase increased injuries and fatalities in Michigan. Other studies, however, provide conflicting results. Chang and Paniati (1990) and Sidhu (1990) find little to no effect on traffic safety. McCarthy (1991) even find evidence of safety improvements. In 1995, the federal government allowed states to freely set their own speed limits on interstates.

Model

Driving Style Decision

The model presented in this paper will make predictions about the change in the fraction of careful drivers given various parameter changes and will account for the strategic interaction among drivers in choosing a driving style. Assume a population of drivers heterogeneous only in the benefit they receive from reckless driving. Assume that the benefit to reckless driving is exogenous and is greater than the benefit to careful driving. Note that the benefit to reckless driving is gross benefit, not net benefit, so does not include any accident risk or cost. I normalize the benefit to careful driving to 0. I denote the benefit to reckless driving by \( b \) and assume that \( b \) is distributed over \([0, \tilde{b}]\) according to a probability density function \( f(b) \). Each driver randomly encounters another driver and the two play a symmetric normal form game where each chooses a driving style, \( R \) (Reckless) or \( C \) (Careful). I assume that each driver knows his own benefit, but not the other driver’s benefit, and that both know the benefit density function \( f(b) \).

The expected payoff from choosing strategy \( R \) for a type \( b \) driver matched with an unknown \( b \) driver is given by the following utility function:

\[
U(R, b) = b - A[\Pi(R, R)Pr(R) + \Pi(R, C)Pr(C)]
\]

(1)

where \( \Pi(i, j) \) is the probability of an accident given a driver pair playing \((i, j) \) \( \in \{R, C\} \), \( Pr(j) \) is the probability that the other driver plays strategy \( j \), and \( A \) is the
accident cost, which we assume to be independent of the other parameters and driver behavior.

The expected payoff from choosing strategy $C$ is then similarly given by:

$$U(C, b) = -A[\Pi(C, R)Pr(R) + \Pi(C, C)Pr(C)]$$

(2)

Let $p_c = \Pi(R, R) - \Pi(C, R)>0$ and $p_c = \Pi(R, C) - \Pi(C, C)>0$. These represent the additional risk to changing from careful to reckless given the other driver’s strategy. We assume that changing from careful to reckless increases accident probability and that $\Pi(C, R) = \Pi(R, C)$.

Let $\hat{b}$ represent the benefit to the driver indifferent between $R$ and $C$, given all other parameters. Mathematically, this means that $\hat{b}$ is the value of $b$ such that $U(C, b) = U(R, b)$. We can also think of this as the threshold benefit to choose the reckless strategy. Since expected accident cost is independent of $b$, drivers with $b>\hat{b}$ strictly prefer $R$ to $C$. Similarly, drivers with $b<\hat{b}$ strictly prefer $C$ to $R$. It then follows that the proportion of careful drivers is given by $Pr(C) = \int_{0}^{\hat{b}} f(b) \, db = F(\hat{b})$ and the proportion of reckless drivers is given by $Pr(R) = 1 - \int_{0}^{\hat{b}} f(b) \, db = 1 - F(\hat{b})$. Then since $\hat{b}$ equates $U(C, b)$ and $U(R, b)$, $\hat{b}$ must satisfy,

$$\hat{b} = A[p_c(1 - F(\hat{b})) + p_c F(\hat{b})]$$

(3)

Represent the derivative of the right hand side of (3) with respect to $\hat{b}$ by,

$$\Delta(\hat{b}) = A[p_c - p_c f(\hat{b})]$$

(4)

$\Delta(\hat{b})$ is positive if $p_c<p_c$ and negative if $p_c>p_c$. Sufficient conditions for (3) to have a unique interior solution are $\hat{b}>Ap_c, \hat{b}<Ap_c$, and $\Delta(\hat{b})<1$.

**Comparative Statics**

To examine which factors increase the fraction of careful drivers, I determine which factors increase the threshold to reckless driving. Any factor that increases the threshold to reckless driving increases the fraction of careful drivers. Below I show how a change in accident cost affects the fraction of careful drivers. The derivative of $\hat{b}$ with respect to $A$ is then,

$$\frac{d\hat{b}}{dA} = \frac{p_c(1 - F(\hat{b})) + p_c F(\hat{b})}{1 - \Delta(\hat{b})}$$

(5)

Since $p_c$ and $p_c$ are positive and $0<F(\hat{b})<1$, the numerator is positive. Since we assumed $\Delta(\hat{b})<1$ to guarantee a unique interior solution, the denominator is also positive. $\frac{d\hat{b}}{dA}$ is then positive. This means that an increase in accident cost $A$ increases the threshold benefit to reckless driving, therefore increasing the fraction of careful drivers.

I will now analyze how a change in accident probabilities affects $\hat{b}$. I show below that a proportional increase in all accident probabilities is equivalent to that same accident cost, which we assume to be independent of the other parameters and driver behavior.

The expected payoff from choosing strategy $C$ is then similarly given by:

$$U(C, b) = -A[\Pi(C, R)Pr(R) + \Pi(C, C)Pr(C)]$$

(2)

Let $p_c = \Pi(R, R) - \Pi(C, R)>0$ and $p_c = \Pi(R, C) - \Pi(C, C)>0$. These represent the additional risk to changing from careful to reckless given the other driver’s strategy. We assume that changing from careful to reckless increases accident probability and that $\Pi(C, R) = \Pi(R, C)$.

Let $\hat{b}$ represent the benefit to the driver indifferent between $R$ and $C$, given all other parameters. Mathematically, this means that $\hat{b}$ is the value of $b$ such that $U(C, b) = U(R, b)$. We can also think of this as the threshold benefit to choose the reckless strategy. Since expected accident cost is independent of $b$, drivers with $b>\hat{b}$ strictly prefer $R$ to $C$. Similarly, drivers with $b<\hat{b}$ strictly prefer $C$ to $R$. It then follows that the proportion of careful drivers is given by $Pr(C) = \int_{0}^{\hat{b}} f(b) \, db = F(\hat{b})$ and the proportion of reckless drivers is give by $Pr(R) = 1 - \int_{0}^{\hat{b}} f(b) \, db = 1 - F(\hat{b})$. Then since $\hat{b}$ equates $U(C, b)$ and $U(R, b)$, $\hat{b}$ must satisfy,

$$\hat{b} = A[p_c(1 - F(\hat{b})) + p_c F(\hat{b})]$$

(3)

Represent the derivative of the right hand side of (3) with respect to $\hat{b}$ by,

$$\Delta(\hat{b}) = A[p_c - p_c f(\hat{b})]$$

(4)

$\Delta(\hat{b})$ is positive if $p_c<p_c$ and negative if $p_c>p_c$. Sufficient conditions for (3) to have a unique interior solution are $\hat{b}>Ap_c, \hat{b}<Ap_c$, and $\Delta(\hat{b})<1$.

**Comparative Statics**

To examine which factors increase the fraction of careful drivers, I determine which factors increase the threshold to reckless driving. Any factor that increases the threshold to reckless driving increases the fraction of careful drivers. Below I show how a change in accident cost affects the fraction of careful drivers. The derivative of $\hat{b}$ with respect to $A$ is then,

$$\frac{d\hat{b}}{dA} = \frac{p_c(1 - F(\hat{b})) + p_c F(\hat{b})}{1 - \Delta(\hat{b})}$$

(5)

Since $p_c$ and $p_c$ are positive and $0<F(\hat{b})<1$, the numerator is positive. Since we assumed $\Delta(\hat{b})<1$ to guarantee a unique interior solution, the denominator is also positive. $\frac{d\hat{b}}{dA}$ is then positive. This means that an increase in accident cost $A$ increases the threshold benefit to reckless driving, therefore increasing the fraction of careful drivers.

I will now analyze how a change in accident probabilities affects $\hat{b}$. I show below that a proportional increase in all accident probabilities is equivalent to that same accident cost, which we assume to be independent of the other parameters and driver behavior.

The expected payoff from choosing strategy $C$ is then similarly given by:

$$U(C, b) = -A[\Pi(C, R)Pr(R) + \Pi(C, C)Pr(C)]$$

(2)

Let $p_c = \Pi(R, R) - \Pi(C, R)>0$ and $p_c = \Pi(R, C) - \Pi(C, C)>0$. These represent the additional risk to changing from careful to reckless given the other driver’s strategy. We assume that changing from careful to reckless increases accident probability and that $\Pi(C, R) = \Pi(R, C)$.

Let $\hat{b}$ represent the benefit to the driver indifferent between $R$ and $C$, given all other parameters. Mathematically, this means that $\hat{b}$ is the value of $b$ such that $U(C, b) = U(R, b)$. We can also think of this as the threshold benefit to choose the reckless strategy. Since expected accident cost is independent of $b$, drivers with $b>\hat{b}$ strictly prefer $R$ to $C$. Similarly, drivers with $b<\hat{b}$ strictly prefer $C$ to $R$. It then follows that the proportion of careful drivers is given by $Pr(C) = \int_{0}^{\hat{b}} f(b) \, db = F(\hat{b})$ and the proportion of reckless drivers is give by $Pr(R) = 1 - \int_{0}^{\hat{b}} f(b) \, db = 1 - F(\hat{b})$. Then since $\hat{b}$ equates $U(C, b)$ and $U(R, b)$, $\hat{b}$ must satisfy,

$$\hat{b} = A[p_c(1 - F(\hat{b})) + p_c F(\hat{b})]$$

(3)

Represent the derivative of the right hand side of (3) with respect to $\hat{b}$ by,

$$\Delta(\hat{b}) = A[p_c - p_c f(\hat{b})]$$

(4)

$\Delta(\hat{b})$ is positive if $p_c<p_c$ and negative if $p_c>p_c$. Sufficient conditions for (3) to have a unique interior solution are $\hat{b}>Ap_c, \hat{b}<Ap_c$, and $\Delta(\hat{b})<1$.
proportional increase in accident cost. To do so, it is useful to total differentiate (3), the equation that the indifferent benefit satisfies. This gives,

\[
\frac{dB}{\bar{b}} = \left(\frac{1}{1 - \Delta(b)}\right) \left[\frac{dA}{A} + \frac{dp_c}{p_c} \left(\frac{p_r (1 - F(\bar{b}))}{p_r (1 - F(\bar{b})) + p_c F(\bar{b})}\right)\right] + (6)
\]

I will show that the effect of a given proportional increase in all accident probabilities on \(\bar{b}\) equals the effect of that same given proportional increase in \(A\) on \(\bar{b}\). First, let all accident probabilities increase by some given proportion \(c\) and let \(A\) remain unchanged. We can show that the proportional change in \(p_c\) is also \(c\).

\[
\frac{dp_c}{p_c} = \frac{\Pi(R, R) + c\Pi(R, R) - [\Pi(C, R) + c\Pi(C, R)] - [\Pi(R, R) - \Pi(C, R)]}{\Pi(R, R) - \Pi(C, R)} = (8)
\]

Since from the above equation, \(c = \frac{dp_c}{p_c}\), by symmetry \(c = \frac{dp_r}{p_r}\). Using (6), we can determine the effect on \(\bar{b}\),

\[
\frac{\bar{b}}{\bar{b}} = \left(\frac{1}{1 - \Delta(b)}\right) \left[0 + c \left(\frac{p_r (1 - F(\bar{b}))}{p_r (1 - F(\bar{b})) + p_c F(\bar{b})}\right) + c \left(\frac{p_r (1 - F(\bar{b})) + p_c F(\bar{b})}{p_r (1 - F(\bar{b})) + p_c F(\bar{b})}\right)\right] + (10)
\]

Simplifying this expression,

\[
\left(\frac{1}{1 - \Delta(b)}\right) \left[0 + c \left(\frac{p_r (1 - F(\bar{b}) + p_c F(\bar{b}))}{p_r (1 - F(\bar{b})) + p_c F(\bar{b})}\right)\right] = \left(\frac{1}{1 - \Delta(b)}\right) [c]
\]

I then repeat the same process for \(A\). Let \(A\) increase by some given proportion \(c\) and let \(p_r\) and \(p_c\) remain unchanged. This means that, \(c = \frac{dA}{A}\). Using (6), we can determine the effect on \(\bar{b}\),

\[
\frac{\bar{b}}{\bar{b}} = \left(\frac{1}{1 - \Delta(b)}\right) [c + 0] = \left(\frac{1}{1 - \Delta(b)}\right) [c]
\]

Since (9) and (10) are equivalent, a given proportional increase in both \(p_r\) and \(p_c\) and the same proportional increase in \(A\) have the same effect on \(\bar{b}\). Therefore since proportional increase in accident cost. To do so, it is useful to total differentiate (3), the equation that the indifferent benefit satisfies. This gives,

\[
\frac{dB}{\bar{b}} = \left(\frac{1}{1 - \Delta(b)}\right) \left[\frac{dA}{A} + \frac{dp_r}{p_r} \left(\frac{p_r (1 - F(\bar{b}))}{p_r (1 - F(\bar{b})) + p_r F(\bar{b})}\right)\right] + (6)
\]

