The Carroll Round at Georgetown University

Box 571016 Washington, DC 20057

Tel: (202) 687-5696
Fax: (202) 687-1431
carrollround@georgetown.edu

http://carrollround.georgetown.edu
CARROLL ROUND PROCEEDINGS

The Sixteenth Annual Carroll Round
An Undergraduate Conference Focusing on Contemporary International Economics Research and Policy

Editors in Chief:
Arjun Krishnan, Jacob Witt

Associate Editors:
Olivia Bisel, Shine Aung, Grady Killeen, Arjun Krishnan, Joshua Levy, Duy Mai, Parker Malarkey, Meggie Underwood, Elinor Walker, Crystal Zhu

A Carroll Round Publication
The Sixteenth and Seventeenth Carroll Round Steering Committees
Georgetown University, Washington, DC
2018
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the Carroll Round?</td>
<td>iv</td>
</tr>
<tr>
<td>Notes on Paper and Conference Participation</td>
<td>iv</td>
</tr>
<tr>
<td>Notes on Published Papers</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>v</td>
</tr>
<tr>
<td>A Brief History of the Carroll Round</td>
<td>vi</td>
</tr>
<tr>
<td>Why I Support the Carroll Round</td>
<td>ix</td>
</tr>
<tr>
<td><strong>Carroll Round Proceedings</strong></td>
<td></td>
</tr>
<tr>
<td>(*Denotes Abstract Only)</td>
<td></td>
</tr>
<tr>
<td>Trade creation and Diversion in the U.S. - Colombia Trade Promotion Agreement:</td>
<td>1</td>
</tr>
<tr>
<td>A Preliminary Analysis</td>
<td></td>
</tr>
<tr>
<td>AMY YANG</td>
<td></td>
</tr>
<tr>
<td><em>Dartmouth College</em></td>
<td></td>
</tr>
<tr>
<td>A Performance Analysis of Managed Futures</td>
<td>11</td>
</tr>
<tr>
<td>CHRISTIAN FERNANDO</td>
<td></td>
</tr>
<tr>
<td><em>Georgetown University</em></td>
<td></td>
</tr>
<tr>
<td>What are the Effects of Specific Trade Agreement Provisions on Bilateral Trade Flows?*</td>
<td>33</td>
</tr>
<tr>
<td>DUNCAN HOBBS</td>
<td></td>
</tr>
<tr>
<td><em>Georgetown University</em></td>
<td></td>
</tr>
<tr>
<td>The Cost of Conflict: Refugee Flows in Sub-Saharan Africa</td>
<td>34</td>
</tr>
<tr>
<td>GABRIELLE COHEN</td>
<td></td>
</tr>
<tr>
<td><em>University of Cape Town</em></td>
<td></td>
</tr>
<tr>
<td>The Effect of Scandal on the Value of Political Connections in South Korean Chaebol</td>
<td>46</td>
</tr>
<tr>
<td>GRACE KIM</td>
<td></td>
</tr>
<tr>
<td><em>Georgetown University</em></td>
<td></td>
</tr>
<tr>
<td>Attribute-Based Regulations: The Case of Corporate Average Fuel Economy (CAFE) Standards*</td>
<td>65</td>
</tr>
<tr>
<td>GRADY KILLEEN</td>
<td></td>
</tr>
<tr>
<td><em>Georgetown University</em></td>
<td></td>
</tr>
<tr>
<td>Transparency and the Effect of Natural Resources on Economic Growth*</td>
<td>66</td>
</tr>
<tr>
<td>OLIVIA BISEL, YUOU WU</td>
<td></td>
</tr>
<tr>
<td><em>Georgetown University</em></td>
<td></td>
</tr>
</tbody>
</table>
Power and Petropolitics: How Oil Prices Affect Clout
GRiffin COHEN
Georgetown University

The Economic Efficacy of Reintegration Assistance for Former Child Soldiers
JONATHAN KAUFMANN
American University

Why Uncertainty Matters: An Empirical Approach for Spain
ROBERT PÉREZ GARCIA, ALEXANDRE DÍAZ CODINA, ADAM LAAZIZ
ALAOUI DESCARREGA, VICTOR M. QUINTAS MARTÍNEZ, ADRIA ROCA VALLS
Universitat Pompeu Fabra

Why don't countries sue more often at the WTO? An Analysis of Specific Trade Concerns in the SPS Committee*
MATHISON S. CLORE
Georgetown University

Do Language and Cultural Barriers Enable Firms to Price Discriminate Between Countries in the European Union?
TIM MAECKER
University of Warwick

Exploring conflict in Nigeria: Does the rapacity effect dominate across local government areas in Nigeria?*
DAVID HENNING
University of Warwick

Pharmaceutical Royalty Licensing Transaction Impacts on Shareholder Wealth: An Event Study
ALEXANDRA MALINDI COLYER
Georgetown University

Financial Fragility and Contagion in Emerging Economies: A Panel Analysis
STEFANI KOSTADINOVA
American University in Bulgaria

Working Women: The Impact of Financial Crises on Female Labor Force Participation Rates in Developing Countries
ALYSON MATTHEWS
Georgetown University

Voting for Credibility: The Political Economy of Sovereign Credit Ratings
TY GREENBERG*
Georgetown University
MARCEL SCHLEPPER  
*University of Warwick*

Country-Level Crime Rates and Sanctuary Jurisdictional Status  
THOMAS M. SEGERSTROM  
*Gettysburg College*

Relative Deprivation, Rural-Urban Migration and Rural Inequality in China  
MOHAMMAD TAIMUR ALI AHMAD  
*Georgetown University Qatar*

Bank Consolidation and Systemic Risk: M&A During the 2008 Financial Crisis*  
GREGORY D. MASLAK, GONCA SENEL  
*Bowdoin College*

An Application of the Phillips Curve to the Indian Economy  
SHAMBHAVI TIWARI  
*Columbia University*

Goodbye Lenin, Hello Murat? The Effect of Communism on Individual Attitudes Toward Immigration*  
MATTHEW CARL  
*Washington & Lee University*

On the Effects of Intellectual Property Rights on Foreign Direct Investments in Developing Economies: A System GMM Approach  
VINCENT JERALD R. RAMOS  
*University of the Philippines*

Appendix  
iii
What is the Carroll Round?
The Carroll Round is an international economics conference for undergraduate students held each spring 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 500 students from universities and colleges in North America, Western and Eastern Europe, Asia, the Middle East, South America, 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 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 economics are encouraged to apply. Special preference is granted to papers focusing on international issues.

Notes on Published Papers
Many of the papers published in this journal have been shortened due to length requirements. In most cases, changes were minor. However, some essays were significantly abridged due to these constraints.
**Acknowledgments**

Organizing an economics conference as an undergraduate student is no small feat. The Carroll Round has prevailed, in part, thanks to the work of previous committees. Inspired by the founding mission to expand the frontiers of knowledge and policy applications through academic discourse, each successive committee harnessed an astonishing legacy in building the conference each year. At the core of this legacy, of course, is the numerous individuals and organizations without whom the Carroll Round would not be possible.

The Carroll Round would like to acknowledge special individuals who have cared deeply about our cause. Alumna Marianne Keler and her spouse Michael Kershow have graced us with their support and presence every year. The Carroll Round Endowed Program Fund that they created for us provides us with a perpetual income stream to support the annual conference. We thank Ms. Keler and Mr. Kershow deeply for having been such an important part of the Carroll Round over all these years.

Our deepest gratitude will forever go to alumnus Yunho Song, who has personally supported the Carroll Round from its very first year. The first committee had the privilege of sitting down with Mr. Song at the Tombs to convey our dreams. Mr. Song designated for us his endowment fund, which now supports the Carroll Round year after year.

Among the Carroll Round alumni, Mr. Scott Pedowitz has provided tremendous guardianship and support. Mr. Pedowitz was a member of the founding committee, and the Carroll Round was able to gain his support for each of the sixteen years since his graduation. Mr. Pedowitz meets with the committee members every year to gain updates and to share the vision that has continued since the founding year.

The Carroll Round would have not been possible without the support of many other individuals. We would like to recognize Mr. Mario Espinosa, Mr. Oleg Nodelman, Ms. Colleen Murphy, Mr. and Mrs. Kenneth Kunkel, Ms. Sarah Osborne, Mr. Jonathan Prin, Mr. Jon Skillman and Ms. Luanne Selk, Mr. Geoffrey Yu, and former Carroll Round Steering Committee members Mr. James Arnold, Ms. Meredith Ballotta, Mr. Stephen Brinkmann, Ms. Amanda Delp, Ms. Stacey Droms, Mr. Brandon Feldman, Ms. Yasmine Fulena, Mr. Christopher Griffin, Dr. Andrew Hayashi, Ms. Rebecca Heide, Dr. Anna Klis, Mr. Michael Kunkel, Ms. Marina Linhart, Ms. Nancy Lee, Mr. Shuo Tan, Dr. Erica Yu Wright, and Dr. Ariell Zimran.

In addition, the Sallie Mae Corporation significantly funded the first five conferences, and we are most grateful for their foresight in supporting our conception and our growth into an established undergraduate research conference. Moreover, 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 Carroll Round was helped by past and present members of the advancement office: Mr. Mohamed Abdel-Kader, Mr. Thomas Esch, Ms. Carma Fauntleroy, Ms. Elizabeth Franzino, Ms. Reema Ghazi, Ms. Gail Griffith, Mr. Richard Jacobs, Dr. Venilde Jeronimo, Ms. Katerina Kulagina, Ms. Christine Smith, and Ms. Cara Sodos. The Carroll Round is presently pursuing an institutional research grant, and the guidance of Mr. Clark Bonilla has been invaluable. We would also like to provide special recognition to all the former steering committee members, beyond those already mentioned, who have contributed very generously to help the Carroll Round. Among them, we would especially like to thank Ms. Sue Bai, Mr. Albert Chiang, Ms. Daphney Francois, Mr. Edward Hedke, Mr. Dennis Huggins, Ms. Cindy Jin, Mr. Michael Karno, Mr. Dan Leonard, Mr. Jonathan McClure, Mr. Brendan Mullen, Dr. Emy Reimao, and Ms. Kristen Skillman.

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. We deeply thank each of the successive deans of the School of Foreign Service: Robert Gallucci, Carol Lancaster, Jim Reardon-Anderson, and Joel Hellman. Administratively, the Carroll Round was helped by SFS Dean’s Office members Dean Kendra Billingslea, Ms. Denisse Bonilla-Chaoui, Mr. Beau Boughamer, Ms. Rebecca Ernest, Mr. Franz Hartl, Dr. Dan Powers, Mr. Michael Volk, and Mr. Benjamin Zimmerman.

The Carroll Round has been fortunate for the last seventeen years to enjoy the substantive quality of the economics undergraduates from across the world. We are particularly grateful to those professors that steer their best students to the Carroll Round, especially Professor Nancy Marion of Dartmouth College, Professor Judith Shapiro of the London School of Economics, Professor Michael Seeborg of Illinois Wesleyan University, Dr. Gianna Boero of Warwick University, and Professor Ian Walker of Lancaster University.

We receive the professional experience and wisdom of some of the most respected economists in the field. For the Sixteenth Annual Carroll Round, we were particularly fortunate to have keynote lectures from Dr. Jason Furman of Harvard University and Dr. Nobuhiro Kiayotaki of Princeton University. Also critical to the substantive development of the Carroll Round and our participants’ work are the session chairs who take the time to read participants’ papers and critique their presentations at the conference. We would like to thank the 2017 session chairs for their contributions to the conference: Professors Dan Cao (Georgetown), Robert Cumby (Georgetown), Christopher Griffin (Harvard), Marko Klasnja (Georgetown), Anna Klis (Northern Illinois), Anna Maria Mayda (Georgetown), Olga Timoshenko (George Washington), and Charles Udomsaph (Georgetown).

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 its 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 March 2018)

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 steering 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 each year and of course the global contingent that gathers in 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 perpetual need to look ahead, reviewing its origins is equally important. During the 2001-2002 school 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 always has been and can be.

The conference’s birthplace, as many know by now, was an Oxford pub called the Radcliffe Arms. Even though that fact 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 the next year.

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 a 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 and watering hole, 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 mature 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 after returning to Georgetown.

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 group, 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 after returning to Georgetown.

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 Pembrokeians 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 preclude 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 the first two decades years have advanced to graduate study at Berkeley, Chicago, Cornell, Duke, Harvard, 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, law, and public policy faculties. 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 possible the annual publication of these Proceedings. 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 successor Chairs from Seth Kundrot in 2003 to Olivia Bisel in 2018. 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 several years 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
Why I Support the Carroll Round

“Ideas are more powerful than guns.” - Joseph Stalin

You are probably asking what Stalin – and his idealistic, bloody, and brutal rule in Russia – has to do with the Carroll Round. Good question – the belief that ideas are the most powerful force at work in our world was the impetus behind the Carroll Round. It is a belief that has been closely held by philosophers, politicians, artists, businessmen, dictators, and academics across the world. That’s why I support the Carroll Round.

The Carroll Round was founded to “foster the exchange of ideas” among students of international economics and political economy around the world and with leading minds in the field. Why does that matter? Universities are already centers of discussion and innovation. They are built to encourage learning, debate, creativity, and expression. But individual universities, especially within a particular discipline, often cultivate complementary perspectives and can, unintentionally, train students in a school of thought. At the University of Chicago, for example, the school of economics is associated with the neoclassical school of thought popularized by several of its distinguished faculty. Top schools across the U.S. divided along different approaches in macroeconomics (saltwater vs. freshwater) with philosophical differences on the role of rational behavior, fiscal, and discretionary policy. As a student of economics, the school where you study implies an orientation – at least perceived – even if not truly adopted. By bringing together the brightest undergraduate economists from all of these institutions we test those philosophies, break out of the expected molds, and shape the type of brilliant, sharp, world-changing ideas that can only be honed through smart, challenging, and contentious debates.

From the time – more years ago than I care to count – when I served as Chair of the Carroll Round to today, the conference has exceeded my expectations and imagination as a place to share, hone, and challenge ideas. The Carroll Round has built bridges and collaborations among the student participants. The Carroll Round has challenged students by debate, dinner, and discussion with leading economists in the field. I remember so very clearly the presentation by John Nash on an old overhead projector with typed transparent slides that I struggled to comprehend even deep in my own economic coursework. Or that thread of game theory that was picked up two years later with Thomas Schelling whose application to conflict resolution has continued to touch my career in arenas well outside of economics. The Carroll Round has produced ideas – great ideas. Just by browsing through the proceedings this year – as I’m sure you are anxious to dive into – you can appreciate the variety, depth, and thoughtfulness of the upcoming generation of economists tackling urgent societal issues.

So I won’t keep you any longer except to say that ideas are powerful. And the Carroll Round continues to be one of the most incredible exchanges of ideas for undergraduates in economics in the U.S. As you read the papers and abstracts housed here, imagine the questions that discussants posed to their peers, the crazy ideas sparked over dinner and drinks, and how excited the authors must have been to ask questions and just chat with renowned speakers this year. That is what makes the Carroll Round so powerful. And that is why I continue to support them and encourage you to do the same.

Marina Linhart
CEO at Next Street
Carroll Round Proceedings

The Sixteenth Annual Carroll Round Undergraduate Research Conference
Trade Creation and Diversion in the
U.S. - Colombia Trade Promotion Agreement:
A Preliminary Analysis

AMY YANG
Dartmouth College

ABSTRACT
This paper provides a preliminary analysis of the changes in Colombian trade patterns following the 2012 U.S.-Colombia Trade Promotion Agreement (TPA). Based on Colombian import data available for over 6,800 product categories from 2011 through 2014, I investigate the extent to which the TPA has resulted in trade creation and/or diversion during the first two years of its implementation. Variation in the extent of tariff reductions across different product groups is used to measure the impact of regional trade liberalization on import flows. The results indicate that the TPA led to trade creation effects during the initial two years, with no evidence of trade diversion.

1 I would like to thank Professor Nancy Marion for providing insightful guidance throughout the formulation of both this paper and the subsequent thesis that came out of this research. I would also like to thank Professor Andrew Bernard for generously offering a rich set of Colombian import data with product-level details.
1. Introduction
On May 15, 2012, the United States and Colombia implemented a bilateral Trade Promotion Agreement (TPA) that eliminated tariffs on 80% of goods exported from the U.S. to Colombia. The TPA is one of 14 free trade agreements (FTAs) implemented by the U.S. since the 1980s, and one of 70+ regional trade agreements (RTAs) that are currently in force. Although multilateral liberalization promotes gains from trade, economic literature suggests that regional liberalization – in which tariffs are lowered relative to member nations but maintained against the rest of the world – may not necessarily improve net welfare.