I will show that the effect of a given proportional increase in all accident probabilities on \(\bar{b}\) equals the effect of that same given proportional increase in \(A\) on \(\bar{b}\). First, let all accident probabilities increase by some given proportion \(c\) and let \(A\) remain unchanged. We can show that the proportional change in \(p_r\) is also \(c\),

\[
\frac{dp_r}{p_r} = \frac{\Pi(R, R) + c\Pi(R, R) - [\Pi(C, R) + c\Pi(C, R)] - [\Pi(R, R) - \Pi(C, R)]}{\Pi(R, R) - \Pi(C, R)} = (8)
\]

Since from the above equation, \(c = \frac{dp_r}{p_r}\), by symmetry \(c = \frac{dp_c}{p_c}\). Using (6), we can determine the effect on \(\bar{b}\),

\[
\frac{\bar{b}}{\bar{b}} = \left(\frac{1}{1 - \Delta(b)}\right) \left[0 + c \left(\frac{p_r (1 - F(\bar{b}))}{p_r (1 - F(\bar{b})) + p_c F(\bar{b})}\right) + c \left(\frac{p_r (1 - F(\bar{b}) + p_c F(\bar{b})}{p_r (1 - F(\bar{b})) + p_c F(\bar{b})}\right)\right] + (10)
\]

Simplifying this expression,

\[
\left(\frac{1}{1 - \Delta(b)}\right) \left[0 + c \left(\frac{p_r (1 - F(\bar{b}) + p_c F(\bar{b})}{p_r (1 - F(\bar{b})) + p_c F(\bar{b})}\right)\right] = \left(\frac{1}{1 - \Delta(b)}\right) [c]
\]

I then repeat the same process for \(A\). Let \(A\) increase by some given proportion \(c\) and let \(p_r\) and \(p_c\) remain unchanged. This means that, \(c = \frac{dA}{A}\). Using (6), we can determine the effect on \(\bar{b}\),

\[
\frac{\bar{b}}{\bar{b}} = \left(\frac{1}{1 - \Delta(b)}\right) [c + 0] = \left(\frac{1}{1 - \Delta(b)}\right) [c]
\]

Since (9) and (10) are equivalent, a given proportional increase in both \(p_r\) and \(p_c\) and the same proportional increase in \(A\) have the same effect on \(\bar{b}\). Therefore since
an increase in $A$ increases $\bar{b}$, a given proportional increase in both $p_1$ and $p_2$ increases $\bar{b}$. This means that increasing both accident risk differentials by the same proportion increases the fraction of careful drivers.

If the accident differentials do not increase by the same proportion, then this conclusion is not always true. Depending on the different probability increases, increasing all accident probabilities could lead to a decrease in the proportion of careful drivers. This idea is useful because it has the potential to explain counterintuitive or inconsistent results for the effect of any factor linked to accident probability on careful driving.

Welfare

Though not the focus of this paper, I will briefly discuss the effects of various parameter changes on welfare. Welfare is given by the expected payoff to a given driver:

$$w = \int f(b) \, db - A[\Pi(R, R)(1 - F(\bar{b})) + 2\Pi(R, C)F(\bar{b})(1 - F(\bar{b})) + \Pi(C, C)F(\bar{b})^2]$$  

(13)

The derivative of welfare with respect to $\bar{b}$, is given by,

$$\frac{dw}{d\bar{b}} = f(\bar{b}) \left[ -\bar{b} + 2A \left( p_1(1 - F(\bar{b})) + p_2F(\bar{b}) \right) \right] = f(\bar{b})\bar{b}$$  

(14)

$\frac{dw}{d\bar{b}}$ is then positive. All else equal, this means that increasing the threshold benefit to choose recklessness, and as a result increasing the fraction of careful drivers, affects welfare positively. However, all else is almost never equal since a change in the threshold benefit is almost always a response to a change in accident cost or probability. Therefore we need to determine the effects of accident cost and probability on welfare as these are more meaningful.

The derivative of welfare with respect to $A$ is given by,

$$\frac{dw}{dA} = \frac{\partial w}{\partial A} + \frac{\partial w}{\partial \bar{b}} \frac{\partial \bar{b}}{dA}$$  

(15)

$\frac{dw}{dA}$ is positive if the positive expression $\frac{\partial w}{\partial A}$ outweighs the negative expression $\frac{\partial w}{\partial \bar{b}} \frac{\partial \bar{b}}{dA}$. However, $\frac{dw}{dA}$ is negative if the opposite is true. We therefore expect the effect of increasing $A$ on welfare to be ambiguous. According to (9) and (10) the effect of an increase in $A$ on welfare is equivalent to the effect of a proportional increase in all accident probabilities on welfare. We similarly expect the effect of a proportional increase in all accident probabilities on welfare to be ambiguous.

Measuring Careful Driving

In practice we do not observe the fraction of careful drivers. Since the data comes from fatal accident reports, we can only find the probability that a driver was reckless given an increase in $A$ increases $\bar{b}$, a given proportional increase in both $p_1$ and $p_2$ increases $\bar{b}$. This means that increasing both accident risk differentials by the same proportion increases the fraction of careful drivers.

If the accident differentials do not increase by the same proportion, then this conclusion is not always true. Depending on the different probability increases, increasing all accident probabilities could lead to a decrease in the proportion of careful drivers. This idea is useful because it has the potential to explain counterintuitive or inconsistent results for the effect of any factor linked to accident probability on careful driving.

Welfare

Though not the focus of this paper, I will briefly discuss the effects of various parameter changes on welfare. Welfare is given by the expected payoff to a given driver:

$$w = \int f(b) \, db - A[\Pi(R, R)(1 - F(\bar{b})) + 2\Pi(R, C)F(\bar{b})(1 - F(\bar{b})) + \Pi(C, C)F(\bar{b})^2]$$  

(13)

The derivative of welfare with respect to $\bar{b}$, is given by,

$$\frac{dw}{d\bar{b}} = f(\bar{b}) \left[ -\bar{b} + 2A \left( p_1(1 - F(\bar{b})) + p_2F(\bar{b}) \right) \right] = f(\bar{b})\bar{b}$$  

(14)

$\frac{dw}{d\bar{b}}$ is then positive. All else equal, this means that increasing the threshold benefit to choose recklessness, and as a result increasing the fraction of careful drivers, affects welfare positively. However, all else is almost never equal since a change in the threshold benefit is almost always a response to a change in accident cost or probability. Therefore we need to determine the effects of accident cost and probability on welfare as these are more meaningful.

The derivative of welfare with respect to $A$ is given by,

$$\frac{dw}{dA} = \frac{\partial w}{\partial A} + \frac{\partial w}{\partial \bar{b}} \frac{\partial \bar{b}}{dA}$$  

(15)

$\frac{dw}{dA}$ is positive if the positive expression $\frac{\partial w}{\partial A}$ outweighs the negative expression $\frac{\partial w}{\partial \bar{b}} \frac{\partial \bar{b}}{dA}$. However, $\frac{dw}{dA}$ is negative if the opposite is true. We therefore expect the effect of increasing $A$ on welfare to be ambiguous. According to (9) and (10) the effect of an increase in $A$ on welfare is equivalent to the effect of a proportional increase in all accident probabilities on welfare. We similarly expect the effect of a proportional increase in all accident probabilities on welfare to be ambiguous.

Measuring Careful Driving

In practice we do not observe the fraction of careful drivers. Since the data comes from fatal accident reports, we can only find the probability that a driver was reckless given
that the driver was involved in a fatal accident. Any effect we find will represent the effect only on drivers involved in fatal accidents. Using Bayes’ rule, we can deduce the effect on the general driver population.

\[
Pr(R|acc) = \frac{Pr(acc|R)Pr(R)}{Pr(acc|R)Pr(R) + Pr(acc|C)Pr(C)} = \tag{16}
\]

\[
P(\Pi, R, R)(1 - F(\hat{b}))^3 + 2\Pi(R, C)\hat{b}(1 - F(\hat{b})) + \Pi(C, C)\hat{b}^2
\]

I assume that all accident probabilities increase by the same proportion. I will use the above equation (14) to show that a proportional increase in all accident probabilities has no direct effect on the probability of observing recklessness in an accident \( Pr(R|acc) \). Assume that each probability changes by some proportion \( c \). This means each \( \Pi \) becomes \( \Pi + c\Pi \). Then the new level of \( Pr(R|acc) \) is given by,

\[
Pr(R|acc) = \frac{(1 + c)\Pi(R, R)(1 - F(\hat{b}))^2 + (1 + c)\Pi(R, C)\hat{b}(1 - F(\hat{b}))}{(1 + c)\Pi(R, R)(1 - F(\hat{b}))^2 + 2(1 + c)\Pi(R, C)\hat{b}(1 - F(\hat{b})) + (1 + c)\Pi(C, C)\hat{b}^2} \tag{18}
\]

Factoring out \((1 + c)\) from each term gives,

\[
\frac{(1 + c)}{} \cdot \frac{\Pi(R, R)(1 - F(\hat{b}))^2 + \Pi(R, C)F(\hat{b})(1 - F(\hat{b}))}{(1 + c)\Pi(R, R)(1 - F(\hat{b}))^2 + (1 + c)\Pi(R, C)\hat{b}(1 - F(\hat{b})) + (1 + c)\Pi(C, C)\hat{b}^2} \tag{19}
\]

This establishes that a change in all accident probabilities by the same proportion has no direct effect on the probability of observing recklessness in a fatal accident. Although there is no direct effect, there is an indirect effect. Since a change in accident probability also affects \( F(\hat{b}) \), a change in a accident probabilities will indirectly affect \( Pr(R|acc) \) through \( F(\hat{b}) \). Then, since a proportional change in accident probabilities does not directly affect \( Pr(R|acc) \), if we observe a change in \( Pr(R|acc) \), we can conclude that this change must be due to a change in \( F(\hat{b}) \).

Next, we can use \( Pr(R|acc) \) to deduce the direction of the change in \( F(\hat{b}) \). From (14), \( Pr(R|acc) \) is decreasing in \( F(\hat{b}) \). So if we observe a decrease in \( Pr(R|acc) \), then there must have been an increase in \( F(\hat{b}) \). From earlier discussion of (9) and (10), we expect a proportional increase in all accident probabilities to increase \( F(\hat{b}) \). Therefore, if we observe a decrease in \( Pr(R|acc) \), we can conclude that the accident probability increase led to an increase in the fraction of careful drivers, \( F(\hat{b}) \).

Now I will show that if we relax the proportional probability change assumption, a decrease in \( Pr(R|acc) \) no longer implies an increase in \( F(\hat{b}) \) and vice versa. This is because we would no longer be able to distinguish the effect of \( F(\hat{b}) \) on \( Pr(R|acc) \) from the effect of the individual probabilities on \( Pr(R|acc) \). An example is the case where

\[
Pr(R|acc) = \frac{Pr(acc|R)Pr(R)}{Pr(acc|R)Pr(R) + Pr(acc|C)Pr(C)} = \tag{16}
\]

\[
P(\Pi, R, R)(1 - F(\hat{b}))^3 + 2\Pi(R, C)\hat{b}(1 - F(\hat{b})) + \Pi(C, C)\hat{b}^2
\]

I assume that all accident probabilities increase by the same proportion. I will use the above equation (14) to show that a proportional increase in all accident probabilities has no direct effect on the probability of observing recklessness in an accident \( Pr(R|acc) \). Assume that each probability changes by some proportion \( c \). This means each \( \Pi \) becomes \( \Pi + c\Pi \). Then the new level of \( Pr(R|acc) \) is given by,

\[
Pr(R|acc) = \frac{(1 + c)\Pi(R, R)(1 - F(\hat{b}))^2 + (1 + c)\Pi(R, C)\hat{b}(1 - F(\hat{b}))}{(1 + c)\Pi(R, R)(1 - F(\hat{b}))^2 + 2(1 + c)\Pi(R, C)\hat{b}(1 - F(\hat{b})) + (1 + c)\Pi(C, C)\hat{b}^2} \tag{18}
\]

Factoring out \((1 + c)\) from each term gives,

\[
\frac{(1 + c)}{} \cdot \frac{\Pi(R, R)(1 - F(\hat{b}))^2 + \Pi(R, C)F(\hat{b})(1 - F(\hat{b}))}{(1 + c)\Pi(R, R)(1 - F(\hat{b}))^2 + (1 + c)\Pi(R, C)\hat{b}(1 - F(\hat{b})) + (1 + c)\Pi(C, C)\hat{b}^2} \tag{19}
\]

This establishes that a change in all accident probabilities by the same proportion has no direct effect on the probability of observing recklessness in a fatal accident. Although there is no direct effect, there is an indirect effect. Since a change in accident probability also affects \( F(\hat{b}) \), a change in a accident probabilities will indirectly affect \( Pr(R|acc) \) through \( F(\hat{b}) \). Then, since a proportional change in accident probabilities does not directly affect \( Pr(R|acc) \), if we observe a change in \( Pr(R|acc) \), we can conclude that this change must be due to a change in \( F(\hat{b}) \).

Next, we can use \( Pr(R|acc) \) to deduce the direction of the change in \( F(\hat{b}) \). From (14), \( Pr(R|acc) \) is decreasing in \( F(\hat{b}) \). So if we observe a decrease in \( Pr(R|acc) \), then there must have been an increase in \( F(\hat{b}) \). From earlier discussion of (9) and (10), we expect a proportional increase in all accident probabilities to increase \( F(\hat{b}) \). Therefore, if we observe a decrease in \( Pr(R|acc) \), we can conclude that the accident probability increase led to an increase in the fraction of careful drivers, \( F(\hat{b}) \).

Now I will show that if we relax the proportional probability change assumption, a decrease in \( Pr(R|acc) \) no longer implies an increase in \( F(\hat{b}) \) and vice versa. This is because we would no longer be able to distinguish the effect of \( F(\hat{b}) \) on \( Pr(R|acc) \) from the effect of the individual probabilities on \( Pr(R|acc) \). An example is the case where...
we assume a proportional increase in all accident probabilities when there is actually a fixed increase. In this example assuming a proportional increase could lead us to reject the model even if people did behave according to the model’s prediction. Since a fixed increase in each of the three accident probabilities has no effect on \( p_c \) or \( p_n \), from (3) we know there is no effect on \( b \) or \( F \). However, I will show that there is a change in \( Pr(R|Acc) \). Assume that all accident probabilities increase by some fixed amount \( c \). Then the value of \( Pr(R|Acc) \) will change:

\[
Pr(R|Acc)_{\text{new}} = \frac{\Pi(R,R) + c(1 - F(b)) + \Pi(R,C) + cF(\bar{b})(1 - F(\bar{b}))}{\Pi(R,R)(1 - F(b))^2 + 2\Pi(R,C)F(b)(1 - F(\bar{b})) + \Pi(C,C)F(\bar{b})^2 + c}
\]

(21) shows that a fixed increase raises the numerator by \( c \cdot [1 - F(\bar{b})] \) and the denominator by \( c \). Unless the initial value of \( Pr(R|Acc) \) is exactly the same as \( 1 - F(\bar{b}) \) then we will observe a change in \( Pr(R|Acc) \). This effect would be due to changes in the accident probabilities rather than due to a change in \( F(\bar{b}) \). If we wrongly assume a proportional increase we would attribute any effect on \( Pr(R|Acc) \) to \( F(\bar{b}) \) even though there actually was no change in \( F(\bar{b}) \). We would conclude that since a fixed increase in all accident probabilities led to a change in \( F(\bar{b}) \) that the model’s prediction was wrong. Another example is the case where the proportional changes in \( \Pi(RR) \) and \( \Pi(CC) \) are very different. We can similarly show that \( Pr(R|Acc) \) does not always accurately predict the change in \( F(\bar{b}) \).

To reiterate, wrongly assuming a proportional increase in all accident probabilities is problematic in testing the model for two reasons. The first reason lies in the model. According to the model, a factor that increases accident probabilities by uneven proportions could actually decrease careful driving. However, with the proportional probability change assumption, the model predicts an increase in careful driving. Therefore, if we observe a decrease in careful driving, it could be that the probabilities changed unevenly, not that the model made wrong predictions. The second reason lies in using \( Pr(R|Acc) \) as an indicator of the fraction of careful drivers involved in fatal accidents. The issue instead lies in using it to deduce the fraction of careful drivers out of all drivers on the road. Even if an increase in the accident probabilities did lead to an increase in the fraction of careful drivers, we could observe a positive \( Pr(R|Acc) \), which we would conclude to mean that the fraction of careful drivers decreased.

we assume a proportional increase in all accident probabilities when there is actually a fixed increase. In this example assuming a proportional increase could lead us to reject the model even if people did behave according to the model’s prediction. Since a fixed increase in each of the three accident probabilities has no effect on \( p_c \) or \( p_n \), from (3) we know there is no effect on \( b \) or \( F(\bar{b}) \). However, I will show that there is a change in \( Pr(R|Acc) \). Assume that all accident probabilities increase by some fixed amount \( c \). Then the value of \( Pr(R|Acc) \) will change:

\[
Pr(R|Acc)_{\text{new}} = \frac{\Pi(R,R) + c(1 - F(\bar{b})) + \Pi(R,C) + cF(\bar{b})(1 - F(\bar{b}))}{\Pi(R,R)(1 - F(b))^2 + 2\Pi(R,C)F(b)(1 - F(\bar{b})) + \Pi(C,C)F(\bar{b})^2 + c}
\]

(22) shows that a fixed increase raises the numerator by \( c \cdot [1 - F(\bar{b})] \) and the denominator by \( c \). Unless the initial value of \( Pr(R|Acc) \) is exactly the same as \( 1 - F(\bar{b}) \) then we will observe a change in \( Pr(R|Acc) \). This effect would be due to changes in the accident probabilities rather than due to a change in \( F(\bar{b}) \). If we wrongly assume a proportional increase we would attribute any effect on \( Pr(R|Acc) \) to \( F(\bar{b}) \) even though there actually was no change in \( F(\bar{b}) \). We would conclude that since a fixed increase in all accident probabilities led to a change in \( F(\bar{b}) \) that the model’s prediction was wrong. Another example is the case where the proportional changes in \( \Pi(RR) \) and \( \Pi(CC) \) are very different. We can similarly show that \( Pr(R|Acc) \) does not always accurately predict the change in \( F(\bar{b}) \).

To reiterate, wrongly assuming a proportional increase in all accident probabilities is problematic in testing the model for two reasons. The first reason lies in the model. According to the model, a factor that increases accident probabilities by uneven proportions could actually decrease careful driving. However, with the proportional probability change assumption, the model predicts an increase in careful driving. Therefore, if we observe a decrease in careful driving, it could be that the probabilities changed unevenly, not that the model made wrong predictions. The second reason lies in using \( Pr(R|Acc) \) as an indicator of the fraction of careful drivers involved in fatal accidents. The issue instead lies in using it to deduce the fraction of careful drivers out of all drivers on the road. Even if an increase in the accident probabilities did lead to an increase in the fraction of careful drivers, we could observe a positive \( Pr(R|Acc) \), which we would conclude to mean that the fraction of careful drivers decreased.
Data and Empirical Strategy

Data

Each year starting in 1975 The US Department of Transportation’s Fatality Analysis Reporting System (FARS) has collected data from police reports from 50 states, District of Columbia, and Puerto Rico on accidents that occurred on a public roadway and resulted in a fatality within 30 days of the accident. The data set includes information on various road characteristics, environmental conditions, vehicle information, violations issued, accident causes, and more. FARS offers different variables at the accident, vehicle and person levels. Each observation has a crash number to make merging the data files possible.

Empirical Strategy

This empirical analysis aims to test the the general model on a few forms of recklessness, including distracted driving, alcohol use, and speeding. To do this I limit the analysis to two driver encounters on interstate highways in the sample and measure how factors that change accident probability affect the probability that the accident report mentioned recklessness. I use data from 1995 to 1997 to take advantage of law changes that produce variation in speed limits, one of the factors I am testing. In 1995 the government lifted the 65 mph maximum and allowed states to set their own limits. In the period from end of 1995 through the end of 1996 some states raised the limit by 5 to 10 mph while others kept the limit at 65 mph. After I limit the data to two driver accidents on interstate highways there are 3,226 observations total, 964 in 1995, 1,118 in 1996, and 1,144 in 1997.

I define the dependent variable in three different ways. I use the regression below for the first two types of recklessness. This regression uses accident level data from 1995-1997 to predict the probability of each type of recklessness given a fatal accident. I use both linear probability fixed effects and logit fixed effects models.

\[
P(\text{reck}/\text{acc})_{ist} = \beta_0 + \beta_1 \text{speedlimit}_{ist} + \beta_2 \text{sign}_{ist} + \beta_3 \text{badweather}_{ist} + \beta_4 \text{dark}_{ist} + \beta_5 X_{ist} + \phi_i + \gamma_s + \epsilon_{ist} \] (24)

The dependent variable \(P(\text{reckless}/\text{accident})\) equals 1 if accident \(i\) in state \(s\) in year \(t\) involved recklessness and 0 if not. I will first define recklessness as distracted driving, which specifically means if the accident report lists causes such as talking, eating, cell phone, fax machine, computer, on-board navigations system, radio, or front window display. 14% of the observations involved these characteristics. I will then define recklessness as alcohol use. 16% of the of the accidents involved alcohol and 14% involved a BAC of at least .08. The reasoning for testing alcohol involvement separately is that the decision to drive while distracted may be different from the decision to drive while drunk. Under the influence of alcohol, a person may not consider risk factors in the same way as a sober driver does. This means that the coefficients may be biased.

\[
P(\text{reck}/\text{acc})_{ist} = \beta_0 + \beta_1 \text{speedlimit}_{ist} + \beta_2 \text{sign}_{ist} + \beta_3 \text{badweather}_{ist} + \beta_4 \text{dark}_{ist} + \beta_5 X_{ist} + \phi_i + \gamma_s + \epsilon_{ist} \] (25)

\[
P(\text{reck}/\text{acc})_{ist} = \beta_0 + \beta_1 \text{speedlimit}_{ist} + \beta_2 \text{sign}_{ist} + \beta_3 \text{badweather}_{ist} + \beta_4 \text{dark}_{ist} + \beta_5 X_{ist} + \phi_i + \gamma_s + \epsilon_{ist} \] (26)

The dependent variable \(P(\text{reckless}/\text{accident})\) equals 1 if accident \(i\) in state \(s\) in year \(t\) involved recklessness and 0 if not. I will first define recklessness as distracted driving, which specifically means if the accident report lists causes such as talking, eating, cell phone, fax machine, computer, on-board navigations system, radio, or front window display. 14% of the observations involved these characteristics. I will then define recklessness as alcohol use. 16% of the of the accidents involved alcohol and 14% involved a BAC of at least .08. The reasoning for testing alcohol involvement separately is that the decision to drive while distracted may be different from the decision to drive while drunk. Under the influence of alcohol, a person may not consider risk factors in the same way as a sober driver does. This means that the coefficients may be biased.

Empirical Strategy

This empirical analysis aims to test the the general model on a few forms of recklessness, including distracted driving, alcohol use, and speeding. To do this I limit the analysis to two driver encounters on interstate highways in the sample and measure how factors that change accident probability affect the probability that the accident report mentioned recklessness. I use data from 1995 to 1997 to take advantage of law changes that produce variation in speed limits, one of the factors I am testing. In 1995 the government lifted the 65 mph maximum and allowed states to set their own limits. In the period from end of 1995 through the end of 1996 some states raised the limit by 5 to 10 mph while others kept the limit at 65 mph. After I limit the data to two driver accidents on interstate highways there are 3,226 observations total, 964 in 1995, 1,118 in 1996, and 1,144 in 1997.

I define the dependent variable in three different ways. I use the regression below for the first two types of recklessness. This regression uses accident level data from 1995-1997 to predict the probability of each type of recklessness given a fatal accident. I use both linear probability fixed effects and logit fixed effects models.

\[
P(\text{reck}/\text{acc})_{ist} = \beta_0 + \beta_1 \text{speedlimit}_{ist} + \beta_2 \text{sign}_{ist} + \beta_3 \text{badweather}_{ist} + \beta_4 \text{dark}_{ist} + \beta_5 X_{ist} + \phi_i + \gamma_s + \epsilon_{ist} \] (24)

\[
P(\text{reck}/\text{acc})_{ist} = \beta_0 + \beta_1 \text{speedlimit}_{ist} + \beta_2 \text{sign}_{ist} + \beta_3 \text{badweather}_{ist} + \beta_4 \text{dark}_{ist} + \beta_5 X_{ist} + \phi_i + \gamma_s + \epsilon_{ist} \] (25)

\[
P(\text{reck}/\text{acc})_{ist} = \beta_0 + \beta_1 \text{speedlimit}_{ist} + \beta_2 \text{sign}_{ist} + \beta_3 \text{badweather}_{ist} + \beta_4 \text{dark}_{ist} + \beta_5 X_{ist} + \phi_i + \gamma_s + \epsilon_{ist} \] (26)
towards zero if drunk drivers fail to consider the risks that sober drivers do. If variables that are significant under the distracted driving definition are no longer significant under the drunk driving definition, then we have evidence to support the above reasoning.

I include state and year fixed effects, $\gamma_1$ and $\gamma_2$, to account for potential endogeneity issues. State fixed effects will control for any differences among states that are fixed over time, like demographic characteristics or general state traffic environments. Year fixed effects will control for differences over time that are fixed across states, like population growth or general economic conditions.

I use a logit fixed effects model because the dependent variable is a binary 0 or 1 variable, not continuous on the 0 to 1 interval. The OLS fixed effects regression could generate predicted probability values of less than 0 or greater than 1, which would certainly be wrong as no probability can be greater than 1 or less than 0. Here, the probabilities of distracted and intoxicated driving are 14, which is close enough to 0 to warrant concern about predicted probabilities lower than 0 from the OLS regression.

I use the regression below for the third type of recklessness, speeding. I use OLS fixed effects with tobit and tobit fixed effects models as robustness checks.