A traditional way to assess the welfare effects of an RTA is to look at trade creation and trade diversion. These concepts were pioneered by Jacob Viner in 1950 to analyze the merits of a customs union. Specifically, trade creation occurs when the removal of tariffs allows partner-country imports to replace high-cost domestic production (Viner, 1950). On the other hand, trade diversion occurs when partner-country imports displace lower-cost production from non-partner nations around the world (Viner, 1950). The net welfare effect of a regional trade arrangement can therefore be positive or negative, depending on the relative magnitude of trade creation and diversion.

In this paper, I examine the changes in Colombian trade patterns following the 2012 TPA to provide a preliminary assessment of the welfare effects resulting from the agreement. Based on Colombian import data for over 6,800 product categories available from 2011 through 2014, I investigate the extent to which the TPA has led to trade creation and/or trade diversion during the first two years of its implementation.

Exploring the welfare effects of the U.S.-Colombia TPA is important in an era where RTAs are shaping the landscape of international trade. In spite of the prevalence of RTAs, there remains to be much debate over the merits of these regional arrangements. Previous studies have found that the welfare effects of RTAs may differ across time periods and geography. Consequently, each RTA may require unique analysis to determine its impact on trade creation and diversion. Since the TPA is relatively recent, its welfare effects have yet to be studied in detail. An empirical analysis of the TPA therefore adds a valuable data point to existing literature. With a greater repertoire of case studies on various RTAs, economists can better outline the conditions that make certain agreements more conducive to welfare improvement than others. Understanding these conditions will be helpful in informing policy decisions about trade deals.

In this paper, I find that the U.S.-Colombia TPA resulted in trade creation effects during the initial two years of its implementation, with no evidence of trade diversion. Thus, the preliminary impact of the TPA appears to be welfare-improving.

2. Overview of Literature
In 1950, Viner challenged the assumption that regional trade agreements are always welfare-improving. By introducing the concepts of trade creation and trade diversion, Viner (1950) established a framework for analyzing the welfare effects of RTAs.

Since then, authors have attempted to determine the conditions under which RTAs are more or less conducive to welfare improvement, but have reached little consensus. Krugman (1991), for example, argues that net welfare would improve if neighboring countries (who are “natural” trading partners even without special arrangements) form an RTA; under this scenario, Krugman (1991) claims that potential losses from trade diversion are limited, while potential gains from trade creation are large. Meanwhile, Bhagwati (2008) questions the “natural trading partners” theory and contends that any regional trade agreement should be circumscribed because it threatens the ultimate goal of an open multilateral system.

Empirical studies have also run into difficulty reaching firm conclusions about the impact of RTAs. For example, Yeats (1998) finds evidence that Mercosur, a regional trading block consisting of Argentina, Brazil, Paraguay, Uruguay and Venezuela (at the time), was not internationally competitive. Yeats (1998) presents evidence that the increase in intra-Mercosur trade between 1988 and 1994 was largely caused by trade diversion from lower-cost sources outside the region. On the other hand, Musila (2005) finds empirically that two African RTAs (ECOWAS and COMESA) experience net trade
creation between 1991 and 1998. Based on existing literature, it appears that the intensity of trade creation or diversion varies across regions and periods. Thus, each RTA may require unique analysis to determine whether it improves or hurts overall welfare.

To date, there is great diversity in empirical approaches to studying the effects of regional trade agreements. One empirical strategy, introduced by Clausing (2001), exploits the variation in the extent of tariff liberalization under the Canada-U.S. FTA (CUSFTA) to assess the degree of trade creation and diversion. After the implementation of CUSFTA in 1989, tariffs on certain products were eliminated immediately, while others were reduced in equal increments over five or ten years. Thus, there was substantial variation in the level of initial protection as well as the degree of tariff liberalization. Clausing (2001) uses this variation to address two key questions: (1) first, do product groups with larger tariff reductions experience greater increases in trade with the partner country? (2) Secondly, do product groups that experience disproportionately gains in trade with the partner country experience disproportionately slower growth in trade with the rest of the world? Using empirical methods, Clausing (2001) finds that the CUSFTA was responsible for over half of the $42 billion increase in U.S. imports from Canada between 1989 and 1994, and that there was little evidence of trade diversion. Whereas the gravity model fails to discern an effect of CUSFTA on trade flows, Clausing (2001) teases out several insights by disaggregating the data.

This paper applies the empirical methodology used by Clausing (2001) to explore the welfare implications of the U.S.-Colombia TPA. Although I borrow the empirical methods of Clausing (2001), my paper differs from previous work in several ways. First, I provide a novel analysis of whether the bilateral agreement between the U.S. and Colombia has resulted in greater trade creation or diversion. To my understanding, there are no other papers that provide an ex-post analysis of the TPA along this dimension. Furthermore, my paper contributes to the topic of regional liberalization by expanding the repertoire of RTAs that have been empirically studied. Given the range of welfare effects observed in existing literature, it is helpful to add an additional data point to see whether a particular RTA generates greater trade creation or diversion.

3. Data
In this paper, I use panel data on Colombian imports between the years of 2011 and 2014. Since the TPA was implemented in 2012, the dataset provides enough coverage to assess how the bilateral agreement affected trade flows during the first two years of its implementation.

The dataset was obtained from Colombian customs officials (DIAN), and access was granted courtesy of Professor Andrew Bernard at the Tuck School of Business. Data is provided for each import transaction during the aforementioned time frame, and covers information on the value of the import transaction, the product category of the imported good, the ad valorem tariff rate, the date of the import, and the exporting country. The import values are reported in nominal dollars (CIF – USD), and have been crosschecked with reports on Trading Economics and the World Integrated Trade Solution Database to confirm that the numbers match those of other reputable sources.

The unit of observation is the product category of an import. Product categories are recorded using the Harmonized Classification System at the ten-digit level (HS-10). The dataset, in its full detail, contains over 6,800 unique ten-digit categories.

Tariff rates are given as the duties collected relative to the total value of trade in a given HS-10 category. In each import transaction, the tariffs levied are reported as a percentage point value. Tariff liberalizations taking place under the U.S.-Colombia TPA are scheduled according to the Harmonized system, and began on May 15, 2012. While some tariffs were eliminated immediately, others are to be lowered annually over a period of five to 17 years. The variation in the extent of tariff reductions, which are predetermined by the TPA, allows us to identify the effects of tariff liberalization on import growth.

Although it would be preferable to measure import values in real terms, the dataset does not contain price information for each HS-10 category. Thus, I do not have enough information to convert nominal import values to real ones. As an alternative solution, I use nominal import values in my regressions (see Section 4), but include year dummies to control for annual product price changes.
Section 4 goes into greater detail about how the dependent and independent variables are constructed from the data.

4.1. Trade Creation: Do Colombian imports from the U.S. increase due to the TPA?

4.1.1 Methodology

A preliminary look at Figure 1 reveals that the total value of U.S. products imported by Colombia has increased each year between 2011 and 2014. However, to what extent can post-2012 increases be ascribed to the TPA? Given the multitude of factors that influence trade flows, it is difficult to assess how tariff reductions actually impact the stream of imports into a country.

In order to isolate the effects of the TPA on raising Colombian imports of U.S. goods, I utilize the methodology from a paper by Clausing (2001), who assesses the impacts of the Canada-U.S. FTA on trade creation and diversion. Specifically, Clausing (2001) uses the varying extent of tariff liberalization across product groups to distinguish the effects of a trade agreement from other influences affecting trade flows.

The main regression, as modeled after Clausing (2001), is shown below:

\[
\% \Delta \text{Imports of U.S. Products}_{i,t} = \alpha + \beta_1 \% \Delta \text{U.S. Tariff}_{i,t} + \beta_\text{Year} (1.1)
\]

Specifically, this regression tests the hypothesis that U.S. products facing the greatest degree of tariff liberalization from the TPA will experience the largest increase in Colombian import demand. The dependent variable is the year-over-year percent change in Colombian imports of U.S. products \((\% \Delta \text{Imports of U.S. Products}_{i,t})\) across each HS-10 category. Specifically, the dependent variable will be approximated by:

\[
\ln(\text{Imports of U.S. Products}_{i,t}) - \ln(\text{Imports of U.S. Products}_{i,t-1}),
\]

where "i" represents each HS-10 category, "t" represents the year of interest, and "t-1" represents the preceding year.

The key independent variable is the year-over-year percent change in ad valorem tariff rates levied by Colombia on U.S. products \((\% \Delta \text{U.S. Tariff}_{i,t})\) for each HS-10 category. In particular, the independent variable will be approximated by:

\[
\ln(1 + \text{U.S. Tariff}_{i,t}) - \ln(1 + \text{U.S. Tariff}_{i,t-1}),
\]

where "i" – as before – represents each HS-10 category, "t" represents the year of interest, and "t-1" represents the preceding year. This approximation is fairly accurate when the tariff rate is small, which is typically the case in this dataset.²

Prior to running the regression, one would expect the key independent variable to have a negative coefficient. In other words, a fall in Colombian tariffs on U.S. products would correspond to an increase in Colombian imports of those products from the United States. A negative coefficient would indicate trade creation.

I run the regression using HS-10 product fixed effects. Moreover, I include fixed year effects to control for time-variant features, such as exchange rates, product prices, income changes, and cyclical factors.³ It is important to control for fixed effects given the panel structure of the dataset.

As a robustness check, I add another variable that may affect the responsiveness of imports to tariff liberalizations. The variable that I include is the share of Colombian imports that originated in the U.S. during the prior period. The main

---

² As mentioned in Section 3, the average U.S. tariff rate is small (3.81 percentage points) with a small standard deviation (6.38 percentage points).

³ The data covers four years of observations. Normally, controlling for fixed year effects would require the inclusion of three year dummies. However, since the dependent variable and key independent variable are computed as year-over-year changes, the regression technically only has three years of observations. Thus, I use two year dummies in my regression.
regression, after including this factor, is shown below:

\[
\% \Delta \text{Imports of U.S. Products}_{i,t} = \alpha + \beta_1 \% \Delta \text{U.S. Tariff}_{i,t} + \beta_2 \text{U.S. Share}_{i,t-1} + \beta_7 \text{Year} \quad (1.2)
\]

The new variable, U.S. Share\(_{i,t-1}\), is denoted as a decimal value between 0 and 1, and represents the U.S. share of total Colombian imports of commodity “\(i\)” from the previous period (“\(t-1\)”). I include this variable as a robustness check because the degree of trade creation due to tariff reductions may depend on the prior competitiveness of the partner country.

In all regressions, standard errors are clustered by HS-10 groupings to correct for heteroskedasticity and autocorrelation.

4.1.2 Results

Column (1) in Table 1 shows the output from regressing on \(\% \Delta \text{Imports of U.S. Products}_{i,t}\). Column (2) adds in fixed year effects. Column (3) includes U.S. Share\(_{i,t-1}\) as a control to check for robustness. The results from Column (2) indicate that a 1 percent decrease in tariffs on U.S. products is associated with a 3.29% increase in Colombian imports from the U.S.

In all three columns, the coefficient on the \(\% \Delta \text{U.S. Tariff}_{i,t}\) variable is negative and statistically significant at the 1% level. These results suggest that lower tariffs on U.S. goods are associated with higher imports from the U.S., providing evidence in support of trade creation. The trade creation effects are robust to the inclusion of the U.S. Share\(_{i,t-1}\) variable.

Moreover, it is interesting to note that the 2012-year dummy (the year in which the TPA was implemented) is significant and positively correlated with Colombian imports from the United States. This coefficient may reflect an increase in U.S. imports following the passage of the TPA on May 15, 2012, after which 80% of American imports became duty-free immediately.

Overall, the regression results provide evidence of trade creation effects during the first two years of the TPA’s implementation.

4.2. Trade Diversion: Do Colombian imports from the ROW decrease due to the TPA?

4.2.1 Methodology

After identifying trade creation effects, I now turn to a related question about trade diversion: do Colombian imports from the rest of the world fall due to the bilateral trade agreement? Once again, I modify a regression used by Clausing (2001):

\[
\% \Delta \text{Imports of ROW Products}_{i,t} = \alpha + \beta_1 \% \Delta \text{U.S. Tariff}_{i,t} + \beta_2 \% \Delta \text{ROW Tariff}_{i,t} + \beta_3 \text{U.S. Share}_{i,t-1} + \beta_7 \text{Year} \quad (2)
\]

This regression tests the hypothesis that the greatest reduction in ROW imports following the TPA are in product categories where the U.S. received the greatest tariff cuts.

Note that equation (2) is nearly identical to equation (1.2), but with two key differences. First, the dependent variable is no longer the year-over-year percent change in Colombian imports of U.S. products; rather, it is the year-over-year percent change in Colombian imports of rest-of-world products (%\( \Delta \text{Imports of ROW Products}_{i,t}\)) across each HS-10 category. Rest-of-world (ROW) is defined as all countries except the U.S. The dependent variable will be approximated by:

\[
\ln(\text{Imports of ROW Products}_{i,t}) - \ln(\text{Imports of ROW Products}_{i,t-1}),
\]

where “\(i\)” represents each HS-10 category, “\(t\)” represents the year of interest, and “\(t-1\)” represents the preceding year.

The second key difference is that an additional regressor is included in the equation. As in Clausing (2001), I add the year-over-year percent change in tariff rates levied by Colombia on ROW products (%\( \Delta \text{ROW Tariff}_{i,t}\)) for each HS-10 category\(^4\). The reason for including this variable is that, all else equal, an increase in average tariffs on ROW countries will

\(^4\) The ROW tariff is a trade-weighted average of the tariffs levied by Colombia on all of its trading partners, excluding the U.S. The ROW tariff is computed for each HS-10 category in all relevant years (2011-2014).
presumably lower total imports of ROW products, and failing to control for this factor will result in omitted variables bias. The $\%\Delta ROW$ Tariff$_i,t$ variable will be approximated by:

$$\ln(1 + ROW \text{ Tariff}_{i,t}) - \ln(1 + ROW \text{ Tariff}_{i,t-1}),$$

where “$i$” represents each HS-10 category, “$t$” represents the year of interest, and “$t-1$” represents the preceding year. Based on the reasoning described above, one would expect the coefficient on this regressor to be negative.

As in equations (1.1) and (1.2), the main independent variable in the trade diversion equation is the year-over-year percent change in tariff rates levied by Colombia on U.S. products ($%\Delta U.S.\text{Tariff}_{i,t}$). If the TPA leads to trade diversion, I would expect the coefficient on this variable to be positive and significant. In other words, a fall in Colombian tariffs levied on U.S. products would correspond to a fall in Colombian imports of these products from ROW countries. A positive coefficient would indicate that Colombian firms are likely switching their import purchases away from ROW countries to the United States.

In all regressions, standard errors are clustered by HS-10 groupings to correct for heteroskedasticity and autocorrelation.

### 4.2.2 Results

Column (1) in Table 2 shows the output from regressing on $%\Delta Imports$ of ROW Products$_i,t$. Column (2) adds $%\Delta ROW$ Tariff$_i,t$ as a control. Column (3) includes fixed year effects. Column (4) incorporates $%\Delta ROW$ Tariff$_i,t$ as a control to check for robustness. In all columns, a change in tariffs on U.S. products is not significantly related to a change in imports from ROW countries. The lack of statistical significance of the coefficient suggests that the TPA – in its first two years – has not led to notable trade diversion effects.

### 5. Discussion

One of the key strengths of using the methodology developed by Clausing (2001) is that it provides a sensitive measure of the TPA’s impact on import flows. Oftentimes, it is difficult to tease out the relationship between tariff liberalization and import growth using aggregate data, since a multitude of factors can affect a country’s nominal value of imports. However, by disaggregating the data and using the variation in tariff reductions across different product groups, this paper is able to identify the degree of trade creation and diversion that resulted from the TPA’s tariff liberalizations.

It is also important to address some of the limitations of the paper. First, the R-squared values are low for all of the regressions. As discussed by Clausing (2001), low explanatory power is one of the tradeoffs of using highly disaggregated data. Many difficult-to-measure influences likely account for the bulk of variation in annual Colombian imports. If the data is grouped into fewer buckets of commodities (i.e. using HS-2 levels), the explanatory power of the regressions increases, but this increase comes at a cost: aggregating the data makes it more difficult to isolate the effects of tariff liberalization on trade flows.

Including additional control variables that proxy for factors such as changing supply conditions and varying tastes would imaginably increase the explanatory power. However, these variables are difficult to measure. Constructing and including these variables could be a valuable area of exploration in the future.

Given the relatively recent implementation of the Colombia-U.S. TPA, another caveat is that the paper cannot draw conclusions about the long-term merits of the trade agreement. Since the TPA was passed in 2012 and import data was only available up until 2014, I could only assess the TPAs impact on trade flows during the first two years of its implementation. It is difficult to predict whether trade creation or diversion effects will dominate in the future. For example, there may be a stickiness effect, where Colombian firms may stay with existing trade partners during the first two years (due to contractual obligations, etc.), but as time passes, firms may divert their purchases away from ROW countries to the U.S.

---

5. This approximation is fairly accurate when the ROW tariff rate is small. This is generally the case in the dataset. The mean ROW tariff rate is 3.89 percentage points, and the standard deviation is 6.43 percentage points.
Thus, although this paper finds no effect of the TPA on generating trade diversion during the initial years, it is possible that diversion will occur in the long run.

If an updated dataset (up through 2017) becomes available, it would be interesting to examine the longer-term implications of the TPA on trade flows. Having a dataset with a longer time period would provide more degrees of freedom, which would allow me to include specific controls for time-varying factors (such as exchange rates, income changes, and other cyclical factors) instead of simply using year dummies.

6. Conclusion
The findings suggest that the U.S.-Colombia TPA had notable trade creation effects during the first two years of its implementation, with no evidence of trade diversion. Although it is not possible to draw conclusions about the longer-term welfare effects of the TPA based on the data in this study, it is still valuable to discuss some of the policy implications that arise from the paper’s findings.

In an era where RTAs are shaping the landscape of international trade, it is important to understand the conditions that make certain regional trade agreements more conducive to welfare improvement than others. So far, this paper reveals that the U.S.-Colombia TPA had positive welfare effects in the initial two years after its implementation. If follow-up studies can show that the TPA’s welfare effects persist in the long run, it would be valuable for policymakers to understand which factors contribute to the TPA’s success.

One potential aspect that could influence the success of a regional trading agreement is whether the countries involved are natural trading partners. Interestingly enough, the U.S. was Colombia’s largest trading partner, even before the TPA was implemented. As a topic of future research, it would be intriguing to explore whether RTAs between natural trading partners are more likely to be welfare improving. To approach this question empirically, one could start by classifying a broad sample of RTAs based on the degree to which member countries are natural trading partners. Researchers can then assess the extent to which each RTA has led to trade creation and diversion, and determine whether RTAs formed among natural trading partners yield better welfare effects.

Overall, learning about the conditions that make RTAs more conducive to welfare improvement would be useful in guiding policy decisions about the types of trade deals that countries should pursue. This paper serves as a small building block towards this broader objective.
References
Appendix

Figure 1. Annual Value of Colombian Imports of U.S. Products

<table>
<thead>
<tr>
<th>Year</th>
<th>Import Value (CIF USD - Billions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>13</td>
</tr>
<tr>
<td>2012</td>
<td>14</td>
</tr>
<tr>
<td>2013</td>
<td>16</td>
</tr>
<tr>
<td>2014</td>
<td>19</td>
</tr>
</tbody>
</table>

In all years, the total annual value of Colombian imports from the U.S. increases.

TABLE 1

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tariff change on U.S. products</td>
<td>-3.505***</td>
<td>-3.289***</td>
<td>-2.489***</td>
</tr>
<tr>
<td></td>
<td>(0.864)</td>
<td>(0.883)</td>
<td>(0.863)</td>
</tr>
<tr>
<td>Year dummy 2012</td>
<td>0.078**</td>
<td>0.120***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.030)</td>
<td></td>
</tr>
<tr>
<td>Year dummy 2013</td>
<td>0.016</td>
<td>0.016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.031)</td>
<td></td>
</tr>
<tr>
<td>Share of imports from the U.S. in previous period</td>
<td>-4.007***</td>
<td></td>
<td>(0.176)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.024***</td>
<td>-0.053***</td>
<td>1.027***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.018)</td>
<td>(0.052)</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.01</td>
<td>0.01</td>
<td>0.13</td>
</tr>
<tr>
<td>Number of observations</td>
<td>13,875</td>
<td>13,875</td>
<td>13,875</td>
</tr>
</tbody>
</table>

NOTES:
The dependent variable is the percent change in Colombian imports from the U.S., year-over-year. Robust standard errors (clustered by HS-10 codes) are reported in parentheses. ***, **, and * indicate significance at 1, 5, and 10% levels, respectively.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tariff change on U.S. products</td>
<td>-0.307</td>
<td>-0.309</td>
<td>-0.086</td>
<td>0.204</td>
</tr>
<tr>
<td></td>
<td>(0.252)</td>
<td>(0.254)</td>
<td>(0.253)</td>
<td>(0.245)</td>
</tr>
<tr>
<td>Tariff change on ROW products</td>
<td>0.102</td>
<td>0.562</td>
<td>-1.133**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.493)</td>
<td>(0.508)</td>
<td>(0.459)</td>
<td></td>
</tr>
<tr>
<td>Year dummy 2012</td>
<td></td>
<td></td>
<td>0.135***</td>
<td>0.108***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Year dummy 2013</td>
<td></td>
<td></td>
<td>-0.079***</td>
<td>-0.066***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.023)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>Share of imports from the U.S. in previous period</td>
<td></td>
<td></td>
<td></td>
<td>2.854***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.158)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.056***</td>
<td>0.056***</td>
<td>0.042***</td>
<td>-0.532***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.012)</td>
<td>(0.034)</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.10</td>
</tr>
<tr>
<td>Number of observations</td>
<td>17,398</td>
<td>17,398</td>
<td>17,398</td>
<td>17,398</td>
</tr>
</tbody>
</table>

NOTES:
The dependent variable is the percent change in Colombian imports from the ROW, year-over-year. Robust standard errors (clustered by HS-10 codes) are reported in parentheses. ***, **, and * indicate significance at 1, 5, and 10% levels, respectively.
A Performance Analysis of Managed Futures

CHRISTIAN FERNANDO
Georgetown University

ABSTRACT
Managed futures is a systematic hedge fund strategy that relies on quantitative models to analyze and follow price trends across diversified markets. The goal of this paper is to analyze the performance of managed futures in alternative contexts. I first examine whether or not managed futures can produce alpha and if it can time markets. In the latter half of the paper, I test if the addition of managed futures to traditional and hedge fund portfolios will improve portfolio performance. The results show that managed futures produces positive alpha and can time equity and commodity markets. The strategy cannot, however, time currency and bond markets. I also conclude that managed futures is a highly effective diversifier of traditional portfolios yet failed to positively improve the performance of hedge fund portfolios.
1. Introduction
The 2008 global financial crisis highlighted the weakness of traditional 60/40 stock and bond portfolios as they failed to effectively protect investor capital (Robaton, 2016). In response, investors have prioritized building more diversified portfolios with higher allocations to alternative asset classes which tend to be negatively or weakly correlated with stocks and bonds. Indeed, hedge funds have attracted substantial institutional interest, as assets under management (AUM) have doubled from $1.5 trillion in 2009 to nearly $3 trillion as of 2016 (Hedge Fund Industry Assets Under Management, n.d.). From an investor’s perspective, hedge funds offer a unique opportunity to access differentiated investment strategies that have unique return distributions and ideally will generate superior performance in the long-run performance (Harper, n.d.). Managed futures is projected to experience rapid growth in AUM as investors increasingly see it as an effective diversifier that can protect portfolios from periods of volatility and downturn (The Benefits of Managed Futures). I choose to focus my research on examining whether or not the newfound interest in managed futures is rational and whether or not the strategy does, in fact, produce positive performance benefits. Namely, I will explore whether or not managed futures can generate alpha, time markets and improve the risk-adjusted returns of traditional and hedge fund portfolios. In short, managed futures is simply the professional management of futures contracts and has several notable characteristics that differentiate it from traditional asset classes as well as other alternatives. These include being exchange traded, regulated, liquid, transparent, cash efficient, and non-directional. Trend following is the core of the investment process employed by managed futures managers. Trend followers believe markets are not efficient and over time will exhibit trends in price over time. These patterns are a result of market sentiment, economic cycles, seasonality and other momentum factors (Covel, 2006). Ultimately, trend following seeks to systematically and quantitatively identify price trends and then buy into these trends (go long if the trend is positive or go short if the trend is negative) for as long as they persist (Covel, 2006). In effect, this means that as a trend reverses, there will inevitably be a brief give-back in profits. However, so as long as the initial trend persists for some time, the strategy will be profitable on a net basis. In the case of a sharp reversal in trend reversal or a lack of trends altogether — due to a risk on – risk off sentiment amongst market participants —, trend followers will likely lose money (Covel, 2006). While the strategy is intuitive to understand, identifying and deciding how much to allocate to trends across different markets is an extremely complex process. Trend following systems are mechanical and rules-based, entering, allocating, risk managing and exiting trades based on a set of preconceived predetermined conditions. These rules are usually programmed into computer systems and are automatically executed. Managers favor this systematic approach because it removes human emotion from trading and allows for a more objective and quantitative analysis of price and market risks. Despite being a systematic strategy, trend followers are active managers as they continuously seek to optimize and improve the quantitative models used in futures trading.

I plan to outline my paper as follows: First, I will review the academic literature that discusses the performance of managed futures in various contexts. I will subsequently proceed to discuss the data, empirical strategy and relevant methods of evaluating investment performance. Lastly, I will provide the results of my research along with an interpretation of these findings which, hopefully, will provide useful insights for industry practitioners and academics with an interest in managed futures, hedge funds, and portfolio construction.

My empirical method is focused on calculating several key performance statistics in order to evaluate managed futures. In particular, I will first calculate alpha to determine the returns of the strategy adjusted for the relevant systematic risk

1 Managed futures is a systematic hedge fund strategy that uses quantitative models to identify and follow price trends across several markets.
2 Alpha is measurement of returns, adjusted for systematic (or market) risk. A traditional portfolio has a 60% allocation to global equities and a 40% allocation to global bonds. A hedge fund portfolio has an equally weighted allocation to global macro, equity long/short, equity market neutral, fixed income arbitrage, convertible bond arbitrage, distressed securities, merger arbitrage and event-driven multi-strategy.
3 A futures contract is an agreement to buy or sell a particular commodity or financial instrument at a specified time in the future. Futures, unlike forwards, are standardized and traded on exchanges. A forward contract is traded over-the-counter (OTC) and can be customized to suit the preferences of the negotiating parties.
4 The role of active management in a quantitative hedge fund will focus on improving and developing quantitative models to more effectively follow and capture trends.
factors. I also test whether or not managed futures can successfully time markets which I evaluate through using the
Treynor-Mazuy model. Lastly, I test whether or not the addition of managed futures to a traditional portfolio and to a
hedge fund portfolio will significantly improve the Sharpe Ratio in either case. The results of my research demonstrate that
managed futures produces positive, statistically significant alpha, can successfully time equity and commodity markets but
not bond and currency markets and significantly improves the Sharpe Ratio of a traditional portfolio at various allocations
but not the Sharpe Ratio of a hedge fund portfolio.

2. Literature Review

2.1 Theoretical Background on Portfolio Construction
The theoretical underpinnings of my research lie in modern portfolio theory (MPT). According to MPT, absolute returns
are ineffective measures of performance because they do fail to account for risk. Since most investors are risk averse, high
absolute returns are only favorable if the risk associated with these returns is tolerable. Markowitz discovered that since the
market positively compensates investors based on the level of risk they take, managers can construct portfolios to optimize
or maximize expected return based on a given level of market risk, emphasizing that higher risk is an inherent part of
higher reward (Jaffarian, 2007). In particular, Markowitz noted that the key goal in portfolio construction is seeking the
efficient frontier or the combination of assets that have a minimum combined variance at all possible levels of returns
(Markowitz, 1952). MPT also highlights the portfolio benefits of diversification as the addition of an uncorrelated asset
to the portfolio can result in an outward shift in the efficient frontier, effectively improving portfolio returns for every
level of risk (Modern Portfolio Theory, n.d.). In many ways, MPT is the theoretical cornerstone of portfolio management
as it highlights the benefits of diversification. The goal of my research is to apply the insights of MPT to evaluate the
performance of managed futures in various contexts.

2.2 Existing Literature on the Portfolio Impact of Managed Futures
Indeed, much of the literature on managed futures draws upon MPT to analyze the impact it has on traditional portfolios
of stocks and bonds. The study that first examined the portfolio impact of managed futures found that “the combined
portfolios of stock (or stock and bonds) after including judicious investment in managed futures accounts (or funds) show
substantially less risk at every possible level of expected return than portfolios of stocks (and bonds) alone” (Lintner, 1983).
Lintner’s findings were both challenged and confirmed by several subsequent academic studies. For example, a series of
studies by Elton, Gruber, and Rentzler (1987, 1989, 1990) examined public commodity pools, and, in contrast to Linter’s
findings, determined that there was little evidence for the benefits of managed futures. However, these studies (known
colloquially as the EGR studies) were criticized for several data issues, namely their reliance on public commodity pools in
their analysis, which is one of the most expensive ways to invest in managed futures (Black, 2016). Indeed, other analyses
of managed futures found evidence that, on average, the strategy provides attractive risk-adjusted returns, especially if the
performance is measured in the context of a diversified portfolio of stocks and bonds. For example, a 1999 study from the
Chicago Mercantile Exchange (CME) in 1999 found that “Portfolios with as much as 20% of assets in managed futures
yielded up to 50% more than a portfolio of stocks and bonds alone” (Hurst, n.d.). Similarly, a 1994 report from J.P. Morgan
concluded that a 15% or more allocation of managed futures to traditional stock and bond portfolios would reduce risk
and increase returns (Gregoriou and Zhu, 2005). While my research will also examine the impact that managed futures
has on traditional portfolios, it is differentiated in that it evaluates the statistical significance of the improvement in Sharpe
Ratios whereas many existing papers do not. In addition, it examines the impact that managed futures has on a hedge fund
portfolio as well as whether the strategy can produce positive alpha and time markets which are performance issues that
have not been adequately explored in the existing literature.

Throughout the aforementioned studies, performance analysis is a key component of the methodology, and these studies
share notable similarities in their empirical strategy. For instance, indices are typically used to represent the traditional
Brothers Government Bond index to represent fixed income. Other studies use broader indices such as MSCI World or
the Barclay Global Aggregate Bond Index to represent equities and bonds, respectively. Using a broader index is helpful
as it provides a complete representation of the asset class that is not specific to a certain geographic region. There is
more variation in how these studies represent different hedge fund strategies. Using hedge fund databases is a popular approach as they allow researchers to eliminate funds with incomplete or ambiguous data and give returns net of fees. Other empirical studies have resorted to different indices to represent hedge funds and specific hedge fund strategies. For example, researchers have represented managed futures using the Stark 300 Index or the MLM Index and have represented hedge funds with the CISDM Hedge Fund Index and the HFRI Index. The empirical process of testing the portfolio impact of managed futures has a fair degree of uniformity across the literature, despite slight deviations in thematic focus. Typically, papers will compare two portfolios: a traditional portfolio consisting only of stocks and bonds and the other being a traditional portfolio with an allocation to managed futures. If the portfolio with managed futures experiences superior performance over the same time period as the stock- and- bond portfolio, researchers can conclude that managed futures positively impacted the portfolio. The existing literature will be critical to informing my empirical approach, particularly with respect to finding data to represent the various asset classes that are pertinent to the study.

3. Data & Methodology

3.1.1 Representing Managed Futures and Hedge Funds
I use indices in order to represent hedge funds and managed futures in my analysis. Doing so prevents idiosyncratic biases such as poor management, hedge fund size, style drift or overcapacity from influencing performance as an index is an aggregation of all managers within the asset class. In addition, finding data on individual hedge funds or managed futures funds would be very challenging as the information is proprietary and rarely disclosed. In choosing the appropriate index, there are several factors that are worth considering. First, it is critical to use indices that are maintained by reliable institutions as doing so will minimize the probability that index data will have calculation errors that could create biases and discredit the results of this research. It is also important that the index provides a reasonable amount of historical data that is available at reasonable frequencies (quarterly, monthly or daily). Additionally, the index construction methodology must be able to represent the underlying strategy as effectively as possible. Lastly, the index data should be reasonably accessible and capable of being manipulated and analyzed. Given this criterion, I select the following indices: 1) CISDM Equal Weighted Hedge Fund Index to represent a portfolio of hedge funds 2) CISDM Equal Weighted CTA Index to represent managed futures 3) CISDM Global Macro Index for Global Macro 4) CISDM Equity Long / Short Index for Equity Long / Short 5) CISDM Equity Market Neutral for Equity Market Neutral 6) CISDM Fixed Income Arbitrage for Fixed Income Arbitrage 7) CISDM Convertible Arbitrage for Convertible Arbitrage 8) CISDM Distressed Securities for Distressed Securities 9) CISDM Merger Arbitrage for Merger Arbitrage 10) CISDM Event Driven Multi-Strategy for Event Driven Multi-Strategy.

Both the CISDM Equal Weighted Hedge Fund Index and the CISDM Equal Weighted CTA index use data from the Morningstar CISDM Database (formerly the MAR Database), the oldest hedge fund and managed futures database in the market with performance information from over 5,000 hedge funds, fund of funds and managed futures funds. The CISDM Equal Weighted Hedge Fund Index and CISDM CTA Equal Weighted Index are constructed on a monthly basis by taking the average performance of hedge funds and managed futures funds that have reported performance to the CISDM Morningstar Database. To prevent the results from being skewed, the index does not include outliers which are at least +/-3 standard deviations away from the average and equally weights the underlying funds constituting the index. By removing outliers, the index provides a more accurate representation of the typical performance of a managed futures fund or hedge fund portfolio and prevents the returns from being skewed upward or downward by outlier funds.