$speeding = \beta_0 + \beta_1 speedlimit_{stat} + \beta_2 sign_{stat} + \beta_3 badweather_{stat} + \beta_4 dark_{stat} + \beta_5 X_{stat} + \phi_t + \gamma_1 + \epsilon_{stat}$

The dependent variable, $speeding$, is defined as the total number of miles over the speed limit of both cars involved in the accident. Since 62% of the accidents did not involve a vehicle driving over the speed limit, I also include tobit models as robustness checks.

There are several explanatory variable of interest in these three analyses. Speed limit is the numerical speed limit on the road where the accident occurred. Bad weather is a dummy variable for if the accident occurred in adverse weather conditions including rain, snow, and sleet. Sign is a dummy variable for if there was a regulatory or warning sign. Dark is a dummy variable for if the accident occurred when it was dark outside. Number of lanes is the number of traffic lanes on the road. Construction, intersection, ramp, and curved road are dummy variables for if the road possessed these characteristics. Summary statistics for all variables are presented in Table 1. In Table 1, speeding has a different number of observations than the rest of the variables because the police reports did not list travel speed for every accident.

**Speed Limits and Road Signs**

There are certain issues specific to speed limits and road signs that make analyzing these variables more complicated.

An endogeneity issue is that regulators may set speed limits and decide road sign locations based on the same road conditions that influence driver behavior. First, regulators may choose speed limits based on road conditions such that a higher speed limit might not increase accident probability. For example, a divided highway might towards zero if drunk drivers fail to consider the risks that sober drivers do. If variables that are significant under the distracted driving definition are no longer significant under the drunk driving definition, then we have evidence to support the above reasoning.

I include state and year fixed effects, $\gamma_1$ and $\gamma_2$, to account for potential endogeneity issues. State fixed effects will control for any differences among states that are fixed over time, like demographic characteristics or general state traffic environments. Year fixed effects will control for differences over time that are fixed across states, like population growth or general economic conditions.

I use a logit fixed effects model because the dependent variable is a binary 0 or 1 variable, not continuous on the 0 to 1 interval. The OLS fixed effects regression could generate predicted probability values of less than 0 or greater than 1, which would certainly be wrong as no probability can be greater than 1 or less than 0. Here, the probabilities of distracted and intoxicated driving are 14, which is close enough to 0 to warrant concern about predicted probabilities lower than 0 from the OLS regression.

I use the regression below for the third type of recklessness, speeding. I use OLS fixed effects with tobit and tobit fixed effects models as robustness checks.

$speeding = \beta_0 + \beta_1 speedlimit_{stat} + \beta_2 sign_{stat} + \beta_3 badweather_{stat} + \beta_4 dark_{stat} + \beta_5 X_{stat} + \phi_t + \gamma_1 + \epsilon_{stat}$

The dependent variable, $speeding$, is defined as the total number of miles over the speed limit of both cars involved in the accident. Since 62% of the accidents did not involve a vehicle driving over the speed limit, I also include tobit models as robustness checks.

There are several explanatory variable of interest in these three analyses. Speed limit is the numerical speed limit on the road where the accident occurred. Bad weather is a dummy variable for if the accident occurred in adverse weather conditions including rain, snow, and sleet. Sign is a dummy variable for if there was a regulatory or warning sign. Dark is a dummy variable for if the accident occurred when it was dark outside. Number of lanes is the number of traffic lanes on the road. Construction, intersection, ramp, and curved road are dummy variables for if the road possessed these characteristics. Summary statistics for all variables are presented in Table 1. In Table 1, speeding has a different number of observations than the rest of the variables because the police reports did not list travel speed for every accident.

**Speed Limits and Road Signs**

There are certain issues specific to speed limits and road signs that make analyzing these variables more complicated.

An endogeneity issue is that regulators may set speed limits and decide road sign locations based on the same road conditions that influence driver behavior. First, regulators may choose speed limits based on road conditions such that a higher speed limit might not increase accident probability. For example, a divided highway might...
have a lower accident probability and higher speed limit compared to a local, heavily travelled road. To mitigate this issue, we restrict this analysis to interstate roads.

Another endogeneity issue is that lower speed limits and more traffic signs may be more prevalent on more dangerous roads, which already witness little reckless driving. These would then be associated with less reckless driving due to the dangerous road conditions that discourage reckless driving, not due to the speed limit or traffic sign. The coefficients would also be representing the effects of safer road conditions on careful driving. The coefficient on speed limit would be biased towards zero or even become positive when we predict a negative sign. The coefficient on traffic sign would be negatively biased, the direction that we predict, leading us to falsely conclude that there is a significant negative effect. Restricting the analysis to interstate roads also helps to control for this issue by reducing the variation in the road conditions. Many interstate roads are similar, but since conditions still do vary more controls are needed. The data set is very detailed and allows us to include many of these road condition related controls. Some include number of lanes, construction, intersection, ramp, vehicle type, etc. In the above regression equation, X represents this vector of controls.

Results

Distracted Driving

Table 2 shows different types of regressions for distracted driving. The first version shows OLS with controls. Such controls include number of lanes, construction, intersection, ramp, curved road, urban road, truck, bus, motorcycle, number of occupants, week day, rush hour, and teen driver. After adding controls, the coefficients on bad weather, dark lighting, and number of lanes are significant at the 1% level and speed limit is significant at the 5% level. However, these estimates could be biased. As stated before, there may be certain factors specific to each state or time period that are also correlated with both our explanatory variables of interest and recklessness. The second version uses state and year fixed effects to account for some endogeneity concerns in the OLS. After accounting for state and year fixed effects, the results are similar to (1) but most coefficients are significant at 5% level instead of the 1% level. However, there is still an econometric issue since the dependent variable is a discrete 0 or 1. The third version uses a fixed effects logit regression, which can account for the issue in the OLS fixed effects. The results are similar to (2).

Overall, these results support the model’s predictions about several variables. The coefficient on bad weather is consistently negative and significant, supporting previous research and indicating that distracted driving decreases in bad weather. The coefficients on dark lighting and the number of lanes are also consistently negative and significant, indicating that distracted driving decreases under these conditions as well.
Alcohol Use

Next, I define reckless accidents differently. Instead of representing distracted driving the dependent variable represents alcohol involvement. In Table 3, regression (1) shows logit fixed effects under the distracted definition, (2) shows logit fixed effects with accidents involving any amount of alcohol, and (3) shows logit fixed effects with accidents involving a BAC of at least .08. In both (2) and (3) the only significant variable is number of lanes. This supports the hypothesis laid out in the empirical strategy section that the decision to be distracted while driving is different from the decision to use alcohol and drive.

Speeding

Next, I use speeding as the reckless driving behavior in the dependent variable. I measure this as as total miles per hour over the speed limit involved in the accident. In Table 4, regression (1) shows an OLS regression, (2) shows fixed effects, (3) shows tobit, and (4) shows fixed effects tobit. The signs of the marginal effects and significance levels of the coefficients are generally consistent across all specifications for the variables of interest, most of which are significant at the 1% level and all at the 5% level. We see the expected results that bad weather, traffic signs, and ramps are associated with less speeding. Although we also see that speed limits are associated with less speeding, we should conclude that higher speed limits induce drivers to reduce their speed. All else equal, a higher speed limit would simply mean that a driver is less likely to be above the speed limit, not that the driver actively reduced speed.

These results differ somewhat from the distracted driving results. Here, traffic signs and ramps are associated with a negative, significant effect on reckless behavior while they were not significant in the distracted driving case. Dark lighting is associated with a positive significant effect on reckless behavior while it was associated with a negative significant effect in the distracted driving case. These results suggest that the avenue through which the driver changes behavior may depend on what factor increased accident probability. It may be the case, for example, that drivers respond to traffic signs by reducing their speed but not by reducing their level of distraction.

Speed Limit

I conduct a separate analysis of speed limit to address empirical issues specific to speed limit. Table 5 shows the logit fixed effects regression with various combinations of controls. The marginal effects of speed limit at the means are .00548, .00383, .00379, .00417, and .00461 respectively. In (1) speed limit is the only variable and is significant. The first set of controls represents various road characteristics, like construction or number of lanes. After adding these controls, speed limit is no longer significant. Regression (3) adds a second set of controls representing various vehicle characteristics and speed limit remains insignificant. This slightly raises the marginal effect of speed limit at the means and the coefficient becomes significant at the 10% level. Regression

Speeding

Next, I use speeding as the reckless driving behavior in the dependent variable. I measure this as as total miles per hour over the speed limit involved in the accident. In Table 4, regression (1) shows an OLS regression, (2) shows fixed effects, (3) shows tobit, and (4) shows fixed effects tobit. The signs of the marginal effects and significance levels of the coefficients are generally consistent across all specifications for the variables of interest, most of which are significant at the 1% level and all at the 5% level. We see the expected results that bad weather, traffic signs, and ramps are associated with less speeding. Although we also see that speed limits are associated with less speeding, we should conclude that higher speed limits induce drivers to reduce their speed. All else equal, a higher speed limit would simply mean that a driver is less likely to be above the speed limit, not that the driver actively reduced speed.

These results differ somewhat from the distracted driving results. Here, traffic signs and ramps are associated with a negative, significant effect on reckless behavior while they were not significant in the distracted driving case. Dark lighting is associated with a positive significant effect on reckless behavior while it was associated with a negative significant effect in the distracted driving case. These results suggest that the avenue through which the driver changes behavior may depend on what factor increased accident probability. It may be the case, for example, that drivers respond to traffic signs by reducing their speed but not by reducing their level of distraction.

Speed Limit

I conduct a separate analysis of speed limit to address empirical issues specific to speed limit. Table 5 shows the logit fixed effects regression with various combinations of controls. The marginal effects of speed limit at the means are .00548, .00383, .00379, .00417, and .00461 respectively. In (1) speed limit is the only variable and is significant. The first set of controls represents various road characteristics, like construction or number of lanes. After adding these controls, speed limit is no longer significant. Regression (3) adds a second set of controls representing various vehicle characteristics and speed limit remains insignificant. This slightly raises the marginal effect of speed limit at the means and the coefficient becomes significant at the 10% level. Regression
(4) adds bad weather and the coefficient on speed limit remains positive but becomes significant. Regression (5) adds additional controls, producing similar results to (4).

The model and regression results differ on speed limits. The model predicts a negative relationship between speed limits and the probability that we observe recklessness in an accident, but we observe either no effect or a positive significant effect. There are a few reasons for this. First, our speed limit variable is the effect of the observed speed limit where the accident took place, but we are using it as a measure of posted speed limit. Any variable that changes where fatal accidents take place will affect the observed speed limit even though it does not change the posted speed limit. From regression (4) in Table 5, adding the bad weather variable makes the speed limit coefficient significant. Bad weather is exogenous so does not affect the posted speed limit, but does affect the observed speed limit. This means that when we put bad weather into the model we control for a correlation that does not exist between posted speed limit and bad weather and introduce a positive bias to posted speed limit.

### Conclusion

I have presented a way to analyze driver behavior by taking into account accident probability, accident cost, driver benefit, and other drivers’ behavior. The model predicts that drivers respond to traffic variables. With a higher accident cost, drivers become more careful. With a higher accident probability, the behavior change depends on relative changes in accident probability. I then empirical tested the effects of several traffic variables on three forms of recklessness, distracted driving, speeding, and alcohol use. I find evidence that drivers may become more careful in response to bad weather, dark lighting, ramps, traffic signs, and congestion.