5 Style drift is a divergence from a stated investment strategy or objective for the purposes of achieving a superior return or reducing losses. Style drift is generally viewed negatively because a shift in investment strategy may change the fund’s risk-return profile and, by extension, the impact that the fund has on a portfolio, even if it positively benefits performance in the short run. Overcapacity is when an investment manager’s assets under management get so large that it begins to negatively impact returns as there aren’t enough investment opportunities to target. As a result, the manager is forced to take highly concentrated positions or hold the new investment funds in cash, both of which are inefficient outcomes.

6 Fund of Funds (FoFs) are hedge funds that invest in other hedge funds. Investors who then invest in the FoFs gain exposure to the hedge funds that the FoFs is invested in. FoFs are popular with hedge fund investors that are looking to diversify across managers at a low cost.
The strategies represented in the CISDM Equal Weighted Hedge Fund Index are very comprehensive, making it an effective strong proxy for a hedge fund portfolio. What makes the hedge fund index even more advantageous for my analysis is that it doesn’t include managed futures as an underlying strategy, making it possible to test the performance impact of adding managed futures to a hedge fund portfolio. Furthermore, given that all the indices are constructed by CISDM and use the same rules, there is no risk that differing index construction methodologies would create biases in the analysis. Lastly, the returns for these indices are net of fees, which is important in order to provide a realistic measure of investment performance (Center for International Securities and Derivatives Markets, n.d). In short, the construction methodology makes the indices effective proxies for the typical performance of a hedge fund portfolio and managed futures strategy.

In order to calculate other performance metrics such as alpha and market timing, it will be necessary to incorporate the following asset classes: 1) Equities – MSCI World Index 2) Bonds – Barclays Global Aggregate Bond Index 3) Currencies – DB G10 Currency Futures Index 4) Commodities – GS Commodities Index

Given that managed futures managers are trend followers and macro-oriented investors, they typically do not trade individual equities. Instead, they trade futures on equity indices in order to capture trends in different equity markets. Typically, managers will focus on developed market equities as the futures for these indices are significantly more liquid and therefore reduce trading and transaction fees. The MSCI World Index is an effective way to represent equity exposure, as it is a broad global equity index that represents large and mid-cap companies across twenty-three developed markets. Managers also trade futures on a global portfolio of bonds that are diversified by geography, maturity, interest rate and originator (for example government versus corporate), making it a good idea to represent the asset class through the Barclays Global Aggregate Bond Index which is a broad, fixed-income index that tracks global government, corporate and securitized fixed-income investments. Managers also trade futures on currencies and commodities, and these two assets typically make up the largest proportion of managed futures portfolios. Typically, managers will focus on the more liquid currency and commodity contracts, yet there is considerable variety with respect to the specific markets that are traded in. For example, smaller managers will tend to focus on trading a specific commodity market such as energies or metals while as larger managers trade a more diversified portfolio with exposure to several commodity markets. Such differences are also typically true for currencies, with the exception that most managers will trade futures on developed market currencies as they are more liquid and cheaper to trade. Due to the variation across managers, I use broad-based commodity and currency indices — the Goldman Sachs Commodity Index and the DB G10 Currency Futures Index — to represent these respective sources of market risk.

These indices will be sufficient to evaluate market timing as they represent the core markets that managed futures managers trade in. However, the indices are likely to be ineffective when calculating alpha since they are long only. This is problematic as managed futures managers can short when markets decline, and therefore benefit from systematic risk during bear markets. To control for this, I include the S&P 500 Diversified Trend Index (DTI), in addition to the aforementioned core markets, as a fifth regressor when calculating alpha. S&P 500 DTI is a passive index of managed futures that makes trades based on price signals given by a seven-month moving average. This methodology allows the index to take long and short positions based on price trends (similar to managed futures managers) but by using less sophisticated indicators compared with managed futures (Looking Under the Hood “Passive Managed Futures Mutual Funds” n.d.). The main benefit of incorporating the S&P 500 DTI, in addition to equities, bonds, currencies and commodities, when calculating alpha is that it controls for systematic risk during both bull and bear markets because it can go long or short.

3.2.1 Empirical Strategy – Jensen’s Alpha

The objective of my research is to evaluate managed futures in alternative contexts; namely, to determine if managed futures can produce alpha, time markets and improve the risk-adjusted returns of traditional and hedge fund portfolios. Alpha is simply a measurement of a portfolio’s return adjusted for its systematic risk and is an effective way of quantifying manager skill as it controls for the proportion of returns that can be attributed to favourable market movements (luck). With that said, the metric’s effectiveness largely depends on how effectively the benchmark(s) control for systematic risk and the considered time frame. In addition, it does not account for the idiosyncratic risk of an investment, making it possible for a relatively poor manager to achieve positive alpha by taking on excessive risk by, for example, leveraging
investment positions (Aragon and Ferson, 2007). Below is the formula for Jensen's alpha:

\[ \alpha = (R_p - R_f) - B_p(R_m - R_f) \]

As demonstrated in the equation above, Jensen's Alpha is calculated by regressing the excess returns of the portfolio against the excess returns of various systematic risk factors. I use the 3-month Treasury bill rate as a measure of the risk-free rate. In the case of managed futures, the source of systematic risk comes from the core markets that it trades: equities, bonds, currencies and commodities. I represent these markets through indices and include them as regressors in my calculation of alpha. However, as previously discussed, using long-only indices is insufficient as they will only control for systematic risk in bull markets when in fact managed futures managers benefit from systematic risk during bear markets as managers can short. In order to control for this, I use the S&P 500 DTI as the fifth regressor since it is a long/short index and therefore will control for systematic risk during bear markets as well. One of the benefits of Jensen's alpha is that it provides a measure of statistical significance to the calculated excess return, which gives it an interpretable value (Aragon and Ferson, 2007). I also test whether or not the S&P 500 DTI, a passive benchmark of managed futures, can produce positive and statistically significant alpha. Given that the S&P 500 DTI primarily trades commodities and currencies, I represent these markets through indices and include them as the first two regressors in order to control for systematic risk. However, long only indices are insufficient as they only control for systematic risk in bull markets while as the S&P 500 DTI benefits from exposure to bear markets through its ability to short. In order to control for this, I use managed futures as a third regressor as it is a long/short strategy that trades commodities and currencies, among other asset classes, and therefore will control for market exposure in down markets as well.

3.2.2 Empirical Strategy – Market Timing

Having examined alpha, I analyze performance by testing the strategy’s market timing ability. Given that the central investment goal of managed futures managers is to systematically and quantitatively identify and follow trends in price, a key element of the strategy is determining when trends begin and when trends start to falter so that positions can be risked and de-risked appropriately. Given the nature of the strategy, it is reasonable to expect that managed futures managers demonstrate some degree of market timing skill. To test this, I use the Treynor-Mazuy Model which is calculated through using the following regression:

\[ r_{pt + 1} = a + b r_{pt + 1} + A_1 r^2_{pt + 1} + v + 1 \]

The Treynor-Mazuy model is essentially a quadratic extension of the basic CAPM and is estimated using multiple regression. The second term in the regression, \( r_{pt + 1} \), is the value of excess return squared. Treynor and Mazuy (1966) argue that if the strategy increases (decreases) the portfolio's market exposure prior to an increase (decrease) in the market, then the estimated equation will describe a convex upward-sloping regression “line” and will generate a Lambda coefficient that is positive (Aragon and Ferson, 2007). Typical managed futures managers trade currencies, commodities, bonds and equities and, if the strategy can successfully time markets, it makes sense to test timing ability in the markets that the strategy actually trades in. Since managers can effectively go long and short in any market, I hypothesize that the TM model should indicate that the strategy shows successful market timing ability so long as there is a persistent degree of price trends during the examined time horizon.

I also examine whether or not market timing is uniquely inherent to active managers by also computing the market timing ability of the S&P 500 DTI. In particular, I test whether or not S&P 500 DTI can time currency, commodity, equity, and bond markets. Testing market timing of the S&P 500 DTI is useful as it contextualizes the value-add of the sophisticated, quantitative algorithms used by managed futures managers. Effectively, if a semi-passive benchmark which uses a simple, 7-month moving average to make buy and sell decisions can time markets just as effectively as an actively managed managed futures fund, then the market timing ability of managed futures managers is less impressive. However, if managed futures managers display superior market skill compared to the S&P 500 DTI, then market timing is a unique

---

7 Excess return is calculated as the return of the security (in this case managed futures) minus the risk free rate which is typically represented by short term (3-month) U.S. treasuries.
property of the strategy as it cannot be replicated with a simplistic and cheaper option.

3.2.3 Empirical Strategy – Portfolio Impact of Managed Futures
Having calculated alpha and market timing for managed futures, I now turn to examine whether or not the strategy can add value to a traditional portfolio and a hedge fund portfolio. I represent the portfolio of hedge funds using the CISDM Equal Weighted Hedge Fund Index and managed futures from using the CISDM CTA Equal Weighted Index using data from 1994. The traditional portfolio is represented by a 60% allocation to global equities, represented by the MSCI World Index, and a 40% allocation to global bonds, represented by the Barclays Global Aggregate Bond Index. To analyze the portfolio impact of managed futures, I conduct a difference in Sharpe Ratios test of a hedge fund portfolio with and without managed futures. Doing so will determine whether or not the addition of managed futures to hedge fund portfolios can result in statistically significant improvements in performance. In short, my null hypothesis (H₀) is that managed futures do not impact portfolio performance, and my alternative hypothesis (Hₐ) is that there is, in fact, a difference in portfolio performance when managed futures is added. According to convention, I test for statistical significance at the 10%, 5%, and 1% significance levels. To calculate the p-values for the differences in Sharpe Ratio, I derive a distribution of the differences in the Sharpe Ratios of Portfolio A, a pure hedge fund portfolio, and Portfolio B, a hedge fund portfolio with an allocation to managed futures. Creating this distribution of differences in Sharpe Ratios is empirically valid because the returns of portfolio A and B are independently and identically distributed (IID); or, in other words, the individual portfolio returns have a low serial correlation. According to the Central Limit Theorem, this means that their linear combination will also be IID and approach a normal distribution, allowing for tests of statistical significance. However, the primary advantage of this methodology is its practical utility. In particular, Opdyke (2007) shows that the greater the correlation between the returns of Portfolio A and Portfolio B, the lower the variance will be in the Sharpe Ratio distribution. This improves the statistical power (or the probability of correctly rejecting the null hypothesis that the Sharpe Ratios are equal when they are in fact not equal) of the difference in Sharpe Ratios test. This is pertinent to my research as it is reasonable to expect portfolio A and portfolio B to be highly correlated as they hold the same assets, with the only difference being that portfolio B has an allocation to managed futures while as portfolio A is a pure hedge fund portfolio. Due to this attractive property, the methodology provides an extremely robust strategy to perform significance testing on differences in Sharpe Ratios (Opdyke, 2007). Despite the noted advantages with this methodology, one issue with comparing Sharpe Ratio improvements for different portfolio weights is that these are not independent tests as the t-statistic for different allocations will have a different sampling and, by extension, different sampling variance. As a result, caution is due since a lack of independent tests may result in a slight over or understatement of Type I errors, and, by extension, the p-values.

The formula for the Sharpe Ratio is given below:

\[
\text{Sharpe Ratio} = \frac{R - R_f}{\sigma}
\]

Effectively, it is a measure that evaluates a portfolio’s return in excess of the risk free rate per unit of total risk.\(^8\) The Sharpe Ratio is traditionally thought to be most effective when applied to an investor’s total portfolio as opposed to a specific security or asset class. Applying the Sharpe Ratio to a single asset class in isolation ignores the correlation of the asset with other investments in the portfolio, so it may not correspond in any meaningful way to the desirability of the asset as an investment in any meaningful way (Aragon and Ferson, 2007). A downside with of the Sharpe Ratio is that it uses standard deviation to measure risk which can be misleading as it punishes both positive and negative deviations from the mean. The final step in my analysis is to test the impact that the S&P 500 DTI has on traditional and hedge fund portfolios. Similar to the managed futures tests, my null hypothesis (H₀) is that S&P 500 DTI does not impact the traditional/hedge fund portfolio and my alternative hypothesis (Hₐ) is that there is, in fact, a difference in portfolio performance when S&P 500 DTI is added. In order to remain consistent, I test for statistical significance at the 10%, 5%, and 1% levels. The main reason for running a similar test on the S&P 500 DTI is to determine whether or not the portfolio impact on managed futures is unique to the strategy or whether a similar outcome can be achieved through the S&P 500 DTI, a far less sophisticated, but similar trend following strategy. If the results show that the S&P 500 DTI had a similar impact, then it diminishes the conclusion that there is something unique about investing in an actively managed, systematic trend following strategy like

---

8 Total risk is the combination of both idiosyncratic and systematic risk.
managed futures because the same results can be achieved through a simpler iteration of the strategy. By contrast, if there is a noticeable difference between the two cases, then it is clear that managed futures managers provide unique value to an investment portfolio, potentially justifying active management and the higher fees that are charged to investors.

4. Results and Analysis

4.1 Alpha of Managed Futures

In order to assess the performance of managed futures, I test whether or not the strategy can produce above-market return or alpha.

<table>
<thead>
<tr>
<th>Beta Factors</th>
<th>Alpha</th>
<th>P value</th>
<th>R Squared</th>
<th>Adj R Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stocks</td>
<td>-0.0059</td>
<td>0.011</td>
<td>0.059</td>
<td>0.052</td>
</tr>
<tr>
<td>Stocks and Bonds</td>
<td>-0.0009</td>
<td>0.662</td>
<td>0.295</td>
<td>0.285</td>
</tr>
<tr>
<td>Stocks Bonds and Currencies</td>
<td>0.0018</td>
<td>0.388</td>
<td>0.384</td>
<td>0.371</td>
</tr>
<tr>
<td>Stocks Bonds Currencies and Commodities</td>
<td>0.0022</td>
<td>0.307</td>
<td>0.388</td>
<td>0.370</td>
</tr>
<tr>
<td>Stocks Bonds Currencies Commodities and DTI</td>
<td>0.0399</td>
<td>0.027</td>
<td>0.650</td>
<td>0.638</td>
</tr>
</tbody>
</table>

Source: Bloomberg

I first regress the returns of managed futures against equities as it is common for managers to trade futures on equity indices as part of capturing directional momentum trends in different equity markets. This generates a negative alpha of -.0059, and the result is significant at the 5% significance level. Given the statistical significance of the result, at first, it seems as if managed futures does not generate alpha. However, the result also has an R-squared of .059, which suggests that equities are not a major driver of returns and that there are other sources of systematic risk that haven't yet been accounted for yet. To address this shortcoming, I add bonds to the regression, as it is common for managers to trade futures on fixed income instruments as a means of capturing trends in bond markets. Even with this result, the alpha remains negative at -.0009. However, the result is no longer statistically significant as the P value rose to 66%. The result implies that even if stocks and bonds were the only systematic risk factors impacting managed futures, we couldn't conclude (with any reasonable level of statistical significance) that the strategy generates negative alpha. Yet, even with the inclusion of stocks and bonds, not all sources of systematic risk have been accounted for. I next add currencies to the regression, and the result is the first positive alpha of 0.018. However, this result is not statistically significant as the P value is 39.5%. I next add commodities to the regression, anticipating that this will have a significant impact on alpha as commodities are one of the most popular asset classes for managed futures managers to trade. The output shows that alpha improves to .0022 yet is not a statistically significant improvement as the P value remains above 10%. It is only when the final regressor, the S&P 500 DTI, is added to control for systematic risk during bear markets that alpha jumps to .0399 which is statistically significant at the 5% level demonstrating that managed futures does, in fact, generate statistically significant alpha.

The strategy’s positive alpha can be explained by its favorable return distribution. In particular, managed futures have superior cumulative performance in comparison with the markets that it trades in (Figure 1).
The results demonstrate that, initially, managed futures produces relatively modest returns that are comparable with commodity, currency, equity, and bond markets. The performance diverges after 2008 as managers generate positive returns during the year, starkly contrasting major markets like equities and commodities which experience high drawdowns. Similarly, during 2011, when equities experienced a slight downturn due to Eurozone concerns, and, in 2014, when commodity prices collapsed, managed futures was unfazed and produced positive returns. When examined over a longer period, it becomes clear that managed futures is able to generate relatively high levels of return for less risk, allowing it to produce alpha.