Opportunities for future research include relaxing certain assumptions in the model and empirically testing welfare effects. First, the model could account for an accident cost that depends on accident probability. It may be the case that higher accident probability makes accident cost higher. Second, the model could account for the idea that a factor that increases accident probability may be likely change each driver’s inherent benefit to careful and reckless driving. Many traffic safety variables not only increase safety, but also decrease the inherent benefit to careful or reckless driving. Speed limits, for example, allow people to travel faster. This may increase the benefit to careful driving since drivers can reach their destinations faster. It may also decrease the benefit to distracted driving if the original motivation for distracted driving was boredom from traveling too slow. Third, another study could empirically test for changes in welfare given various traffic safety factor changes.
References


APPENDIX A:
Thirteenth Annual Carroll Round Presentation Schedule

Session 1A
Chair: Robert Cumby (Professor, Georgetown University)
Saugata Sen (London School of Economics)
The Indian Inflation Puzzle: Is the Reserve Bank of India using the right price index for its headline inflation measure?
Discussants: Benjamin Shoesmith

Benjamin Shoesmith (University of North Carolina Wilmington)
Federal Reserve Stimulus Effects on Treasury and Mortgage Rates
Discussants: Viktoryia Pilinko and Andrei Romancenco

Viktoryia Pilinko and Andrei Romancenco (Stockholm School of Economics in Riga)
A Macro-financial Model for Credit Risk Stress Testing: The Case of Latvia
Discussant: Russell Black

Russell Black (Oxford University)
Finance and Industry in Interwar India
Discussant: Saugata Sen

Session 1B
Chair: Raj Desai (Associate Professor, Georgetown University)
Meredith Strike (Georgetown University)
Subnational Politics and Foreign Direct Investment in India
Discussant: Eric Aldenhoff

Eric Aldenhoff (University of Maryland, College Park)
Unemployment Benefit Increases and Mortality
Discussant: Rachel Szymanski

Rachel Szymanski (Georgetown University)
Driving Style as a Game: Behavioral Responses to Road Safety Factors
Discussant: Joshua Walker

Joshua Walker (Lancaster University)
The Self-employed and Saving: New Evidence from Great Britain
Discussants: Benjamin Shoesmith

Benjamin Shoesmith (University of North Carolina Wilmington)
Federal Reserve Stimulus Effects on Treasury and Mortgage Rates
Discussants: Viktoryia Pilinko and Andrei Romancenco

Viktoryia Pilinko and Andrei Romancenco (Stockholm School of Economics in Riga)
A Macro-financial Model for Credit Risk Stress Testing: The Case of Latvia
Discussant: Russell Black

Russell Black (Oxford University)
Finance and Industry in Interwar India
Discussant: Saugata Sen

Session 2A
Chair: James Vreeland (Professor, Georgetown University)
Jonathan Pedde (Dartmouth College)
Political Risk and Stock Market Returns: Evidence from International Equity Markets and Option Markets
Discussant: Samsun Knight

Samsun Knight (Oberlin College)
Democracy and the Effects of Negative Exogenous Income Shocks
Discussant: Christopher Stromeyer

Christopher Stromeyer (Georgetown University)
Liberal Objectives: Does the Ideological Distance between the United States and Developing Countries Affect how much Aid they Receive?
Discussant: Jonathan Pedde

Session 2B
Chair: Steven Radelet (Distinguished Professor in the Practice of Development, Georgetown University)
Russell Morton (Princeton University)
On the Other Side of the Tracks: Neighborhood Effects and Segregation in Morocco
Discussant: Eric Aldenhoff

Eric Aldenhoff (University of Maryland, College Park)
Unemployment Benefit Increases and Mortality
Discussant: Rachel Szymanski

Rachel Szymanski (Georgetown University)
Driving Style as a Game: Behavioral Responses to Road Safety Factors
Discussant: Joshua Walker

Joshua Walker (Lancaster University)
The Self-employed and Saving: New Evidence from Great Britain
Discussants: Benjamin Shoesmith

Benjamin Shoesmith (University of North Carolina Wilmington)
Federal Reserve Stimulus Effects on Treasury and Mortgage Rates
Discussants: Viktoryia Pilinko and Andrei Romancenco

Viktoryia Pilinko and Andrei Romancenco (Stockholm School of Economics in Riga)
A Macro-financial Model for Credit Risk Stress Testing: The Case of Latvia
Discussant: Russell Black

Russell Black (Oxford University)
Finance and Industry in Interwar India
Discussant: Saugata Sen

Session 1B
Chair: Raj Desai (Associate Professor, Georgetown University)
Meredith Strike (Georgetown University)
Subnational Politics and Foreign Direct Investment in India
Discussant: Eric Aldenhoff

Eric Aldenhoff (University of Maryland, College Park)
Unemployment Benefit Increases and Mortality
Discussant: Rachel Szymanski

Rachel Szymanski (Georgetown University)
Driving Style as a Game: Behavioral Responses to Road Safety Factors
Discussant: Joshua Walker

Joshua Walker (Lancaster University)
The Self-employed and Saving: New Evidence from Great Britain
Discussants: Benjamin Shoesmith

Benjamin Shoesmith (University of North Carolina Wilmington)
Federal Reserve Stimulus Effects on Treasury and Mortgage Rates
Discussants: Viktoryia Pilinko and Andrei Romancenco

Viktoryia Pilinko and Andrei Romancenco (Stockholm School of Economics in Riga)
A Macro-financial Model for Credit Risk Stress Testing: The Case of Latvia
Discussant: Russell Black

Russell Black (Oxford University)
Finance and Industry in Interwar India
Discussant: Saugata Sen

Session 2A
Chair: Steven Radelet (Distinguished Professor in the Practice of Development, Georgetown University)
Russell Morton (Princeton University)
On the Other Side of the Tracks: Neighborhood Effects and Segregation in Morocco
Discussant: Eric Aldenhoff

Eric Aldenhoff (University of Maryland, College Park)
Unemployment Benefit Increases and Mortality
Discussant: Rachel Szymanski

Rachel Szymanski (Georgetown University)
Driving Style as a Game: Behavioral Responses to Road Safety Factors
Discussant: Joshua Walker

Joshua Walker (Lancaster University)
The Self-employed and Saving: New Evidence from Great Britain
Discussants: Benjamin Shoesmith

Benjamin Shoesmith (University of North Carolina Wilmington)
Federal Reserve Stimulus Effects on Treasury and Mortgage Rates
Discussants: Viktoryia Pilinko and Andrei Romancenco

Viktoryia Pilinko and Andrei Romancenco (Stockholm School of Economics in Riga)
A Macro-financial Model for Credit Risk Stress Testing: The Case of Latvia
Discussant: Russell Black

Russell Black (Oxford University)
Finance and Industry in Interwar India
Discussant: Saugata Sen
APPENDIX A:
Thirteenth Annual Carroll Round Presentation Schedule

Johnny Huynh (Pomona College)
Heterogeneity in Conditional Cash Transfer Effects: Evidence from a Randomized Trial in Malawi
Discussant: Brian Goggin

Brian Goggin (Georgetown University)
Democracy Through Demands or Connections? Local Government Representation and Access to BPL Cards in India
Discussant: Russell Morton

Session 3A
Chair: Carlo Prato (Assistant Professor, Georgetown University)
Thomas Bumberger (University of Cambridge)
An Analysis of Spatial-Based Signalling Games
Discussant: Matthew Klein

Matthew Klein (University of Chicago)
Low-wage Work & the Great Recession: Application of the "Institutional Inclusivity" Model
Discussant: Nikhil Kalathil

Nikhil Kalathil (Oberlin College)
When Technologies Compete: Disruptive Innovations and First Mover Advantage
Discussant: Thomas Bumberger

Session 3B
Chair: Rodney Ludema (Professor, Georgetown University)
Heather Hedges (Georgetown University)
The Challenge of National Treatment in International Trade: Evidence from WTO Dispute Settlement
Discussant: Kyle Coombs

Kyle Coombs (Macalester College)
Do U.S. Border Enforcement Operations Increase Human Smuggling Fees Along the U.S.-Mexico Border?
Discussant: Natalie Langburd

APPENDIX A:
Thirteenth Annual Carroll Round Presentation Schedule

Johnny Huynh (Pomona College)
Heterogeneity in Conditional Cash Transfer Effects: Evidence from a Randomized Trial in Malawi
Discussant: Brian Goggin

Brian Goggin (Georgetown University)
Democracy Through Demands or Connections? Local Government Representation and Access to BPL Cards in India
Discussant: Russell Morton

Session 3A
Chair: Carlo Prato (Assistant Professor, Georgetown University)
Thomas Bumberger (University of Cambridge)
An Analysis of Spatial-Based Signalling Games
Discussant: Matthew Klein

Matthew Klein (University of Chicago)
Low-wage Work & the Great Recession: Application of the "Institutional Inclusivity" Model
Discussant: Nikhil Kalathil

Nikhil Kalathil (Oberlin College)
When Technologies Compete: Disruptive Innovations and First Mover Advantage
Discussant: Thomas Bumberger

Session 3B
Chair: Rodney Ludema (Professor, Georgetown University)
Heather Hedges (Georgetown University)
The Challenge of National Treatment in International Trade: Evidence from WTO Dispute Settlement
Discussant: Kyle Coombs

Kyle Coombs (Macalester College)
Do U.S. Border Enforcement Operations Increase Human Smuggling Fees Along the U.S.-Mexico Border?
Discussant: Natalie Langburd
APPENDIX A:
Thirteenth Carroll Round Presentation Schedule

Session 4B
Chair: Charles Udomsaph (Visiting Assistant Professor, Georgetown University)
Soumyajit Mazumder (Georgetown University)
Can I Stay a BIT Longer? The Pernicious Effect of Bilateral Investment Treaties on Political Survival
Discussant: Robert Dent
Robert Dent (University of Virginia)
Investigating Risk Tolerance Stability using the NLSY79
Discussant: Dawn Chan
Dawn Chan (Georgetown University)
Discussant: Soumyajit Mazumder

APPENDIX B:
Past Speakers

First Annual Carroll Round
(April 5-7, 2002)
Roger W. Ferguson, Federal Reserve Board of Governors
Donald L. Kohn, Federal Reserve Board of Governors
Lawrence B. Lindsey, Assistant to the President and National Economic Council
Edwin M. Truman, Institute for International Economics
John Williamson, Institute for International Economics

Second Annual Carroll Round
(April 11-13, 2003)
R. Glenn Hubbard, Council of Economic Advisers and Columbia University
Donald L. Kohn, Federal Reserve Board of Governors
John Williamson, Institute for International Economics

Third Annual Carroll Round
(April 15-18, 2004)
Donald L. Kohn, Federal Reserve Board of Governors
John F. Nash, Jr., Princeton University (1994 Nobel Laureate)
Peter R. Orszag, The Brookings Institute

Fourth Annual Carroll Round
(April 22-24, 2005)
Ben S. Bernanke, Federal Reserve Board of Governors
William Easterly, New York University
Maurice Obstfeld, University of California at Berkeley
Edwin M. Truman, Institute for International Economics

Fifth Annual Carroll Round
(April 28-30, 2006)
Kemal Dervis, United Nations Development Programme
Thomas C. Schelling, University of Maryland (2005 Nobel Laureate)

Sixth Annual Carroll Round
(April 19-22, 2007)
Grant D. Aldonas, Center for Strategic and International Studies
François Bourguignon, Chief Economist and Senior Vice President of the World Bank
Randall Kroszner, Federal Reserve Board of Governors

Seventh Annual Carroll Round
(April 17-20, 2008)
Dr. Susan C. Athey, Harvard University
Dr. Philip I. Levy, American Enterprise Institute
Dr. Steven Radelet, Senior Fellow at the Center for International Economics

First Annual Carroll Round
(April 5-7, 2002)
Roger W. Ferguson, Federal Reserve Board of Governors
Donald L. Kohn, Federal Reserve Board of Governors
Lawrence B. Lindsey, Assistant to the President and National Economic Council
Edwin M. Truman, Institute for International Economics
John Williamson, Institute for International Economics

Second Annual Carroll Round
(April 11-13, 2003)
R. Glenn Hubbard, Council of Economic Advisers and Columbia University
Donald L. Kohn, Federal Reserve Board of Governors
John Williamson, Institute for International Economics

Third Annual Carroll Round
(April 15-18, 2004)
Donald L. Kohn, Federal Reserve Board of Governors
John F. Nash, Jr., Princeton University (1994 Nobel Laureate)
Peter R. Orszag, The Brookings Institute

Fourth Annual Carroll Round
(April 22-24, 2005)
Ben S. Bernanke, Federal Reserve Board of Governors
William Easterly, New York University
Maurice Obstfeld, University of California at Berkeley
Edwin M. Truman, Institute for International Economics

Fifth Annual Carroll Round
(April 28-30, 2006)
Kemal Dervis, United Nations Development Programme
Thomas C. Schelling, University of Maryland (2005 Nobel Laureate)
APPENDIX B: Past Speakers

Eighth Annual Carroll Round (April 16-19, 2009)
Dr. Eric S. Maskin, Princeton University (2007 Nobel Laureate)
Dr. Nassim Nicholas Taleb, Universa Investments and New York University

Ninth Annual Carroll Round (April 22-25, 2010)
Philip I. Levy, American Enterprise Institute
Lant Pritchett, Harvard Kennedy School

Tenth Annual Carroll Round (April 14-17, 2011)
Jagdish Bhagwati, Columbia University
Joseph Stiglitz, Columbia University

Eleventh Annual Carroll Round (April 19-22, 2012)
Dr. Jonathan Levin, Stanford University
Gene Sperling, Director of the National Economic Council

Twelfth Annual Carroll Round (April 18-21, 2013)
John Taylor, Stanford University
Janet Currie, Princeton University

Thirteenth Annual Carroll Round (April 10-13, 2014)
Peter Diamond, Massachusetts Institute of Technology (2010 Nobel Laureate)
Martin Ravallion, Georgetown University

APPENDIX C: Former Carroll Round Steering Committees

First Annual Carroll Round (April 5-7, 2002)
Christopher L. Griffin, chair (SFS ’02)
William B. Brady (SFS ’02)
Cullen A. Drescher (COL ’04)
Meredith L. Gilbert (COL ’04)
Joshua M. Harris (SFS ’02)
Andrew T. Hayashi (SFS ’02)
Mark R. Longstreth (SFS ’04)
Kathryn E. Magee (SFS ’02)
Ryan F. Michals (SFS ’02)
J. Brendan Mullin (SFS ’02)
Scott E. Pedowitz (SFS ’02)
Waheed A. Sheikh (SFS ’04)