While the results demonstrate that managed futures can generate statistically significant alpha, some may contend that the positive alpha is largely a result of the strategy’s ability to follow momentum which can be replicated successfully through simplistic trend following models which are not actively managed. In order test the validity this counter-argument, I calculate the alpha of S&P 500 DTI and compare it to that of managed futures.9

I first regress the returns of S&P 500 DTI against currencies as the indicator trades futures on several G10 currencies and seeks to capture directional trends in the currency market. The regression output generates a negative alpha of -0.0070 which is statistically significant at the 1% level. While currencies are a major market for the S&P 500 DTI, commodities make up nearly 50% of its portfolio and clearly are also a key source of systematic risk. Indeed, once included in the regression, alpha improves slightly, rising to -0.0037 which is statistically significant at the 10% level. Despite having controlled for the two key markets that the strategy trades in, the R-squared rises to only 29.3%. This implies that systematic risk explains only a third of returns which is unrealistic given that the S&P 500 DTI, like managed futures, is a globally traded and diversified strategy. In order to control for systematic risk during bear markets, I include managed futures as the third regressor. As a result, alpha rises to -0.0029 and remains statistically significant at the 10% significance level. Even though the S&P 500 DTI does not trade equity or bond futures, I include these markets as final regressors for consistency; as expected, the R-squared value barely changes and the negative alpha effectively remains the same, reflecting the fact that the S&P 500

9 I explain why the S&P 500 DTI is a good passive indicator for managed futures in the Data and Methodology section.
DTI does not have market exposure to equities or bonds. The results indicate that S&P 500 DTI cannot produce positive alpha and its positive returns are largely a result of favorable market performance rather than sophisticated investment skill. This is a markedly different result from the managed futures case which demonstrated that the strategy was able to produce alpha that was positive and statistically significant, thereby exemplifying the value-add of active management.

4.2 Market Timing Ability of Managed Futures and S&P 500 DTI
As part of evaluating the performance managed futures, I also consider whether or not the strategy can successfully time markets which I test by using the Treynor- Mazuy model.

Table 3. Market Timing Test – Managed Futures (2004 – 2016)

<table>
<thead>
<tr>
<th>Market Timing Test - Managed Futures</th>
<th>( \Delta )</th>
<th>( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonds</td>
<td>-2.08</td>
<td>0.47</td>
</tr>
<tr>
<td>Global Equities</td>
<td>1.55</td>
<td>0.01</td>
</tr>
<tr>
<td>Currencies</td>
<td>-1.27</td>
<td>0.51</td>
</tr>
<tr>
<td>Commodities</td>
<td>0.76</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Source: Bloomberg

The results show that managed futures can successfully time commodity and equity markets but cannot time bond and currency markets. This result appears, at first, to be perplexing. If managers apply the same quantitative models to all markets, one would expect the strategy to demonstrate the same degree of market timing in all cases. However, upon closer examination, it is clear that market timing skill also depends on the return distribution in different markets which is independent of manager skill. Since the TM Model will reflect market timing if a strategy is able to make money as prices rise and lose less money when prices are falling, if markets do not exhibit a fair degree of cyclicality, then there is less of an opportunity to “time the market” as there isn’t significant directional fluctuation in price. Intuitively, this explanation makes sense as the two markets that managed futures successfully timed — equities and commodities — tend to be more volatile than the two markets that managed futures failed to time (Figure 2).

Figure 2. Cumulative Performance (1993 – 2016)

Source: Bloomberg

As demonstrated in Figure 2, performance for bonds has been a consistent upward trend, providing managers with no opportunities to time the market as a long position in bonds would have been maintained throughout the timeframe. Though currencies exhibit slightly larger fluctuations, the only examples of substantial negative trends are the 2007-2008 financial crisis and the period between 2013-2015. Though these instances would have provided managers with opportunities to switch to a short position on currencies, for most of the time period managers would have been long.
currencies as the trend has been clearly positive. By contrast, managers would have experienced ample opportunities to
time equity and commodity markets as there has been a substantial amount of directional fluctuation in both markets.
This point can be further exemplified by comparing the percentage of positive months of the respective markets as well as
their volatilities (Table 4).

**Table 4. Structural Characteristics of Core Markets (1993 – 2016)**

<table>
<thead>
<tr>
<th>Volatility (1993 - 2016)</th>
<th>Global Stocks</th>
<th>Commodities</th>
<th>Global Bonds</th>
<th>Currencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of Positive Monthly Returns</td>
<td>15%</td>
<td>21%</td>
<td>3%</td>
<td>9%</td>
</tr>
<tr>
<td>59%</td>
<td>56%</td>
<td>70%</td>
<td>63%</td>
<td></td>
</tr>
</tbody>
</table>

*Source: Bloomberg*

The results show that commodities and equities exhibit a roughly even amount number of positive and negative months while
as bonds and currencies have a higher proportion of positive months. In effect, this shows that equities and commodities
have relatively higher degrees of fluctuation, a point that is also demonstrated by their relatively higher volatilities. Since
higher proportions of market fluctuations will provide managers with more opportunities to demonstrate market timing
skill, the structural difference in volatility between equity and commodity versus bond and currency markets is the best
explanation for the discrepancies in market timing.

I also test whether or not the S&P 500 DTI will exhibit similar market timing skill through the Treynor- Mazuy model
(Table 5).


<table>
<thead>
<tr>
<th>Market Timing Test - S&amp;P 500 DTI</th>
<th>A</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonds</td>
<td>-6.09</td>
<td>0.04</td>
</tr>
<tr>
<td>Global Equities</td>
<td>1.77</td>
<td>0.00</td>
</tr>
<tr>
<td>Currencies</td>
<td>-0.40</td>
<td>0.82</td>
</tr>
<tr>
<td>Commodities</td>
<td>0.98</td>
<td>0.00</td>
</tr>
</tbody>
</table>

*Source: Bloomberg*

The results show the S&P 500 DTI behaves very similarly to managed futures with respect to market timing. Specifically,
S&P 500 DTI successfully times commodity and equity markets to the same degree of statistical significance as managed
futures but cannot time currency and bonds markets. These results are particularly interesting as they demonstrate that
market timing skill is a result of the ability to go long and short based on price trends as opposed to the degree of
sophistication in a managers’ trend following models, risk management, and general investment process. Given how
strongly correlated the S&P 500 DTI is with managed futures, it is unsurprising that the semi-passive index exhibited a
similar level of market timing.

### 4.3.1 The Impact of Managed Futures on Traditional Portfolios

Having examined whether managed futures can produce alpha and time markets, I now turn to test the impact that
managed futures has on traditional portfolios. In order to improve the robustness of my results, I test the portfolio impact
at varying allocations of managed futures.
The results convincingly demonstrate that the addition of managed futures to a traditional portfolio improves performance by a statistically significant margin. Specifically, for 6 out of 9 allocations, the Sharpe ratio increase was significant at the 1% significance level, for 1 out of 9 allocations the Sharpe Ratio increase was significant at the 5% level and in a final case, the Sharpe Ratio increase was significant at the 10% level. There was only one case where the addition of managed futures resulted in an improvement that wasn’t statistically significant. The results demonstrate that as increasing proportions of managed futures are initially added to the traditional portfolio, the Sharpe Ratio improves by significant margins. However, as higher and higher proportions of managed futures are added to the portfolio, the benefit starts to diminish at an increasingly rapid rate. With that said, it is worth noting that even when managed futures is significantly over-weighted the Sharpe Ratio improvement is greater than the case where it is significantly under-weighted. Ultimately, the results show that managed futures is a demonstrably valuable asset class that significantly improves the risk-adjusted returns of traditional portfolios of stocks and bonds.

4.3.2. The Impact of Managed Futures on Hedge Fund Portfolios
Having examined the impact of managed futures on traditional portfolios, I now turn to test whether or not a similar Sharpe Ratio improvement can be achieved for hedge fund portfolios. As with the traditional portfolios, I test the portfolio impact of adding managed futures at varying allocations.

The results show that while the addition of managed futures improves hedge fund portfolio performance in a majority of cases, the difference in Sharpe Ratios is not statistically significant at the 10%, 5% or 1% significance levels. To improve the definitiveness of my conclusion, I use mean-variance optimization to determine the allocation of managed futures that maximizes the portfolio Sharpe Ratio. In this case, the optimal allocation of managed futures is 35%, yet even at this point, the Sharpe Ratio improvement is not statistically significant. The result is ina sharp contrast with the impact that
managed futures had on traditional portfolios, particularly because none of the Sharpe Ratio improvements are statistically significant in this case. However, in both the hedge fund and traditional portfolios the initial addition of managed futures adds value, but as managed futures becomes a greater proportion of the total portfolio, the benefit is rapidly eroded. In the hedge fund portfolio, this effect is even more dramatic as significantly overweighting managed futures (70%, 80% or 90%) actually erodes its diversification benefits, resulting in a Sharpe Ratio that, in fact, lower than that of the pure hedge fund portfolio. Furthermore, in the hedge fund portfolio case holding a heavy overweight to hedge funds generates a higher Sharpe Ratio than a heavy overweight to managed futures. It is also worth noting that the optimal amount of managed futures is significantly lower in the hedge fund portfolio compared with the traditional portfolio, thereby further highlighting its limited value the former case.

It is worth noting that the conclusion on whether or not managed futures improves the Sharpe Ratio is entirely dependent on the significance level that is used. In this case, I chose to evaluate differences in Sharpe Ratios with the conventional 10%, 5%, and 1% significance levels. With that said, if readers are more willing to tolerate a higher probability of a Type I error – that is, a higher probability of incorrectly rejecting the null hypothesis – the final conclusion will change. For example, if I raise the significance level to 15% then a 30% and 35% allocations of managed futures does improve the hedge fund portfolio’s Sharpe Ratio by a statistically significant degree. Likewise, if the significance level is raised to 20%, then a 20%, 30%, 35% and 40% allocations of managed futures generates a statistically significant improvements in the Sharpe ratio of the hedge fund portfolio. While my conclusion is valid under statistical conventions, it will be key for readers to consider their own tolerance for false positives (Type I errors) and accordingly interpret the results presented in Table 7.

4.3.3 The Impact of Managed Futures on Portfolio Volatility

Managed futures appears to have been more effective at reducing portfolio volatility for the traditional portfolio than for the hedge fund portfolio (Table 8).

<table>
<thead>
<tr>
<th>Allocation</th>
<th>Traditional Portfolio</th>
<th>Hedge Fund Portfolio</th>
</tr>
</thead>
<tbody>
<tr>
<td>10MF-90 Traditional/HF</td>
<td>0.94%</td>
<td>0.62%</td>
</tr>
<tr>
<td>20MF-80 Traditional/HF</td>
<td>1.77%</td>
<td>1.08%</td>
</tr>
<tr>
<td>30MF-70 Traditional/HF</td>
<td>2.45%</td>
<td>1.36%</td>
</tr>
<tr>
<td>40MF-60 Traditional/HF</td>
<td>2.93%</td>
<td>1.42%</td>
</tr>
<tr>
<td>50MF-50 Traditional/HF</td>
<td>3.19%</td>
<td>1.25%</td>
</tr>
<tr>
<td>60MF-40 Traditional/HF</td>
<td>3.18%</td>
<td>0.88%</td>
</tr>
<tr>
<td>70MF-30 Traditional/HF</td>
<td>2.91%</td>
<td>0.34%</td>
</tr>
<tr>
<td>80MF-20 Traditional/HF</td>
<td>2.41%</td>
<td>-0.33%</td>
</tr>
<tr>
<td>90MF-10 Traditional/HF</td>
<td>1.73%</td>
<td>-1.11%</td>
</tr>
</tbody>
</table>

Source: Bloomberg

Portfolio volatility is driven by the volatility of the underlying assets and the degree to which they are correlated. Portfolios with assets that are highly correlated will experience higher levels of volatility because they are less effectively diversified and move together. By contrast, portfolios that have assets which are negatively or weakly correlated will be better diversified and experience a more stable performance. One possible hypothesis for the relatively stronger volatility reduction in the traditional portfolio is that managed futures proves to be significantly less correlated with traditional assets and therefore is more effective at lowering portfolio volatility. I test whether this is, in fact, the case by comparing the correlations of managed futures with various hedge fund strategies as well as stocks and bonds (Table 9).

| Correlation of Managed Futures w/ Traditional & Alternative Assets (1998 – 2016) |
|-----------------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Source: Bloomberg

<table>
<thead>
<tr>
<th>Managed Futures</th>
<th>Convertible Arbitrage</th>
<th>Distressed Securities</th>
<th>Equity Long/Short</th>
<th>Equity Market Neutral</th>
<th>Event Driven Multi-Strategy</th>
<th>Fixed Income Arbitrage</th>
<th>Global Macro</th>
<th>Global Macro</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3.19%</td>
<td>-9.88%</td>
<td>0.76%</td>
<td>15.75%</td>
<td>-7.55%</td>
<td>-4.11%</td>
<td>49.41%</td>
<td>-8.22%</td>
<td>-8.81%</td>
</tr>
</tbody>
</table>
Contrary to expectation, managed futures does not appear to be significantly less correlated with stocks and bonds as opposed to hedge funds. In fact, managed futures is fairly negatively correlated with fixed income arbitrage, event-driven multi-strategy, merger arbitrage, convertible arbitrage and distressed securities. Likewise, its correlation with bonds is higher than its correlation with any hedge fund strategy with the exception of global macro. Given this evidence, it appears that the volatility reduction in the traditional portfolio was not a function of managed futures’ low correlation to the underlying assets.

However, the answer may lie in differing levels of asset volatility. As one would expect, the addition of more volatile assets to a portfolio will increase overall portfolio risk while as the addition of less risky assets will decrease portfolio volatility. As such, a possible hypothesis for why managed futures was more effective at reducing portfolio volatility in the traditional portfolio is that the strategy is less volatile compared to stocks and bonds. By contrast, managed futures is probably more volatile compared to the underlying hedge fund strategies which is why it did not reduce portfolio volatility as dramatically. In order to test this hypothesis, I simply compare the volatilities of the various hedge fund strategies as well as stocks and bonds to the volatility of managed futures (Table 10).

<table>
<thead>
<tr>
<th>Table 10. Volatility of Traditional and Alternative Assets Relative to Managed Futures (1998 – 2016)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>Volatility</td>
</tr>
</tbody>
</table>

Source: Bloomberg

The results show that managed futures is, in fact, more volatile than each of the underlying hedge fund strategies in the portfolio, yet is nearly half as volatile as stocks and only slightly more volatile than bonds. Effectively, the results imply that because managed futures had a higher volatility than the underlying hedge fund strategies, when it was added to the hedge fund portfolio it actually contributed to the overall portfolio volatility. This impact was easily offset at lower allocations (such as 10%, 20%, and 30%) as managed futures was sufficiently weakly correlated with the underlying hedge fund strategies. However, as the allocation to managed futures grew, its higher volatility tended to overpower its relatively low correlation with the underlying strategies. This was why an 80% or 90% allocation to managed futures actually raised the volatility of the hedge fund portfolio rather than instead of lowering it. For the traditional portfolio, the effect is almost the complete opposite. Since traditional portfolios have a 60% allocation to stocks and a 40% allocation to bonds, they are overweight an asset that has nearly double the volatility of managed futures. This gives managed futures significant potential to act as a risk reducer in the portfolio, a point that is clearly demonstrated by the fact that there is no allocation where managed futures doesn’t reduce the volatility of the traditional portfolio. Though some may correctly point out that bonds are less volatile, the difference is only 200 basis points and bonds are underweighted in the portfolio, so they contribute less to the overall risk compared to stocks. As a result, the inclusion of bonds cannot meaningfully dampen the volatility reduction impact that managed futures has on the portfolio. Furthermore, managed futures has a relatively low correlation with stocks and bonds which further enhances its potential to reduce overall portfolio volatility. In conclusion, managed futures proved to be more effective at reducing volatility in traditional portfolios because it is less volatile than stocks which the portfolio was overweight in.

4.3.4 The Impact of Managed Futures on Portfolio Return
The results demonstrate that increasing allocations of managed futures improves returns for the traditional portfolio but lowers returns for a hedge fund portfolio (Table 11).
Table 11. Return Improvement After Addition of Managed Futures (1990 – 2016)

<table>
<thead>
<tr>
<th>Allocation</th>
<th>Traditional Portfolio</th>
<th>Hedge Fund Portfolio</th>
</tr>
</thead>
<tbody>
<tr>
<td>10MF-90 Traditional / HF</td>
<td>0.22%</td>
<td>-0.31%</td>
</tr>
<tr>
<td>20MF-80 Traditional / HF</td>
<td>0.46%</td>
<td>-0.61%</td>
</tr>
<tr>
<td>30MF-70 Traditional / HF</td>
<td>0.68%</td>
<td>-0.92%</td>
</tr>
<tr>
<td>40MF-60 Traditional / HF</td>
<td>0.91%</td>
<td>-1.23%</td>
</tr>
<tr>
<td>50MF-50 Traditional / HF</td>
<td>1.14%</td>
<td>-1.53%</td>
</tr>
<tr>
<td>60MF-40 Traditional / HF</td>
<td>1.37%</td>
<td>-1.84%</td>
</tr>
<tr>
<td>70MF-30 Traditional / HF</td>
<td>1.60%</td>
<td>-2.14%</td>
</tr>
<tr>
<td>80MF-20 Traditional / HF</td>
<td>1.83%</td>
<td>-2.44%</td>
</tr>
<tr>
<td>90MF-10 Traditional / HF</td>
<td>2.06%</td>
<td>-2.70%</td>
</tr>
</tbody>
</table>

Source: Bloomberg

Given that the Sharpe Ratio is simply excess returns divided by volatility, a greater improvement in the returns of the traditional portfolio versus the hedge fund portfolio’s should result in a higher Sharpe Ratio for the traditional portfolio, all else being equal. Portfolio returns are the weighted average of the returns of the underlying assets. Portfolios that allocate more to higher yielding assets will, correspondingly, have a higher return while as those that allocate more to lower yielding assets will produce lower returns. The most plausible explanation for why managed futures was more impactful on the traditional portfolio is that it had a superior return relative to stocks and bonds and an inferior return relative to the underlying hedge fund strategies.