Second Annual Carroll Round (April 11-13, 2003)
Seth M. Kundrot, chair (SFS ’03)
Nada M. Abdelnour (SFS ’03)
Maria M. Arhancet (SFS ’04)
Victoria E. Bembernista (SFS ’03)
Michael J. Callen (SFS ’05)
Eric M. Fischer (SFS ’03)
Daphne Francois (SFS/GRD ’04)
Meredith L. Gilbert (COL ’04)
Jeffrey M. Harris (COL ’03)
Robert S. Katz (COL ’04)

Marina Laffriere (SFS ’06)
Lu Shi (SFS ’03)
Stacey H. Tsai (SFS ’03)
Robert T. Wrob (SFS ’03)
Erica C. Yu (COL ’05)

Third Annual Carroll Round (April 15-18, 2004)
Meredith L. Gilbert, chair (COL ’04)
Heber M. Delgado-Medrano (SFS ’06)
Ryan V. Fraser (SFS ’04)
Tetyana V. Gaponenko (SFS ’07)
Yun jung Cindy Jin (SFS ’05)
Sarah H. Knupp (SFS ’04)
Robert S. Katz (COL ’04)

First Annual Carroll Round (April 16-19, 2009)
Dr. Eric S. Maskin, Princeton University (2007 Nobel Laureate)
Dr. Nassim Nicholas Taleb, Universa Investments and New York University

Ninth Annual Carroll Round (April 22-25, 2010)
Philip I. Levy, American Enterprise Institute
Lant Pritchett, Harvard Kennedy School

Tenth Annual Carroll Round (April 14-17, 2011)
Jagdish Bhagwati, Columbia University
Joseph Stiglitz, Columbia University

Eleventh Annual Carroll Round (April 19-22, 2012)
Dr. Jonathan Levin, Stanford University
Gene Sperling, Director of the National Economic Council

Twelfth Annual Carroll Round (April 18-21, 2013)
John Taylor, Stanford University
Janet Currie, Princeton University

Thirteenth Annual Carroll Round (April 10-13, 2014)
Peter Diamond, Massachusetts Institute of Technology (2010 Nobel Laureate)
Martin Ravallion, Georgetown University

APPENDIX B: Past Speakers

APPENDIX C: Former Carroll Round Steering Committees

First Annual Carroll Round (April 5-7, 2002)
Christopher L. Griffin, chair (SFS ’02)
William B. Brady (SFS ’02)
Cullen A. Drescher (COL ’04)
Meredith L. Gilbert (COL ’04)
Joshua M. Harris (SFS ’02)
Andrew T. Hayashi (SFS ’02)
Mark R. Longstreth (SFS ’04)
Kathryn E. Magee (SFS ’02)
Ryan F. Michals (SFS ’02)
J. Brendan Mullin (SFS ’02)
Scott E. Pedowitz (SFS ’02)
Waheed A. Sheikh (SFS ’04)

Second Annual Carroll Round (April 11-13, 2003)
Seth M. Kundrot, chair (SFS ’03)
Nada M. Abdelnour (SFS ’03)
Maria M. Arhancet (SFS ’04)
Victoria E. Bembernista (SFS ’03)
Michael J. Callen (SFS ’05)
Eric M. Fischer (SFS ’03)
Daphne Francois (SFS/GRD ’04)
Meredith L. Gilbert (COL ’04)
Jeffrey M. Harris (COL ’03)
Robert S. Katz (COL ’04)

Marina Laffriere (SFS ’06)
Lu Shi (SFS ’03)
Stacey H. Tsai (SFS ’03)
Robert T. Wrob (SFS ’03)
Erica C. Yu (COL ’05)

Third Annual Carroll Round (April 15-18, 2004)
Meredith L. Gilbert, chair (COL ’04)
Heber M. Delgado-Medrano (SFS ’06)
Ryan V. Fraser (SFS ’04)
Tetyana V. Gaponenko (SFS ’07)
Yun jung Cindy Jin (SFS ’05)
Sarah H. Knupp (SFS ’04)
Robert S. Katz (COL ’04)
APPENDICES

APPENDIX C:
Former Carroll Round Steering Committees

Yousif H. Mohammed (SFS ’06)
Amy M. Osekowsky (SFS ’07)
Daniel P. Schier (SFS ’05)

Fifth Annual Carroll Round
(April 27-30, 2006)
Marina Lafferriere, chair (SFS ’06)
Imrank Bademi (SFS ’06)
Stephen Brinkmann (SFS ’07)
Heber Delgado (SFS ’06)
Lucia Franzese (SFS ’07)
Yasmine Fulena (SFS ’08)
Jen Hardy (SFS ’06)
Michael Kunkel (SFS ’08)
Yousif Mohammed (SFS ’06)
Emy Reimao (SFS ’06)
Tamar Tashjian (SFS ’06)

Sixth Annual Carroll Round
(April 19-22, 2007)
Stephen A. Brinkmann, chair (SFS ’07)
Lucia Franzese (SFS ’07)
Nicholas A. Hartman (SFS ’07)
Ian P. Hindsdale (COL ’09)
Alexander P. Costura (SFS ’09)
Jennifer M. Noh (SFS ’07)
Amy M. Osekowsky (SFS ’07)
Allison E. Phillips (SFS ’07)
Sun Yi (SFS ’07)

Seventh Annual Carroll Round
(April 17-20, 2008)
Yasmine Fulena, chair (SFS ’08)
Sue Bai (SFS ’08)
Stacey Droms (COL ’08)
Brandon Feldman (COL ’08)
Li Jia Gong (SFS ’08)
Kory Katenga (SFS ’10)
Sung Kim (SFS ’08)
Michael Kunkel (SFS ’08)
Fuyang Zhang (SFS ’10)

Eighth Annual Carroll Round
(April 16-19, 2009)
Rebecca A. Heide, chair (SFS ’09)
James O. Arnold (SFS ’11)
Henry T. Gillam (SFS ’10)
Tom J. Han (SFS ’10)
Anna A. Klio (SFS ’10)
Daniel T. Leonard (SFS ’09)
Daniel Y. Lim (SFS ’11)
Arjun B. Pani (SFS ’09)
Benjamin D. Simmons (COL ’09)
Ariell E. Zimran (SFS ’10)

Ninth Annual Carroll Round
(April 22-25, 2010)
Ariell E. Zimran, chair (SFS ’10)
Michael A. Couthan (SFS ’11)
Amendia B. Delp (SFS ’12)
Katherine E. Donato (SFS ’10)
Tom J. Han (SFS ’10)
Michael R. Karmon (SFS ’10)
Allison T. Kerm (SFS ’10)
Anna A. Klio (SFS ’10)
Daniel Y. Lim (SFS ’11)
H. Jess Sook (SFS ’12)
Matthew H. Shapiro (SFS ’11)

Tenth Annual Carroll Round
(April 14-17, 2011)
Amanda Delp, chair (SFS ’12)
James Arnold (SFS ’11)
Albert Chiang (SFS ’13)
Malin Hu (SFS ’11)
Katrina Koser (SFS ’12)
Nancy Lee (SFS ’11)
Doug Proctor (SFS ’12)
Vivek Sampathkumar (SFS ’11)
Monica Scheid (SFS ’11)
Matthew Shapiro (SFS ’11)

Eighth Annual Carroll Round
(April 16-19, 2009)
Rebecca A. Heide, chair (SFS ’09)
James O. Arnold (SFS ’11)
Henry T. Gillam (SFS ’10)
Tom J. Han (SFS ’10)
Anna A. Klio (SFS ’10)
Daniel T. Leonard (SFS ’09)
Daniel Y. Lim (SFS ’11)
Arjun B. Pani (SFS ’09)
Benjamin D. Simmons (COL ’09)
Ariell E. Zimran (SFS ’10)

Ninth Annual Carroll Round
(April 22-25, 2010)
Ariell E. Zimran, chair (SFS ’10)
Michael A. Couthan (SFS ’11)
Amenda B. Delp (SFS ’12)
Katherine E. Donato (SFS ’10)
Tom J. Han (SFS ’10)
Michael R. Karmon (SFS ’10)
Allison T. Kerm (SFS ’10)
Anna A. Klio (SFS ’10)
Daniel Y. Lim (SFS ’11)
H. Jess Sook (SFS ’12)
Matthew H. Shapiro (SFS ’11)

Tenth Annual Carroll Round
(April 14-17, 2011)
Amanda Delp, chair (SFS ’12)
James Arnold (SFS ’11)
Albert Chiang (SFS ’13)
Malin Hu (SFS ’11)
Katrina Koser (SFS ’12)
Nancy Lee (SFS ’11)
Doug Proctor (SFS ’12)
Vivek Sampathkumar (SFS ’11)
Monica Scheid (SFS ’11)
Matthew Shapiro (SFS ’11)
APPENDIX C:
Former Carroll Round Steering Committees

Eleventh Annual Carroll Round
(April 19-22, 2012)
Katrina Kosser, Chair (SFS ‘12)
Albert Chiang (SFS ‘13)
Amanda Delp (SFS ‘12)
Nhaca Le (SFS ‘12)
Doug Proctor (SFS ‘12)
Glenn Russo (COL ‘13)
H. Jess Seok (SFS ‘12)
Anusuya Sivaram (SFS ‘12)
Meredith Strike (SFS ‘14)
Shuo Yan Tan (SFS ‘15)

Twelfth Annual Carroll Round
(April 18-21, 2013)
Glenn Russo, chair (COL ‘13)
Albert Chiang (SFS ‘13)
Meredith Strike (SFS ‘14)
Natalie Nah (SFS ‘15)
Brian Goggin (SFS ‘14)
Heather Hedges (SFS ‘14)
Dawn Chan (SFS ‘14)
Edward Hedke (SFS ‘13)
Elena Malik (SFS ‘14)
Stephen McDonald (SFS ‘13)
Emily Oehlsen (SFS ‘13)

Thirteenth Annual Carroll Round
(April 10-13, 2014)
Heather Hedges (SFS ‘14)
Brian Goggin (SFS ‘14)
Dawn Chan (SFS ‘14)
Elena Malik (SFS ‘14)
Natalie Nah (SFS ‘15)
Jill Ni (SFS ‘14)
Kristen Skillman (SFS ‘16)
Meredith Strike (SFS ‘14)
Christopher Stromeyer (SFS ‘14)
Rachel Szymanski (SFS ‘14)

APPENDIX D:
Members of the Advisory Panel

Meredith L. Gilbert, The Chartist Group
Christopher L. Griffin, College of William & Mary
Andrew T. Hayashi, NYU School of Law
Mitch Kaneda, Georgetown University
Robert S. Katz, Stanford University
J. Brendan Mullen, American College of Cardiology
Scott E. Pedowitz, Corporate Executive Board
Erica C. Yu, University of Maryland, College Park
APPENDIX E:
Past Participants

First Annual Carroll Round (April 5-7, 2002)

Azhar Abdul-Quader    Columbia University
Santosh Anagol     Stanford University
William Brady     Georgetown University
Daniel Braun     Oberlin College
Jacqueline Bueso     University of Pennsylvania
Kara Campbell     University of Virginia
Benn Eifert     Stanford University
Courtney Fretz     University of Pennsylvania
Carlos Galvez     Stanford University
Aniruddha Gopalakrishnan   Duke University
Christopher Griffin     Georgetown University
Casey Hanson     Lehigh University
Joshua Harris     Georgetown University
Andrew Hayashi     Georgetown University
Marco Hernandez     Massachusetts Institute of Technology
Katia Hristova     Illinois-Wesleyan University
Maria Jeleseu     Massachusetts Institute of Technology
Fadi Kanaan     Yale University
Avinash Kaza     Stanford University
Vinay Kumar     Duke University
Anisha Madan     Illinois-Wesleyan University
Kathryn Magee     Georgetown University
Ryan Michaels     Georgetown University
Jack Moore     Stanford University
Brendan Mullen     Georgetown University
Andrei Muresianu     Brown University
Scott Orleck     Duke University
Scott Pedowitz     Georgetown University
Jonathan Prin     University of Pennsylvania
Jeremy Sandford     Illinois-Wesleyan University
Deborah Slezkak     Illinois-Wesleyan University
Conan Wong     Brown University

Second Annual Carroll Round (April 11-13, 2003)

Nada Abdelnour     Georgetown University
Amanda Barnett     Emory University
Andrea Bell     Wellesley College
Patrick Byrne     University of Colorado
David Chao     Cornell University
Sylvia Ciesluk     Lehigh University
Adam Doverspike     Georgetown University

APPENDIX E:
Past Participants

First Annual Carroll Round (April 5-7, 2002)

Azhar Abdul-Quader    Columbia University
Santosh Anagol     Stanford University
William Brady     Georgetown University
Daniel Braun     Oberlin College
Jacqueline Bueso     University of Pennsylvania
Kara Campbell     University of Virginia
Benn Eifert     Stanford University
Courtney Fretz     University of Pennsylvania
Carlos Galvez     Stanford University
Aniruddha Gopalakrishnan   Duke University
Christopher Griffin     Georgetown University
Casey Hanson     Lehigh University
Joshua Harris     Georgetown University
Andrew Hayashi     Georgetown University
Marco Hernandez     Massachusetts Institute of Technology
Katia Hristova     Illinois-Wesleyan University
Maria Jeleseu     Massachusetts Institute of Technology
Fadi Kanaan     Yale University
Avinash Kaza     Stanford University
Vinay Kumar     Duke University
Anisha Madan     Illinois-Wesleyan University
Kathryn Magee     Georgetown University
Ryan Michaels     Georgetown University
Jack Moore     Stanford University
Brendan Mullen     Georgetown University
Andrei Muresianu     Brown University
Scott Orleck     Duke University
Scott Pedowitz     Georgetown University
Jonathan Prin     University of Pennsylvania
Jeremy Sandford     Illinois-Wesleyan University
Deborah Slezkak     Illinois-Wesleyan University
Conan Wong     Brown University

Second Annual Carroll Round (April 11-13, 2003)

Nada Abdelnour     Georgetown University
Amanda Barnett     Emory University
Andrea Bell     Wellesley College
Patrick Byrne     University of Colorado
David Chao     Cornell University
Sylvia Ciesluk     Lehigh University
Adam Doverspike     Georgetown University
## APPENDIX E:
### Past Participants

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benn Eifert</td>
<td>Stanford University</td>
</tr>
<tr>
<td>Adam Engberg</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Alexandra Fiorillo</td>
<td>Connecticut College</td>
</tr>
<tr>
<td>Eric Fischer</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Zlata Hajro</td>
<td>Wellesley College</td>
</tr>
<tr>
<td>Samina Jain</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Avinash Kaza</td>
<td>Stanford University</td>
</tr>
<tr>
<td>Eric Kim</td>
<td>The George Washington University</td>
</tr>
<tr>
<td>Seth Kundrot</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Lada Kyi</td>
<td>Rice University</td>
</tr>
<tr>
<td>Lee Lockwood</td>
<td>Northwestern University</td>
</tr>
<tr>
<td>Sunil Mulani</td>
<td>New York University</td>
</tr>
<tr>
<td>Holly Presley</td>
<td>Vanderbilt University</td>
</tr>
<tr>
<td>Duncan Roberts</td>
<td>University of California at Berkeley</td>
</tr>
<tr>
<td>Lu Shi</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Shanaz Taber</td>
<td>Barnard College</td>
</tr>
<tr>
<td>Jiang Wei</td>
<td>University of Michigan</td>
</tr>
</tbody>
</table>

**Third Annual Carroll Round (April 15-18, 2004)**

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jeffrey Arnold</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Julia Berzaneva</td>
<td>Mt. Holyoke College</td>
</tr>
<tr>
<td>Mehmet Cangul</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Richard Carew</td>
<td>University of Virginia</td>
</tr>
<tr>
<td>Ashley Coleman</td>
<td>Vanderbilt University</td>
</tr>
<tr>
<td>Dilyana Dimova</td>
<td>Stanford University</td>
</tr>
<tr>
<td>Fernando Galeana</td>
<td>Stanford University</td>
</tr>
<tr>
<td>M. Blair Garvey</td>
<td>Emory University</td>
</tr>
<tr>
<td>Meredith Gilbert</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Adam Greeney</td>
<td>Oberlin College</td>
</tr>
<tr>
<td>Asim Gunduz</td>
<td>University of Virginia</td>
</tr>
<tr>
<td>Marc Hafstead</td>
<td>Northwestern University</td>
</tr>
<tr>
<td>Andrew Hayashi</td>
<td>University of California at Berkeley</td>
</tr>
<tr>
<td>Katherine Howitt</td>
<td>McGill University</td>
</tr>
<tr>
<td>Sohini Kar</td>
<td>Columbia University</td>
</tr>
<tr>
<td>Josh Lewis</td>
<td>McGill University</td>
</tr>
<tr>
<td>Satish Lohani</td>
<td>Illinois-Wesleyan University</td>
</tr>
<tr>
<td>Alexis Manning</td>
<td>Illinois-Wesleyan University</td>
</tr>
<tr>
<td>Sara Menker</td>
<td>Mt. Holyoke College</td>
</tr>
<tr>
<td>Elizabeth Mielke</td>
<td>Vanderbilt University</td>
</tr>
<tr>
<td>Stratos Pahis</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Alicia Pluta</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Adam Raymakers</td>
<td>Dalhousie University</td>
</tr>
<tr>
<td>Caroline Schmutte</td>
<td>Dartmouth College</td>
</tr>
</tbody>
</table>

280
APPENDIX E:  
Past Participants

Matt Sekerke     Johns Hopkins University  
John Soleanico     Columbia University  
Kai Szakmary     Columbia University  
Brandon Wall     Yale University  
Kenneth Ward     University of Chicago  
Susan Work     Georgetown University  

Fourth Annual Carroll Round (April 22-24, 2005)

Lidia Barabash     Dartmouth College  
Jasmina Beganovic    Georgetown University  
Xun Bian      Illinois-Wesleyan University  
Michael Furchtgott    Columbia University  
Michael Gechter     Pomona College  
Kevin B. Goldstein    Dartmouth College  
Michael Haase     University of Copenhagen  
Dennis Huggins     Georgetown University  
Michael Insel     Claremont McKenna College  
Jonathan Kirschner    Georgetown University  
Shiying Lee     Duke University  
James Liao     Dartmouth College  
Brian Lichten     Washington University  
Wen Loh     Cornell University  
Alice Luo     Duke University  
Katharine Mullock    University of Western Ontario  
Jose Muatre del Rio     The Ohio State University  
Leah Nelson     Georgetown University  
Ee Cheng Ong     Wellesley College  
Matthew Phan     Columbia University  
Nina Rendelstein     Washington University  
David Rogier     Washington University  
Ana Maria Romero     Illinois-Wesleyan University  
Nathan Saperia     Dartmouth College  
Bogdan Terehenko    Georgetown University  
Olga Timoshenko    University of Western Ontario  
Tom Vogl     Princeton University  
Kenneth Ward     University of Chicago  
Jonathan Wolfson     Washington University  
Suzanne Zurkiya     Emory University  

Fifth Annual Carroll Round (April 28-30, 2006)

Sarah Carroll     Stanford University  
Ruth Coffman     Georgetown University  

APPENDIX E:
Past Participants

Matt Sekerke     Johns Hopkins University  
John Soleanico     Columbia University  
Kai Szakmary     Columbia University  
Brandon Wall     Yale University  
Kenneth Ward     University of Chicago  
Susan Work     Georgetown University  

Fourth Annual Carroll Round (April 22-24, 2005)

Lidia Barabash     Dartmouth College  
Jasmina Beganovic    Georgetown University  
Xun Bian      Illinois-Wesleyan University  
Michael Furchtgott    Columbia University  
Michael Gechter     Pomona College  
Kevin B. Goldstein    Dartmouth College  
Michael Haase     University of Copenhagen  
Dennis Huggins     Georgetown University  
Michael Insel     Claremont McKenna College  
Jonathan Kirschner    Georgetown University  
Shiying Lee     Duke University  
James Liao     Dartmouth College  
Brian Lichten     Washington University  
Wen Loh     Cornell University  
Alice Luo     Duke University  
Katharine Mullock    University of Western Ontario  
Jose Muatre del Rio     The Ohio State University  
Leah Nelson     Georgetown University  
Ee Cheng Ong     Wellesley College  
Matthew Phan     Columbia University  
Nina Rendelstein     Washington University  
David Rogier     Washington University  
Ana Maria Romero     Illinois-Wesleyan University  
Nathan Saperia     Dartmouth College  
Bogdan Terehenko    Georgetown University  
Olga Timoshenko    University of Western Ontario  
Tom Vogl     Princeton University  
Kenneth Ward     University of Chicago  
Jonathan Wolfson     Washington University  
Suzanne Zurkiya     Emory University  

Fifth Annual Carroll Round (April 28-30, 2006)

Sarah Carroll     Stanford University  
Ruth Coffman     Georgetown University  

281
**APPENDIX E:**

**Past Participants**

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dubravka Colic</td>
<td>Wellesley College</td>
</tr>
<tr>
<td>Pratik Dattani</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Jennifer Dawson</td>
<td>Illinois-Wesley University</td>
</tr>
<tr>
<td>Héber Delgado-Medrano</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Sherri Haas</td>
<td>Illinois-Wesley University</td>
</tr>
<tr>
<td>Jen Hardy</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Lauren Iacocca</td>
<td>University of California at Los Angeles</td>
</tr>
<tr>
<td>Salifou Issoufou</td>
<td>University of Wisconsin at Madison</td>
</tr>
<tr>
<td>Stella Klemperer</td>
<td>Brown University</td>
</tr>
<tr>
<td>Daniel Kurland</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Corinne Low</td>
<td>Duke University</td>
</tr>
<tr>
<td>Shanthi Manian</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Michael Monteleone</td>
<td>University of Chicago</td>
</tr>
<tr>
<td>John Nesbitt</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Natasha Nguyen</td>
<td>University of California at Berkeley</td>
</tr>
<tr>
<td>Oyebanke Oyeinika</td>
<td>Carleton College</td>
</tr>
<tr>
<td>Evgeniya Petrova</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Emy Reimao</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Svetoslav Roussanov</td>
<td>Columbia University</td>
</tr>
<tr>
<td>Vikram Shankar</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Juan Carlos Suarez</td>
<td>Trinity University</td>
</tr>
<tr>
<td>Austin Vedder</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>David Wiczer</td>
<td>Carleton College</td>
</tr>
<tr>
<td>Geoffrey Yu</td>
<td>Carleton College</td>
</tr>
<tr>
<td>Xiaoti Zhang</td>
<td>University of Warwick</td>
</tr>
</tbody>
</table>

Sixth Annual Carroll Round (April 19-22, 2007)

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matthew Adler</td>
<td>Oberlin College</td>
</tr>
<tr>
<td>Marion Aviad</td>
<td>Princeton University</td>
</tr>
<tr>
<td>Stephen Brinkmann</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Erik Eggum</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Lucia Franzese</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Tanja Groth</td>
<td>University of St. Andrews</td>
</tr>
<tr>
<td>Ashley Halpin</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Nicholas Hartman</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Adrienna Huffman</td>
<td>Washington University</td>
</tr>
<tr>
<td>Abdulla Humaidan</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Mohammad Huq</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Nedko Kysachukov</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Zachary Mahone</td>
<td>New York University</td>
</tr>
<tr>
<td>R. Priya Mathew</td>
<td>Washington University</td>
</tr>
<tr>
<td>Yana Morgulis</td>
<td>University of Chicago</td>
</tr>
<tr>
<td>Jennifer Noh</td>
<td>Georgetown University</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matthew Adler</td>
<td>Oberlin College</td>
</tr>
<tr>
<td>Marion Aviad</td>
<td>Princeton University</td>
</tr>
<tr>
<td>Stephen Brinkmann</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Erik Eggum</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Lucia Franzese</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Tanja Groth</td>
<td>University of St. Andrews</td>
</tr>
<tr>
<td>Ashley Halpin</td>
<td>Nicholas Hartman</td>
</tr>
<tr>
<td>Adrienna Huffman</td>
<td>Abdulla Humaidan</td>
</tr>
<tr>
<td>Mohammad Huq</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Nedko Kysachukov</td>
<td>University of Chicago</td>
</tr>
<tr>
<td>Zachary Mahone</td>
<td>R. Priya Mathew</td>
</tr>
<tr>
<td>Yana Morgulis</td>
<td>Jennifer Noh</td>
</tr>
</tbody>
</table>