Table 12. Return of Traditional and Alternative Assets Relative to Managed Futures (1998 – 2016)

<table>
<thead>
<tr>
<th>Convertible Arbitrage</th>
<th>Distressed Securities</th>
<th>Equity Long/Short</th>
<th>Equity Market Neutral</th>
<th>Event Driven Multi-Strategy</th>
<th>Fixed Income Arbitrage</th>
<th>Global Macro</th>
<th>Merger Arbitrage</th>
<th>Global Stocks</th>
<th>Global Bonds</th>
<th>Managed Futures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns</td>
<td>7%</td>
<td>1%</td>
<td>8%</td>
<td>6%</td>
<td>0%</td>
<td>6%</td>
<td>5%</td>
<td>6%</td>
<td>5%</td>
<td>5%</td>
</tr>
</tbody>
</table>

Source: Bloomberg

The results demonstrate that managed futures produces a higher return relative to both stocks and bonds, thereby explaining why its addition to the portfolio resulted in strong improvements in return. By contrast, managed futures generates a lower return in comparison to convertible arbitrage, distressed securities, equity long/short and event-driven multi-strategy. While it is true that managed futures has a higher return than equity market neutral, fixed income arbitrage, global macro and merger arbitrage, the difference is relatively small. For example, the average return difference between managed futures and the hedge fund strategies it outperforms is 92 basis points; by contrast, the average return difference between managed futures and stocks and bonds is over 200 basis points. As a result of having better returns than stocks and bonds, managed futures is able to unambiguously improve the returns of the traditional portfolio. By contrast, managed futures has a muted effect on the returns of the hedge fund portfolio as it has comparable returns to the underlying hedge fund strategies.

In summary, managed futures generated a statistically significant improvement in the Sharpe Ratio for the traditional portfolio but not the hedge fund portfolio because the strategy reduced portfolio volatility and improved returns by a greater magnitude in the former case. This was due to managed futures being 1) less volatile than stocks and bonds but more volatile than the underlying hedge fund strategies and 2) having higher returns than stocks and bonds but a relatively similar return compared with the underlying hedge fund strategies.

4.4.1 The Impact of S&P 500 DTI on Traditional Portfolios

I now turn to analyzing whether or not a semi-passive benchmark of managed futures — the S&P 500 DTI — will have a positive impact on the performance of traditional portfolios. Given that managed futures significantly improved performance of a traditional portfolio, I expect S&P 500 DTI to also have a positive but less impactful effect. As with the managed futures case, I test the portfolio impact at varying allocations of S&P 500 DTI in order to ensure the results are
sufficiently robust.


<table>
<thead>
<tr>
<th>Traditional Portfolio Allocation</th>
<th>S&amp;P 500 DTI Allocation</th>
<th>Traditional Portfolio Only Sharpe Ratio</th>
<th>Traditional DTI Portfolio Sharpe Ratio</th>
<th>Sharpe Ratio Difference</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>90%</td>
<td>0.1280</td>
<td>0.1062</td>
<td>-0.0228</td>
<td>0.57</td>
</tr>
<tr>
<td>20%</td>
<td>80%</td>
<td>0.1280</td>
<td>0.1272</td>
<td>-0.0008</td>
<td>0.50</td>
</tr>
<tr>
<td>30%</td>
<td>70%</td>
<td>0.1280</td>
<td>0.1466</td>
<td>0.0186</td>
<td>0.42</td>
</tr>
<tr>
<td>40%</td>
<td>60%</td>
<td>0.1280</td>
<td>0.1603</td>
<td>0.0323</td>
<td>0.34</td>
</tr>
<tr>
<td>50%</td>
<td>50%</td>
<td>0.1280</td>
<td>0.1654</td>
<td>0.0374</td>
<td>0.28</td>
</tr>
<tr>
<td>58%</td>
<td>42%</td>
<td>0.1280</td>
<td>0.1637</td>
<td>0.0357</td>
<td>0.24</td>
</tr>
<tr>
<td>60%</td>
<td>40%</td>
<td>0.1280</td>
<td>0.1627</td>
<td>0.0347</td>
<td>0.24</td>
</tr>
<tr>
<td>70%</td>
<td>30%</td>
<td>0.1280</td>
<td>0.1552</td>
<td>0.0272</td>
<td>0.22</td>
</tr>
<tr>
<td>80%</td>
<td>20%</td>
<td>0.1280</td>
<td>0.1459</td>
<td>0.0179</td>
<td>0.25</td>
</tr>
<tr>
<td>90%</td>
<td>10%</td>
<td>0.1280</td>
<td>0.1366</td>
<td>0.0086</td>
<td>0.35</td>
</tr>
</tbody>
</table>

*Source: Bloomberg*

The results demonstrate that the S&P 500 DTI improves performance in 7 out the 9 cases, making it less effective in comparison with managed futures which improved portfolio performance in all cases. Interestingly enough, the results show that overweighting the S&P 500 DTI (80% or 90% allocations) actually reduces risk-adjusted returns; by contrast, in the managed futures case, we observed a reduction in the magnitude of improvement when there was an overweight to managed futures, but there was no case where the addition of managed futures hurt portfolio performance. What is more important, however, is that none of the Sharpe Ratio improvements, in the case of the S&P 500 DTI, are statistically significant, while as in the managed futures case 8 out of 9 cases generated an improvement that was statistically significant.

As with the managed futures analysis, I use mean-variance optimization to determine the allocation of S&P 500 DTI that maximizes the portfolio Sharpe Ratio. In this case, the optimal allocation of S&P 500 DTI is 42%; yet even at this point the Sharpe Ratio improvement is not statistically significant. In summary, the core finding — that the addition of S&P 500 DTI does not add significant value to a traditional portfolio — is unsurprising as the S&P 500 DTI is an inferior derivate of managed futures and, by extension, one would expect it to be less effective at improving performance.

### 4.4.2: The Impact of S&P 500 DTI on Hedge Fund Portfolios

Having examined the impact of the S&P 500 DTI on traditional portfolios, I now turn to analyzing how its addition to a hedge fund portfolio impacts performance. Given that managed futures performs superiorly compared with the S&P 500 DTI yet failed to improve the Sharpe Ratio of the hedge fund portfolio, it is reasonable to expect that the S&P 500 DTI to also fail to improve the risk-adjusted returns of the hedge fund portfolio. I test the portfolio impact at varying allocations of S&P 500 DTI in order to ensure the results are sufficiently robust.

---

10 The Traditional Portfolio only portfolio Sharpe Ratio is different from the one in Table 8 because I use data from 2004 - 2016 in this table (rather than from between 1990 - 2016, as in Table 6) in order to make it comparable to the S&P 500 DTI which has data beginning from 2004.

<table>
<thead>
<tr>
<th>Hedge Fund Allocation</th>
<th>S&amp;P 500 DTI Allocation</th>
<th>HF Only Portfolio Sharpe Ratio</th>
<th>HF - DTI Portfolio Sharpe Ratio</th>
<th>Sharpe Ratio Difference</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>90%</td>
<td>0.2627</td>
<td>0.1103</td>
<td>-0.1523</td>
<td>0.89</td>
</tr>
<tr>
<td>20%</td>
<td>80%</td>
<td>0.2627</td>
<td>0.1318</td>
<td>-0.1246</td>
<td>0.86</td>
</tr>
<tr>
<td>30%</td>
<td>70%</td>
<td>0.2627</td>
<td>0.1696</td>
<td>-0.0931</td>
<td>0.81</td>
</tr>
<tr>
<td>40%</td>
<td>60%</td>
<td>0.2627</td>
<td>0.2031</td>
<td>-0.0596</td>
<td>0.73</td>
</tr>
<tr>
<td>50%</td>
<td>50%</td>
<td>0.2627</td>
<td>0.235</td>
<td>-0.0277</td>
<td>0.62</td>
</tr>
<tr>
<td>60%</td>
<td>40%</td>
<td>0.2627</td>
<td>0.2601</td>
<td>-0.0025</td>
<td>0.51</td>
</tr>
<tr>
<td>70%</td>
<td>30%</td>
<td>0.2627</td>
<td>0.2747</td>
<td>0.0122</td>
<td>0.43</td>
</tr>
<tr>
<td>79%</td>
<td>21%</td>
<td>0.2627</td>
<td>0.2781</td>
<td>0.0154</td>
<td>0.39</td>
</tr>
<tr>
<td>80%</td>
<td>20%</td>
<td>0.2627</td>
<td>0.278</td>
<td>0.0153</td>
<td>0.39</td>
</tr>
<tr>
<td>90%</td>
<td>10%</td>
<td>0.2627</td>
<td>0.2719</td>
<td>0.0092</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Source: Bloomberg

The results show that the S&P 500 DTI improves hedge fund performance only in 4 out of 10 cases, a clear difference from managed futures which improved the Sharpe Ratio in 7 out of 10 cases. As with managed futures, the Sharpe Ratio improvements at each allocation (including the optimized allocation of 21%) were not statistically significant at the 10%, 5% or 1% significance levels. Furthermore, the results show that allocating more than 30% of the hedge fund portfolio to the S&P 500 DTI actually erodes its diversification benefits, resulting in a Sharpe Ratio that is, in fact, lower than the pure hedge fund portfolio. The core finding — that the addition of S&P 500 DTI does not significantly improve the Sharpe Ratio of a hedge fund portfolio — is unsurprising, given that managed futures, which is actively managed and more sophisticated, also failed to improve risk-adjusted returns. However, what makes this result (and the result for the traditional portfolio) interesting is how dramatically less impactful the S&P 500 DTI was on the portfolios. In the next section, I will examine this result in further detail; in particular, I will seek to uncover why the S&P 500 DTI was unable to be as impactful as managed futures on the hedge fund and traditional portfolio.

4.4.3 Impact of S&P 500 DTI vs. Managed Futures on Portfolio Volatility

Given that the S&P 500 DTI is a benchmark for managed futures, the reasons for why it was more impactful on traditional portfolios compared to a hedge fund portfolio are the same as for why managed futures impacted each portfolio differently. This explanation is given in sections 5.3.3-5.3.4. Instead, the more interesting question that will subsequently be explored is why managed futures was more positively impactful in both the hedge fund and traditional portfolios compared with the S&P 500 DTI.

In comparison with managed futures, the S&P 500 DTI appears to have been more effective at reducing portfolio volatility for both the traditional portfolio and hedge fund portfolio (Tables 15 & 16)

\(^{11}\) The Hedge Fund only portfolio Sharpe Ratio is different from the one in Table 7 because I use data from 2004 - 2016 in this table (rather than between 1990 - 2016, as in Table 7) in order to make it comparable to the S&P 500 DTI which has data beginning from 2004.
Table 15. Comparison of Volatility Reduction to Traditional Portfolio from Addition of Managed Futures vs. S&P 500 DTI (2004 – 2016)

<table>
<thead>
<tr>
<th>Traditional Portfolio Allocation</th>
<th>Managed Futures/S&amp;P 500 DTI Allocation</th>
<th>Volatility Reduction - Managed Futures</th>
<th>Volatility Reduction - S&amp;P 500 DTI</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>90%</td>
<td>2.69%</td>
<td>3.14%</td>
</tr>
<tr>
<td>20%</td>
<td>80%</td>
<td>3.13%</td>
<td>3.78%</td>
</tr>
<tr>
<td>30%</td>
<td>70%</td>
<td>3.39%</td>
<td>4.19%</td>
</tr>
<tr>
<td>40%</td>
<td>60%</td>
<td>3.44%</td>
<td>4.30%</td>
</tr>
<tr>
<td>50%</td>
<td>50%</td>
<td>3.27%</td>
<td>4.10%</td>
</tr>
<tr>
<td>60%</td>
<td>40%</td>
<td>2.89%</td>
<td>3.63%</td>
</tr>
<tr>
<td>70%</td>
<td>30%</td>
<td>2.35%</td>
<td>2.92%</td>
</tr>
<tr>
<td>80%</td>
<td>20%</td>
<td>1.67%</td>
<td>2.06%</td>
</tr>
<tr>
<td>90%</td>
<td>10%</td>
<td>0.87%</td>
<td>1.07%</td>
</tr>
</tbody>
</table>

Source: Bloomberg

Table 16. Comparison of Volatility Reduction to Hedge Fund Portfolio from Addition of Managed Futures vs. S&P 500 DTI (2004 – 2016)

<table>
<thead>
<tr>
<th>Hedge Fund Portfolio Allocation</th>
<th>Managed Futures/S&amp;P 500 DTI Allocation</th>
<th>Volatility Reduction - Managed Futures</th>
<th>Volatility Reduction - S&amp;P 500 DTI</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>90%</td>
<td>-0.97%</td>
<td>-0.67%</td>
</tr>
<tr>
<td>20%</td>
<td>80%</td>
<td>-0.39%</td>
<td>0.01%</td>
</tr>
<tr>
<td>30%</td>
<td>70%</td>
<td>0.12%</td>
<td>0.60%</td>
</tr>
<tr>
<td>40%</td>
<td>60%</td>
<td>0.52%</td>
<td>1.05%</td>
</tr>
<tr>
<td>50%</td>
<td>50%</td>
<td>0.80%</td>
<td>1.34%</td>
</tr>
<tr>
<td>60%</td>
<td>40%</td>
<td>0.94%</td>
<td>1.44%</td>
</tr>
<tr>
<td>70%</td>
<td>30%</td>
<td>0.93%</td>
<td>1.34%</td>
</tr>
<tr>
<td>80%</td>
<td>20%</td>
<td>0.76%</td>
<td>1.04%</td>
</tr>
<tr>
<td>90%</td>
<td>10%</td>
<td>0.45%</td>
<td>0.59%</td>
</tr>
</tbody>
</table>

Source: Bloomberg

These results are particularly surprising as the addition of managed futures improved risk-adjusted returns of both the hedge fund and traditional portfolio by a greater magnitude than did the addition of S&P 500 DTI. Given that the results demonstrate that the S&P 500 DTI actually lowers portfolio volatility in both cases by a greater magnitude, managed futures’ superior improvement in risk-adjusted return must then be driven primarily by a stronger impact on portfolio return. One possible explanation for why the S&P 500 DTI was more effective at reducing portfolio volatility than managed futures is that it is less correlated with the underlying assets in both the hedge fund and traditional portfolio (Table 17).

Table 17. Correlation of Managed Futures and Traditional/Alternative Asset Classes vs. S&P 500 DTI and Traditional/Alternative Asset Classes (2004 – 2016)

<table>
<thead>
<tr>
<th>Convertible Arbitrage</th>
<th>Distressed Securities</th>
<th>Equity Long/Short</th>
<th>Equity Market Neutral</th>
<th>Event Driven Multi-Strategy</th>
<th>Fixed Income Arbitrage</th>
<th>Global Macro</th>
<th>Merger Arbitrage</th>
<th>Global Stocks</th>
<th>Global Bonds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managed Futures</td>
<td>-0.74%</td>
<td>1.29%</td>
<td>26.15%</td>
<td>22.63%</td>
<td>2.81%</td>
<td>-12.94%</td>
<td>70.46%</td>
<td>5.22%</td>
<td>7.29%</td>
</tr>
<tr>
<td>S&amp;P 500 DTI</td>
<td>-27.70%</td>
<td>-34.93%</td>
<td>0.62%</td>
<td>-16.91%</td>
<td>-11.17%</td>
<td>-24.29%</td>
<td>40.94%</td>
<td>-14.49%</td>
<td>-12.70%</td>
</tr>
</tbody>
</table>

Source: Bloomberg

This indeed turns out to be correct as the results demonstrate that the S&P 500 DTI has a lower correlation with stocks, bonds and the various hedge fund strategies in comparison with managed futures. Since managed futures is relatively more correlated with the underlying assets in both portfolios, its inclusion diversified away less risk, allowing the S&P 500 DTI to be more impactful in lowering volatility in both instances. Aside from correlation, the volatility of the underlying assets in a portfolio is also a major driver of overall portfolio volatility (Figure 3).

Figure 3. Rolling 12 Month Volatility of Managed Futures & S&P 500 DTI (2004 – 2016)
The results demonstrate that the volatilities between managed futures and S&P 500 DTI are very similar, with some divergence during periods of very high market stress such as during the credit crisis in 2007-2008 and the commodity price collapse in 2014. Since the volatilities are so similar for both managed futures and the S&P 500 DTI, it is unlikely that differences asset volatility can explain why managed futures was less effective at reducing portfolio volatility. Instead, the discrepancy seems to be driven by the fact that the S&P 500 DTI is less correlated than managed futures with the underlying assets in both the hedge fund and traditional portfolio.

### 4.4.4 Impact of S&P 500 DTI vs. Managed Futures on Portfolio Return

In comparison with the S&P 500 DTI, managed futures appears to have been more effective at improving returns for both the traditional and hedge fund portfolios (Tables 18 & 19).

#### Table 18. Return Improvement with Managed Futures vs. S&P500 DTI for a Traditional Portfolio (2004 – 2016)

<table>
<thead>
<tr>
<th>Traditional Portfolio Allocation</th>
<th>Managed Futures/S&amp;P 500 DTI Allocation</th>
<th>Return Improvement - Managed Futures</th>
<th>Return Improvement - S&amp;P 500 DTI</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>90%</td>
<td>-1.15%</td>
<td>-2.00%</td>
</tr>
<tr>
<td>20%</td>
<td>80%</td>
<td>-1.02%</td>
<td>-1.78%</td>
</tr>
<tr>
<td>30%</td>
<td>70%</td>
<td>-0.90%</td>
<td>-1.56%</td>
</tr>
<tr>
<td>40%</td>
<td>60%</td>
<td>0.77%</td>
<td>-1.34%</td>
</tr>
<tr>
<td>50%</td>
<td>50%</td>
<td>-0.64%</td>
<td>-1.12%</td>
</tr>
<tr>
<td>60%</td>
<td>40%</td>
<td>-0.31%</td>
<td>-0.89%</td>
</tr>
<tr>
<td>70%</td>
<td>30%</td>
<td>0.38%</td>
<td>-0.67%</td>
</tr>
<tr>
<td>80%</td>
<td>20%</td>
<td>0.26%</td>
<td>-0.45%</td>
</tr>
<tr>
<td>90%</td>
<td>10%</td>
<td>0.13%</td>
<td>-0.22%</td>
</tr>
</tbody>
</table>

*Source: Bloomberg*

#### Table 19. Return Improvement with Managed Futures vs. S&P500 DTI for a Hedge Fund Portfolio (2004 – 2016)

<table>
<thead>
<tr>
<th>Hedge Fund Portfolio Allocation</th>
<th>Managed Futures/S&amp;P 500 DTI Allocation</th>
<th>Return Improvement - Managed Futures</th>
<th>Return Improvement - S&amp;P 500 DTI</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>90%</td>
<td>-0.09%</td>
<td>-3.25%</td>
</tr>
<tr>
<td>20%</td>
<td>80%</td>
<td>-0.08%</td>
<td>-2.89%</td>
</tr>
<tr>
<td>30%</td>
<td>70%</td>
<td>-0.07%</td>
<td>-2.53%</td>
</tr>
<tr>
<td>40%</td>
<td>60%</td>
<td>-0.06%</td>
<td>-2.18%</td>
</tr>
<tr>
<td>50%</td>
<td>50%</td>
<td>-0.05%</td>
<td>-1.82%</td>
</tr>
<tr>
<td>60%</td>
<td>40%</td>
<td>-0.04%</td>
<td>-1.46%</td>
</tr>
<tr>
<td>70%</td>
<td>30%</td>
<td>-0.03%</td>
<td>-1.09%</td>
</tr>
<tr>
<td>80%</td>
<td>20%</td>
<td>-0.02%</td>
<td>-0.73%</td>
</tr>
<tr>
<td>90%</td>
<td>10%</td>
<td>-0.01%</td>
<td>-0.37%</td>
</tr>
</tbody>
</table>

*Source: Bloomberg*
Given that managed futures had a relatively mediocre small impact on portfolio volatility, it had to have had a superior impact on portfolio returns in order to have generated comparatively larger improvements in portfolio Sharpe Ratio. It is worth noting how much more effective managed futures was in improving returns in the traditional and hedge fund portfolios. In particular, there was no allocation of S&P 500 DTI in either the traditional or hedge fund portfolios which improved portfolio returns; in fact, portfolio returns in both cases go down as the allocation to S&P 500 DTI is raised. By contrast, the addition of managed futures improves the return of the traditional portfolio in every instance, and portfolio returns increase as the allocation to managed futures is raised. However, in the hedge fund case, the addition of managed futures reduces return in every case, yet by a far lesser degree than the S&P 500 DTI.

Since portfolio return is effectively the weighted average of the returns of the underlying assets, the most likely explanation for why managed futures had a better impact on portfolio return in both cases was because it was a higher yielding strategy (Figure 4).

**Figure 4.** Rolling 12 Month Returns of Managed Futures and S&P 500 DTI (2004 – 2016)

Indeed, the results demonstrate that managed futures tends to have higher returns than the S&P 500 DTI as there is rarely an extended period of time where the returns of managed futures are lower than those of the S&P 500 DTI. Most likely the superior return from managed futures is a result of the strategy’s sophistication and active management, features that the S&P 500 DTI lacks as it utilizes a very simplistic trading strategy to make investments. In the end, managed futures has a demonstrably superior return to S&P 500 DTI and, as a result, positively impacts portfolio return by a greater magnitude than does the S&P 500 DTI.

5. Conclusion

In its evaluation of the performance of managed futures, this paper examined whether or not managed futures could produce alpha, time markets and positively improve the risk-adjusted returns of traditional and hedge fund portfolios. The key conclusions are summarized below:

Managed futures was able to generate positive, statistically significant alpha when exposure (both long and short) to equity, bond, currency and commodity markets was controlled for. By contrast, the S&P 500 DTI, a semi-passive managed futures index, failed to produce significant alpha. This result suggests that there is some positive value that comes from active management and the sophisticated models that managed futures managers use.

Managed futures proved to be successful in timing commodity and equity markets but was unsuccessful in timing bond and currency markets. This discrepancy can be explained by the structural differences between commodity and equity markets versus bond and currency markets. Specifically, currency and bond markets tend to be less cyclical, providing
managers fewer opportunities to time markets. In addition, the results demonstrated that the S&P 500 DTI had identical results in market timing skill, suggesting that timing skill was more dependent on the ability to go long and short based on directional trends rather than the sophistication of the trend following models.

The addition of managed futures to a traditional 60/40 portfolio resulted in a statistically significant improvement in risk-adjusted returns. However, the diversification benefits did not carry onto over to the hedge fund portfolio in which the performance improvement from the addition of managed futures was not statistically significant. This was largely because managed futures proved to be less effective at lowering portfolio volatility and improving returns for the hedge fund portfolio than for the traditional portfolio.

By contrast, the S&P 500 DTI failed to improve risk-adjusted returns of both the hedge fund and traditional portfolios. This was largely due to the fact that the S&P 500 DTI had significantly lower returns compared with managed futures and its addition to both the hedge fund and traditional portfolio had a negative impact on portfolio return. The S&P 500 DTI had a slightly better impact on reducing portfolio volatility compared with managed futures. Overall, the S&P 500 DTI’s ineffectiveness as a diversifier of both hedge fund and traditional portfolios further highlights the portfolio benefits of an actively managed strategy such as managed futures.

I believe the results of my research will strengthen and corroborate the growing interest and appreciation for managed futures in the investment world. I also hope that it serves as an effective introduction to managed futures and addresses the concerns people have about the asset class. Ultimately, portfolio construction, managed futures and alternatives are extremely broad topics with incredible potential for future research. For example, it would be interesting to study other alternative investments such as those in private equity or real estate and test their impacts on hedge fund and traditional portfolios. Though this paper examined differences in portfolio Sharpe Ratios, an interesting follow-up to this research would be re-testing the impact of managed futures on hedge fund and traditional portfolios with a different metrics of risk-adjusted-return such as the Sortino Ratio and Calmar Ratios. A separate angle of study would be to examine whether the impact that managed futures has on both the traditional and hedge fund portfolios can be improved by using different asset allocation techniques such as risk parity. As managed futures grows in scope and popularity and investors increasingly seek better diversification, these questions will certainly be important topics for retail and institutional investors to consider when deciding whether or not to invest in the strategy.

---

12 As is typical in empirical studies, the robustness of the results is limited to the timeframe of the data used. That is, the results presented in this study could change if data from a different time frame is used.
6. References
Black, K, narrator. Global Macro and Managed Futures. CAIA Fundamentals of Alternative Investments
What are the Effects of Specific Trade Agreement Provisions on Bilateral Trade Flows?

DUNCAN HOBBS
Georgetown University

ABSTRACT
This paper examines the effects of various trade agreement provisions on bilateral trade flows between member countries. I utilize the Design of International Trade Agreements Dataset which codes over 900 trade agreements for several key provisions and estimate the effects of these provisions using expanded gravity models on a panel of over 500,000 bilateral trade flows ranging from 1962-2015. To deal with the well documented problems of omitted variable and selection bias in cross-section OLS estimates, I employ importer exporter and time fixed effects in panel regressions to estimate the average treatment effect of these agreement features. I find that membership in a trade agreement on average more than doubles bilateral trade between members, a services provision in a trade agreement increases trade by an average of 84 percent, and provisions regulating procurement, product standards and fair competition reduce trade by 30, 47 and 116 percent respectively. I also turn to nonlinear estimation methods including Heckman selection models in an attempt to deal with selection bias. While the effect sizes of provisions are larger using nonlinear estimators, the direction of the effects remain the same, and the standard errors are even smaller reinforcing the results obtained from panel estimators.
The Cost of Conflict:
Refugee Flows in Sub-Saharan Africa

GABRIELLE COHEN
University of Cape Town

ABSTRACT
This paper examines the cost of refugee spillover to economic growth in both the host country and the neighboring countries of conflict in sub-Saharan Africa for the years 1990 to 2010. Using a Solow growth model augmented for human capital, it becomes clear that refugee spillover is a cost to both the neighboring countries of conflict and the host country itself. The implications of these results indicate that spillover costs should be considered in terms of neighboring countries and the secondary effects that they can have on the host of the conflict. Additionally, it is necessary to consider refugee flows as a cost to the host country; the country of origin and the neighboring countries; and the recipients of refugees seeking asylum.
1. Introduction
This paper endeavors to establish the economic costs of refugee flows as a spillover effect of conflict in Sub-Saharan Africa. Global refugee flows are a common outcome of both interstate and intrastate conflict and have become an international concern in recent years. Quantifying and determining the costs of conflict is integral in establishing appropriate and effective post-conflict reconstruction and aid policy. This is particularly important in developing regions that are susceptible to extended and recurring periods of conflict such as Sub-Saharan Africa.

Literature on identifying and quantifying the costs of conflict includes works by Collier (1999; 2004), who identifies costs of conflict to both the host country and the possible spillover effects, and Murdoch and Sandler (2002) and Dunne and Tian (2015), who quantify spillover effects. Furthermore, the flows of refugees and their costs and effects within the international system are investigated by Akokpari (1998) and Gleditsch and Salehyan (2006). However, a clear and cohesive narrative of refugee flows as a spillover and economic cost of conflict for host countries and neighboring countries is missing. This paper aims to contribute to conflict literature by attempting to create a clear narrative of the economic impact of refugee flows as a costs of conflict on the economic growth of the host of the conflict and its neighboring countries. This is attempted by considering refugee flows both by country of asylum and by country of origin in order to establish a comprehensive understanding of refugee flows within the region for the given time period. Two main hypotheses are put forth in this paper; the first is that spillover of conflict should be considered as a cost to neighboring countries as well as the host over and above the costs of the conflict itself to the host. Additionally, refugee flows, while traditionally considered a cost to recipient countries such as contiguous neighbors, ought to be considered as a cost to the host of the conflict or country of origin as well.

The following section considers the existing literature on the costs of conflict, the spillover effects of conflicts, and the potential costs of refugee flows. Section three and four present the theoretical framework and data and empirical methods respectively while the empirical results are discussed in section five. The conclusion of the paper is then presented in the sixth and final section.

2. Literature Review
Cost of conflict literature identifies the economic costs to the host country and its neighbors. Various authors have conducted investigations identifying the many costs incurred by conflict and the methods by which its impact can be tested and measured. De Groot et al (2009) highlight the importance of considering both host country costs and spillover costs in calculating the true economic costs of a conflict. Bozzoli et al (2010) define costs of conflict as a measure of changes in welfare owing to the conflict. Collier (1999; 2004) introduces the importance of spillover in the costs of conflict calculation. Spillover refers to costs that arise as a result of conflict that commonly hinder the welfare of the neighbors to the country hosting the conflict. To provide a consistent and comprehensive analysis of the economic costs of conflict, spillover must be considered (de Groot, Brück and Bozzoli, 2009). The widely accepted "stylized fact" of the costs of conflict is that violent conflict reduces real GDP by up to 2.2 percentage points (Bozzoli, Brück and Sottsas, 2010).

Hoeffler (2003), Collier (1999; 2004) and Brauer and Dunne (2010) highlight the main economic costs of conflict which is then considered with the closely aligned possible spillover effect. First, the diversion of resources towards increased military spending strains economic growth in the host country by reducing public expenditure and thus welfare and additionally strains growth through the destruction of infrastructure and resources due to increased military activity (Hoeffler, 2003). This cost can further be understood as a stream of spillover as conflict management for neighboring countries requires the reallocation of resources to conflict containment, security, border control and other less-productive activities that are required due to conflict (Dunne and Tian, 2015). Additionally, conflict disrupts trade in the host country and the surrounding countries providing a challenge to supply lines (Brauer and Dunne, 2010).

Collier (1999) draws attention to the portfolio substitution effect induced in a host by conflict as a significant cost to the host country that is not improved post conflict (Collier, 1999). The portfolio substitution effect refers to the reallocation of assets out of the host country as a result of conflict. Physical capital is relocated where possible, maintenance expenditure is reduced and financial capital is relocated. Human capital is relocated through migration due to fear and expenditure
on education which is significantly reduced during conflict (Collier, 1999). The outmigration of human capital presents a loss to local economies and local production (Brauer & Dunne, 2010). Additionally, human capital is further destroyed due to death, declining public health and deterioration in the accumulation of skills and workplace experience (Brauer and Dunne, 2010).

Capital destruction can also be understood as a spillover effect of conflict. Spillover can adversely impact the capital stocks of neighboring and host countries through various streams. The first is the destruction of physical capital stocks found to have a negative impact on both the host of the conflict and primary neighbors. The second stream is labor and human capital. Dunne and Tian (2015) posit that theoretically, destruction and displacement of productive labor is likely for both the host and the neighbors of the host (Dunne & Tian, 2015). Additionally, concern arises surrounding refugee flows, particularly for primary neighbors due to the potential for conflict diffusion and additional management and security costs however it is noted that this may, at least in part, be offset by an inflow of human and physical capital.

The economic costs of conflict can alternatively be understood through their sector-specific effects. The costs of transactions and transportation increase during conflict due to the destruction of infrastructure such as roads and bridges and the dysfunction of institutions such as security that lead to economic loss in production (Collier, 1999). Additionally, all sectors that rely on or produce physical and human capital are expected to incur losses as a result of conflict. Decreased demand through expenditure allocation to physical and human capital and decreased supply of capital due to capital flight severely hinder the productive capacities of capital-intensive sectors of the economy (Collier, 1999). This strain on factors of production provides evidence that conflict reduces the growth rate and the actual level of GDP. Additionally, capital losses can continue to occur post conflict despite the restoration of peace (Collier, 1999).

Collier (1999) models the costs of conflict on host countries and neighboring countries using a neoclassical Solow growth model augmented to include human capital. Furthermore, Murdoch and Sandler (2002; 2004) and Dunne and Tian (2015) take account of the importance of border length and accessibility in establishing spillovers and their intensity for neighboring countries (Murdoch and Sandler, 2002). Civil war is found to have small, negative impacts on the steady state levels of GDP per Capita for both the host and neighbors of the conflict which is found to occur through disruption of the economy (Murdoch and Sandler, 2002; Dunne and Tian, 2015). However, spillover research finds that different spillover effects can have different signs suggesting uncertainty about the short run spillover effects of conflict (De Groot, 2010).

Refugee flow is both a direct cost and spillover of conflict that impacts the host of the conflict, or country of origin, as well as neighboring and recipient countries. The economic costs of refugee flows can be understood in terms of the country of origin and the recipient country. It is of particular interest because of its intersectional nature within the cost of conflict classification. However, existing literature fails to address the refugee flows together and disaggregated by country of origin and country of asylum.