282  282
### APPENDIX E: Past Participants

<table>
<thead>
<tr>
<th>Name</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrew O’Brien Penney</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Jessica Oliveri</td>
<td>Monash University</td>
</tr>
<tr>
<td>Matthew Pech</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Allison Phillips</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Angelica da Rocha</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Sören Radde</td>
<td>University of Bayreuth</td>
</tr>
<tr>
<td>Heleri Rande</td>
<td>New York University</td>
</tr>
<tr>
<td>Elena Spatoulas</td>
<td>University of Michigan</td>
</tr>
<tr>
<td>Yi Sun</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Bennett Surajat</td>
<td>Carleton College</td>
</tr>
<tr>
<td>Freddy Tsai</td>
<td>University of British Columbia</td>
</tr>
<tr>
<td>David Wolff</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Jennifer Xi</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Cynthia Yim</td>
<td>Princeton University</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Karl Andres</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Cecil Ang</td>
<td>University of Virginia</td>
</tr>
<tr>
<td>Alaina Antonucci</td>
<td>The Pennsylvania State University</td>
</tr>
<tr>
<td>Sue Bai</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Marinella Boyadzhiev</td>
<td>Oberlin College</td>
</tr>
<tr>
<td>Quentin Brunnet</td>
<td>Illinois-Wesleyan University</td>
</tr>
<tr>
<td>Brendan Cooper</td>
<td>Carleton College</td>
</tr>
<tr>
<td>Gerard DiPippo</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Stacey Drums</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Varun Dutt</td>
<td>Macalaster College</td>
</tr>
<tr>
<td>Yasmine Fulena</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Amish Gandhi</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Katherine Gordon</td>
<td>Mt. Holyoke College</td>
</tr>
<tr>
<td>Yi Kang</td>
<td>Wesleyan College</td>
</tr>
<tr>
<td>Michael Kunkel</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Han Youp Lee</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Claudio LoCascio</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Olivia Lynch</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Amr Moubarak</td>
<td>The George Washington University</td>
</tr>
<tr>
<td>Simone Nitsch</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Saurabh Pant</td>
<td>New York University</td>
</tr>
<tr>
<td>Carson Sherwood</td>
<td>University of Western Ontario</td>
</tr>
<tr>
<td>Tadashi Shirai</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Dominique Shure</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>William Slater</td>
<td>Vanderbilt University</td>
</tr>
<tr>
<td>Shyam Sundaram</td>
<td>Brown University</td>
</tr>
<tr>
<td>Poh Lin Tan</td>
<td>Princeton University</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Karl Andres</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Cecil Ang</td>
<td>University of Virginia</td>
</tr>
<tr>
<td>Alaina Antonucci</td>
<td>The Pennsylvania State University</td>
</tr>
<tr>
<td>Sue Bai</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Marinella Boyadzhiev</td>
<td>Oberlin College</td>
</tr>
<tr>
<td>Quentin Brunnet</td>
<td>Illinois-Wesleyan University</td>
</tr>
<tr>
<td>Brendan Cooper</td>
<td>Carleton College</td>
</tr>
<tr>
<td>Gerard DiPippo</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Stacey Drums</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Varun Dutt</td>
<td>Macalaster College</td>
</tr>
<tr>
<td>Yasmine Fulena</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Amish Gandhi</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Katherine Gordon</td>
<td>Mt. Holyoke College</td>
</tr>
<tr>
<td>Yi Kang</td>
<td>Wesleyan College</td>
</tr>
<tr>
<td>Michael Kunkel</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Han Youp Lee</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Claudio LoCascio</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Olivia Lynch</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Amr Moubarak</td>
<td>The George Washington University</td>
</tr>
<tr>
<td>Simone Nitsch</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Saurabh Pant</td>
<td>New York University</td>
</tr>
<tr>
<td>Carson Sherwood</td>
<td>University of Western Ontario</td>
</tr>
<tr>
<td>Tadashi Shirai</td>
<td>University of Warwick</td>
</tr>
<tr>
<td>Dominique Shure</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>William Slater</td>
<td>Vanderbilt University</td>
</tr>
<tr>
<td>Shyam Sundaram</td>
<td>Brown University</td>
</tr>
<tr>
<td>Poh Lin Tan</td>
<td>Princeton University</td>
</tr>
</tbody>
</table>
APPENDICES

APPENDIX E: Past Participants

Anders Jensen     London School of Economics and Political Science
William Kafoure     The George Washington University
Elira Kuka     Wellesley College
Daniel Leonard     Georgetown University
Chris Lim     Dartmouth College
Juan Ignacio Elorrieta Maira     University of Chile
Nick Marchio     Macalester College
Eliran Nezirit     American University in Bulgaria
Casey Oswald     Georgetown University
Arjun Pant     Georgetown University
Carlin Pierce     Dartmouth College
Isha Salim     Macalester College
Keval Sangani     University of Warwick
Pranita Saxena     University of California, Berkeley
Benjamin Simmons     Georgetown University
Maximilian Sarianni     Macalester College
Seitaro Takarabe     Wesleyan University
Fabien Thayamballi     Georgetown University
Rachel Winograd     Dartmouth College
Woan Foong Wong     Oberlin College

Ninth Annual Carroll Round (April 22-25, 2010)

Jorge Aponte     Georgetown University
Benjamin Arnold     University of Michigan
Courtney Blair     Harvard University
Vera Chau     New York University
Nick Chantraporn     University of San Francisco
Antonina Davydenko     American University in Bulgaria
Katherine Donato     Georgetown University
Yang Du     Dartmouth College
Siddharth George     London School of Economics and Political Science
Takuma Habu     University of Warwick
Kelsey Hample     Illinois Wesleyan University
Tom Han     Georgetown University
Rob Harris     University of Warwick
Sarah Hinkfuss     Harvard University
Peter Hall     Wesleyan University
Michael Karna     Georgetown University
Tod Kawakita     Dartmouth College
Allison Kern     Georgetown University
Anna Klis     Georgetown University
Birgit Leimer     New York University
Daniel Lim     Georgetown University

APPENDICES

APPENDIX E: Past Participants

Anders Jensen     London School of Economics and Political Science
William Kafoure     The George Washington University
Elira Kuka     Wellesley College
Daniel Leonard     Georgetown University
Chris Lim     Dartmouth College
Juan Ignacio Elorrieta Maira     University of Chile
Nick Marchio     Macalester College
Eliran Nezirit     American University in Bulgaria
Casey Oswald     Georgetown University
Arjun Pant     Georgetown University
Carlin Pierce     Dartmouth College
Isha Salim     Macalester College
Keval Sangani     University of Warwick
Pranita Saxena     University of California, Berkeley
Benjamin Simmons     Georgetown University
Maximilian Sarianni     Macalester College
Seitaro Takarabe     Wesleyan University
Fabien Thayamballi     Georgetown University
Rachel Winograd     Dartmouth College
Woan Foong Wong     Oberlin College

Ninth Annual Carroll Round (April 22-25, 2010)

Jorge Aponte     Georgetown University
Benjamin Arnold     University of Michigan
Courtney Blair     Harvard University
Vera Chau     New York University
Nick Chantraporn     University of San Francisco
Antonina Davydenko     American University in Bulgaria
Katherine Donato     Georgetown University
Yang Du     Dartmouth College
Siddharth George     London School of Economics and Political Science
Takuma Habu     University of Warwick
Kelsey Hample     Illinois Wesleyan University
Tom Han     Georgetown University
Rob Harris     University of Warwick
Sarah Hinkfuss     Harvard University
Peter Hall     Wesleyan University
Michael Karna     Georgetown University
Tod Kawakita     Dartmouth College
Allison Kern     Georgetown University
Anna Klis     Georgetown University
Birgit Leimer     New York University
Daniel Lim     Georgetown University
APPENDIX E: Past Participants

Tenth Annual Carroll Round (April 14-17, 2011)

Ahmad Wahdat    Oberlin College
Ariell Zimran    Georgetown University
Dimitri Avramov    American University in Bulgaria
Daniel Boada    Harvard College
Gustavo Camilo    New York University
Daniel Chan    United States Naval Academy
Meryl Ching    University of Warwick
Kimberly Conlon    University of Minnesota
Tess DeLean    Wellesley College
Max Gelb    Dartmouth College
Ben Guttmann-Kenney    University of Warwick
Malin Hu    Georgetown University
Kilian Huber    London School of Economics and Political Science
Tomas Jagelka    Dartmouth College
Shorena Kalanadarishvili    Smith College
Hitode Koizumi    Soka University of America
Krisjanis Krustins    Stockholm School of Economics in Riga
Benjamin Langworthy    Macalester College
Nancy Lee    Georgetown University
Daniel Lim    Georgetown University
Van Nguyen    Washington and Lee University
Nikita Orlov    University of Warwick
Anselm Rink    London School of Economics and Political Science
Vivek Sampathkumar    Georgetown University
Monica Scheid    Georgetown University
Markus Schwedeler    Maastricht University
Matthew Shapiro    Georgetown University
Zane Silina    Stockholm School of Economics in Riga
Anusuya Sivaram    Georgetown University
David Thomas    University of Oxford
Maximilian JC Thormann    London School of Economics and Political Science

Eleventh Annual Carroll Round (April 19-22, 2012)

Mudara Bogdane    Stockholm School of Economics at Riga
Paul Byatta    Harvard College
Nikhil Dugal    New York University
Vladimir Epuri    American University in Bulgaria
Samuel Evans    University of Warwick
Evan Friedman    Brown University
Fabian Gunzinger    University of Bern

APPENDIX E: Past Participants

Tenth Annual Carroll Round (April 14-17, 2011)

Ahmad Wahdat    Oberlin College
Ariell Zimran    Georgetown University
Dimitri Avramov    American University in Bulgaria
Daniel Boada    Harvard College
Gustavo Camilo    New York University
Daniel Chan    United States Naval Academy
Meryl Ching    University of Warwick
Kimberly Conlon    University of Minnesota
Tess DeLean    Wellesley College
Max Gelb    Dartmouth College
Ben Guttmann-Kenney    University of Warwick
Malin Hu    Georgetown University
Kilian Huber    London School of Economics and Political Science
Tomas Jagelka    Dartmouth College
Shorena Kalanadarishvili    Smith College
Hitode Koizumi    Soka University of America
Krisjanis Krustins    Stockholm School of Economics in Riga
Benjamin Langworthy    Macalester College
Nancy Lee    Georgetown University
Daniel Lim    Georgetown University
Van Nguyen    Washington and Lee University
Nikita Orlov    University of Warwick
Anselm Rink    London School of Economics and Political Science
Vivek Sampathkumar    Georgetown University
Monica Scheid    Georgetown University
Markus Schwedeler    Maastricht University
Matthew Shapiro    Georgetown University
Zane Silina    Stockholm School of Economics in Riga
Anusuya Sivaram    Georgetown University
David Thomas    University of Oxford
Maximilian JC Thormann    London School of Economics and Political Science

Eleventh Annual Carroll Round (April 19-22, 2012)

Mudara Bogdane    Stockholm School of Economics at Riga
Paul Byatta    Harvard College
Nikhil Dugal    New York University
Vladimir Epuri    American University in Bulgaria
Samuel Evans    University of Warwick
Evan Friedman    Brown University
Fabian Gunzinger    University of Bern
APPENDIX E: Past Participants

Twelfth Annual Carroll Round (April 18-21, 2013)

Nikola Andreev    American University in Bulgaria
Matthew Bailey    University of Warwick
Albert Chiang    Georgetown University
Bayarkhuu Chinzorigt    American University in Bulgaria
Hadi Elzayn    Columbia University
Yi Jie Gwee    London School of Economics
Rosa Hayes    Wesleyan University
Asher Hecht-Bernstein    Columbia University
Edward Hedke    Georgetown University
Hannah Hill    Georgetown University
Sasha Indarte    Macalester College
Mohandass Kalaichelvan    Dartmouth College
Phoebe Kotlikoff    United States Naval Academy
Weiwen Leung    Singapore Management University
Shawn Lim    University College London
Michael Lopescoilo    Georgetown University
Sara Marcus    Dartmouth College
Stephen McDonald    Georgetown University
Leyla Mocan    University of Pennsylvania
Preston Mui    Georgetown University

APPENDIX E: Past Participants

Nikola Andreev    American University in Bulgaria
Matthew Bailey    University of Warwick
Albert Chiang    Georgetown University
Bayarkhuu Chinzorigt    American University in Bulgaria
Hadi Elzayn    Columbia University
Yi Jie Gwee    London School of Economics
Rosa Hayes    Wesleyan University
Asher Hecht-Bernstein    Columbia University
Edward Hedke    Georgetown University
Hannah Hill    Georgetown University
Sasha Indarte    Macalester College
Mohandass Kalaichelvan    Dartmouth College
Phoebe Kotlikoff    United States Naval Academy
Weiwen Leung    Singapore Management University
Shawn Lim    University College London
Michael Lopescoilo    Georgetown University
Sara Marcus    Dartmouth College
Stephen McDonald    Georgetown University
Leyla Mocan    University of Pennsylvania
Preston Mui    Georgetown University
APPENDIX E:
Past Participants

Emily Oelsen    Georgetown University
Igors Pasuks    Stockholm School of Economics in Riga
Nicolas Powidayko    University of Brasilia
Michael Reher    Georgetown University
Glenn Russo    Georgetown University
Eduards Sidorovics    Stockholm School of Economics in Riga
Fabian Trotter    London School of Economics
Ilyas Zhukov    University of Warwick

Thirteenth Annual Carroll Round (April 10-13, 2014)
Eric Aldenhoff    University of Maryland, College Park
Russell Black    Oxford University
Thomas Bumberger    University of Cambridge
Dawn Chan    Georgetown University
Kyle Coombs    Macalester College
Rob Dent    University of Virginia
Brian Goggin    Georgetown University
Heather Hedges    Georgetown University
Alyssa Huberts    Georgetown University
Johnny Hynd    Pomona College
Nikhil Kalathil    Oberlin College
Matthew Klein    University of Chicago
Samsun Knight    Oberlin College
Vincent La    Dartmouth College
Nataliya Langburd    Yale University
Katherine Loosley    London School of Economics
Soumyajit Mazumder    Georgetown University
Russell Morton    Princeton University
Jill Ni    Georgetown University
Jonathan Pedde    Dartmouth College
Andrei Romanecenco    Stockholm School of Economics in Riga
Saugata Sen    London School of Economics
Benjamin Shoemith    University of North Carolina Wilmington
Meredith Strike    Georgetown University
Chris Stromeyer    Georgetown University
Rachel Syzmanski    Georgetown University
Josh Walker    Lancaster University