The flight of asylum seekers out of the country of origin are both a direct cost and a spillover of conflict. The flight of refugees from their country of origin poses severe economic and security threats to the country of origin (Akokpari, 1998) and threatens the industry of the origin country (Collier, 1999). The outflow of refugees from the country of origin may be a loss of professionals and educated members of the population thus suggesting a loss in human capital that can further provide a challenge to economic growth (Akokpari, 1998). Additionally, the significant outflow of human capital, due to flight, commonly induces a contraction in human capital intensive industries within the host country (Collier, 1999). Commonly the outflow of refugees leads to disruption in economic production, especially in the agricultural sector that commonly leaves the country of origin more vulnerable to threats such as economic depressions and food shortages (Akokpari, 1998). One can further argue that if there is a flight of human capital through migration, there will be significant financial and physical capital losses to the host country of the conflict as refugees take their assets out of the host country contributing to the additional strain to economic growth in the host country. Finally, remittances are commonly sent back to the country of origin by refugees that may be used to fund rebel movements and thus may pose an additional security cost to the country of origin (Koser and Van Hear, 2003).
In the context of Sub-Saharan Africa particularly, neighboring countries face great challenges as a result of refugees seeking asylum as they commonly lack the infrastructure and resources to adequately accommodate their own populations (Akokpari, 1998). Employment and environmental strains are considered as significant outcomes of refugee spillover. The lack of employment opportunities for refugees commonly provokes the development of a refuge-run black market in the recipient country. Additionally, the environmental strain of refugee camps act as a burden on refugee hosts countries as refugees deplete environmental resources such as firewood and livestock (Akokpari, 1998). Refugees seeking asylum impose significant economic costs on neighboring countries. Economic burdens include requirements of assistance, competition for scarce resources and additional maintenance costs that the recipient country would otherwise not have faced (Salehyan, 2008). Additionally, refugees pose threats to public health commonly having experienced some trauma that requires medical attention as is illustrated by the increased rate of morbidity and mortality for the recipient country (Salehyan, 2008). Furthermore, refugee camps and transitory housing is commonly overcrowded and conducive to the spread of disease. This results in competition and a diversion of health resources that otherwise would have been used on the citizens of the recipient country. Refugees further provide a threat to the security of the recipient country by challenging the ethnic balance within the country, introducing new ideologies within the country and increasing tensions between the host of the conflict and the recipient of the refugees increasing the probability that conflict will occur between the two countries (Salehyan, 2008).

Refugee spillover may increase the probability of conflict within the recipient countries by changing the ethnic composition of the recipient country, spreading radical ideologies, arms and even organized terrorist networks and adding additional pressure to economic competition (Salehyan and Gleditsch, 2006). Salehyan and Gleditsch (2006) consider a case study of Rwandan refugees spilling over into the Democratic Republic of Congo, then Zaire, and how their presence catalyzed conflict not only in the Democratic Republic of Congo but also in Burundi and Uganda (Salehyan and Gleditsch, 2006). Author Akokpari (1998) draws attention to refugees as a security threat to their recipient country due to the risk of continued rebel-action within the recipient country by the refugees. Akokpari draws on the case of armed Hutus in Zaire in 1996 who orchestrated attacks on Rwanda from within the refugee camps of Eastern Zaire (Akokpari, 1998). These cases reinforce the argument that refugees can increase the chance of conflict in the recipient country by encouraging outside involvement in the conflict and rebel attacks through mercenaries or by straining relationships between the recipient government and the government of the country of origin (Akokpari, 1998).

Finally, integral to the sub-Saharan African context, Crisp (1999) provides insight into the major flaws in refugee data. Available refugee data is currently wrought with missing information and measurement errors. This uncertainty and unreliability surrounding African refugee statistics can be attributed to the lack of homogenous definitions of a refugee as a barrier to reliable refugee statistics; as well as artificially low statistics due to fear of host country retaliation and resource scarcity (Crisp, 1999). This is an integral caveat to this analysis of the economic impact of refugee spillover in Sub-Saharan Africa as it highlights the limitations to the data available and provides the author with a precursory caution which is necessary to consider to ensure the thoroughness and validity of this cost of conflict analysis.

Significant discourse exists around refugees and their economic impacts however, no cohesive narrative or exhaustive quantification of the economic costs of refugee flows currently exists. Thus it is necessary to contribute to conflict literature by creating a clear narrative of refugee flows in order to establish their true cost as a spillover effect of conflict on both the host country and neighboring countries of the conflict.

3. Theoretical Framework

Following the literature, a basic Solow model, augmented for human capital, will be used to estimate economic growth (Mankiw et al, 1990). The investigation conducted by Mankiw, Romer and Weil inform this method of human capital integration into the Solow model as they find that the inclusion of human capital in the Solow growth model improves the model's performance in estimating economic growth (Mankiw et al, 1990). The basic model uses the Cobb-Douglas production function for diminishing returns in inputs of production: labour (L), physical capital (K) and human capital (H). Main features of this model include constant returns to scale and in steady state savings equals investment in human and physical capital. For the purposes of this analysis the model will be augmented to include conflict and additionally...
introduce refugee flows as independent variables included through the technology parameter \((A)\).

The human capital augmented production function featuring a technical progress can be written as:

\[
Y(t) = K(t)^\alpha H(t)^\beta [A(t)L(t)]^{1-\alpha-\beta}, \ 0 < \alpha + \beta < 1 \quad (1)
\]

\(Y(t)\) is the aggregate real output at time \(t\), \(k(t)\) is the real stock of physical capital, \(H(t)\) represents human capital at time \(t\) and \(A(t)\) and \(L(t)\) are the technology parameter and labour parameter at time \(t\) respectively. \(\alpha\) indicates the elasticity of output with respect to physical capital, \(\beta\) indicates the elasticity of output with respect to human capital and finally the output elasticity of effective units of labour \(A(t)L(t)\) is indicated by \(1-\alpha-\beta\). Finally, the model makes the assumption that the rate of depreciation of physical and human capital is represented by \(\delta\) while labour grows at an exogenous continuous rate, \(n\) and technological progress at rate \(g\).

To perceive the equation in terms of income per effective worker it is necessary to divide equation (1) by effective labour \((AL)\) thus:

Given that \(y = Y/AL\)

\[
y(t) = k(t)^\alpha h(t)^\beta
\]

with \(k = K/AL\) and \(h = H/AL\) in quantities of physical and human capital per effective worker at time \(t\), respectively.

Given the defined rates of growth of the factors, the equations of motion for labour \((L)\), labour-augmenting technology \((A)\), physical \((K)\) and human capital \((H)\) are:

\[
L(t) = nL(t) \text{ and } A(t) = gA(t)
K(t) = s_kY(t) - \delta K(t) \quad \text{and} \quad H(t) = s_hY(t) - \delta K(t)
\]

The transition equations for physical and human capital that accounts for the growth of capital per capita \((k)\) and human capital per capita \((h)\) are:

\[
k(t) = s_kY(t) - (n+g+\delta)k(t)
h(t) = s_hY(t) - (n+g+\delta)h(t)
\]

Steady state equations where \(k(t) = h(t)\):

\[
k^* = (s_k s_h n + g + \delta)^{1/(1-\alpha-\beta)}
h^* = (s_k s_h n + g + \delta)^{1/(1-\alpha-\beta)}
y^* = (s_k s_h n + g + \delta)^{1/(1-\alpha-\beta)} \left( \frac{s_k}{s_h} \right)^{\alpha/(1-\alpha-\beta)}
\]

By taking logs of the equation and linearising via a truncated Taylor series expansion around the steady-state, substituting in \(y^*\) and setting \(A\) equal to a constant \(a\), one is able to obtain the growth of income. When reconstructed in accordance with Mankiw et al (1992) this represents the long run steady state growth of income per capita as:

\[
y(t) = \beta_0 + \beta_1 \ln(y_0) + \beta_2 \ln(s_k) + \beta_3 \ln(s_h) + \beta_4 \ln(n+g+\delta) \quad (2)
\]

The equation above represents the Solow model augmented for human capital. In accordance with the Solow model, \(y(t)\) is the growth rate of income per capita, \(y_0\) the initial level of income per capita, \(s_k\) and \(s_h\) the levels of physical and human capital respectively and finally \(n+g+\delta\) is measured as an amalgamated variable consisting of the growth rate of the population, the growth rate of technological progress and the rate of depreciation (Dunne & Tian, 2015).

Solow growth theory suggests that economic growth measured as income per capita will be negatively impacted by \((n+g+\delta)\)
and the initial level of income $y_0$ but positively impacted by investment in physical and human capital accumulation; measured by investment and education respectively in this analysis. Lower initial levels of income have been found to allow for faster growth while the term $(n+g+\delta)$ captures increases in labour growth and labour efficiency as well as depreciation thus explaining its negative impact. Investment in physical and human capital allow these negative effects to be offset.

$$\Delta \ln y = \alpha_0 + \beta_1 \ln(y_{0,i}) + \beta_2 \ln(s_k_{i,t}) + \beta_3 \ln(s_h_{i,t}) + \beta_4 \ln(n+g+\delta) + \beta_5 \ln(y_{i,t}) + \beta_6 W_{pri} \ln(conf_{pri,i,t}) + \epsilon_i,t$$

Equation (3) represents the Solow model augmented for human capital including conflict measures as used by Dunne and Tian (2015), weighted for primary neighbors. Primary neighbor contiguity matrices are constructed using a dummy variable approach and a border length approach take into consideration contiguity and the degree of contiguity between countries (Dunne and Tian, 2015). These matrices are constructed by dividing the border distance between countries by the sum of all primary neighbors and the total distance of the host country's border length (Dunne and Tian, 2015).

$$\Delta \ln y = \alpha_0 + \beta_1 \ln(y_{0,i}) + \beta_2 \ln(s_k_{i,t}) + \beta_3 \ln(s_h_{i,t}) + \beta_4 \ln(n+g+\delta) + \beta_5 \ln(conf_{i,t}) + \beta_6 W_{pri} \ln(conf_{pri,i,t}) + \beta(refugee_{a,i,t}) + \beta(refugee_{o,i,t}) + \epsilon_i,t$$

Equation (4) represents the Solow model augmented for human capital, as specified in equation (3) including refugees by country of origin and refugees by country of asylum. The significance of the paper lies in the addition of refugee flows as a determinant for economic growth.

4. Data and Empirical Methods
The dataset used to conduct this analysis is a combination of variables from the Penn World Tables Version 7, Education data from Barro and Lee and Penn World Tables Version 8 and the measures of armed conflict used in this analysis comes from the UCDP/PRIO Armed Conflict Database 4, updated to 2010 by Themner and Wallensteen as used by Tian and Dunne (Dunne and Tian). Additionally, a measure of refugee by country or territory of asylum is used from the World Bank Databank. The weights used in this analysis are provided by Nan Tian and Prof. JP Dunne and were constructed using CIA World Factbook, Google Earth and Gleditsch and Ward (2001)'s minimum distance dataset. Three conflict dummy variables are used to measure conflict: conflict, intensity and civil. Conflict includes all conflicts including civil and interstate conflicts. Civil refers only to the presence of intrastate conflict and finally intensity refers to conflicts that experience a minimum of 1000 conflict related deaths per year (Dunne and Tian, 2015).

Data availability dictates the time period and the sample of countries within sub-Saharan Africa used in this analysis. As indicated by Crisp (1999) various issues surround adequate data in Africa. Thus a caveat should be considered in terms of the potential of missing data and measurement errors. As a result, this analysis will first be conducted from 1960-2010 and thereafter from 1990 to 2010 due to limited data on refugee flows in Sub-Saharan Africa.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variable Description</th>
<th>Mean 1960-2010</th>
<th>Std. Dev</th>
<th>Mean 1990-2010</th>
<th>Std. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>rgdp_pc</td>
<td>Real GDP per capita</td>
<td>1291</td>
<td>1605</td>
<td>1498</td>
<td>1850</td>
</tr>
<tr>
<td>A (Inv)</td>
<td>Investment as a share of GDP or the change in Capital</td>
<td>16.62</td>
<td>10.70</td>
<td>17.26</td>
<td>8.75</td>
</tr>
<tr>
<td>edu</td>
<td>% of secondary education attained for population 25 years and older</td>
<td>11.93</td>
<td>12.64</td>
<td>19.89</td>
<td>14.43</td>
</tr>
<tr>
<td>pop</td>
<td>Population in 1000s</td>
<td>10933</td>
<td>12839</td>
<td>15147</td>
<td>15976</td>
</tr>
<tr>
<td>conflict</td>
<td>Conflict Indicator</td>
<td>0.17</td>
<td>0.37</td>
<td>0.21</td>
<td>0.41</td>
</tr>
<tr>
<td>intensity</td>
<td>Intense conflict indicator</td>
<td>0.06</td>
<td>0.23</td>
<td>0.05</td>
<td>0.23</td>
</tr>
<tr>
<td>civil</td>
<td>Civil war indicator</td>
<td>0.13</td>
<td>0.33</td>
<td>0.17</td>
<td>0.37</td>
</tr>
</tbody>
</table>
Table 1 indicates summary statistics for the 36 sub-Saharan African countries included in the analysis for the years 1960-2010. As indicated by the summary statistics in column 1 and 2 for the full sample, average income per capita is $1290 and 11.93% of the population above 25 years old has a secondary education. Additionally, as found by Dunne and Tian, 17% of all observations indicate a period of conflict with the majority of that conflict indicated as “intense”. The average values of the dummy variables for the presence of conflict, intense conflict and intrastate disputes indicate their respective prevalence in the sub-Saharan region.

For the reduced sample, 1990-2010, the average change in capital as a share of the GDP and the percentage of society above the age of 25 years old which has attained a secondary education is marginally higher than in the case of the full sample. This indicates higher levels of investment in capital and education within this period. However, the growth rates of real GDP, change in capital and education are notably lower during this period. Finally, the percentage of countries experiencing conflict is significantly larger in the reduced sample; 21% of all observations fall under periods of conflict with 80% of those conflicts being civil in nature. This revised framework indicates that this period is characterized by more conflict on average and slower rates of growth in education and investment in capital which is not uncommon in conflict regions.

5. Empirical Results
In order to test the impact of refugee flows as a form of spillover that impacts both the country of origin and country of asylum it is necessary to consider a dynamic growth model in the presence of conflict as a benchmark. The empirical results can be considered in three separate parts. First, a comparison is conducted of the dynamic regression results with contiguity and geographical weights for the full and reduced sample. Second, refugee flows by country of origin and country of asylum are introduced and analyzed and finally the statistical relationship between conflict and refugee flows is investigated.

For the purpose of ensuring an accurate benchmark, a dynamic regression including the contiguity and geographical weights is conducted for both the full sample (1960-2010) and the reduced sample (1990-2010). This provides a base that can be augmented to fully capture the flows of refugees and ensures that the results between samples are consistent with one another. Regression results (1-3) of table 2 indicates results for the full sample while regressions (4-6) indicate results for the reduced sample. The rate of growth of capital and the change capital has a positive economic and significant effect on income as a share of GDP for the full sample but is notably less significant for the reduced sample. Initial levels of income take on the expected sign for both samples and are statistically significant1.

---
1 Additionally, two empirical effects occur for both samples that contradict the theory of the Solow model: education, the measure for human capital, has a negative effect on income and population plus 0.05 has a positive effect on income as per the empirical results. The latter is considered a common result in developing countries and low income countries. The negative result of education on economic growth can be understood through the varying conclusions of Murdoch and Sandler (2002) and De Groot (2010). The authors note that the direction and significance of education may be due to the fact that no real growth in human capital has occurred or potentially that any increase in education has not truly contributed to increased productivity (Dunne and Tian, 2015).
Table 2. Dynamic Regression including Contiguity and Geographical Weights for Full Sample and Reduced Sample

<table>
<thead>
<tr>
<th>Conflict: Sample:</th>
<th>Δln Conflict</th>
<th>Δln Intense</th>
<th>Δln Civil</th>
<th>Δln Conflict</th>
<th>Δln Intense</th>
<th>Δln Civil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δln(A)</td>
<td>0.0348***</td>
<td>0.0335***</td>
<td>0.0346***</td>
<td>0.0090</td>
<td>0.0084</td>
<td>0.0092</td>
</tr>
<tr>
<td></td>
<td>(6.13)</td>
<td>(5.91)</td>
<td>(6.10)</td>
<td>(1.00)</td>
<td>(0.93)</td>
<td>(1.02)</td>
</tr>
<tr>
<td>Δln(edu)</td>
<td>-0.0762*</td>
<td>-0.0740*</td>
<td>-0.0761*</td>
<td>0.0379*</td>
<td>-0.211</td>
<td>-0.234*</td>
</tr>
<tr>
<td></td>
<td>(-2.55)</td>
<td>(-2.48)</td>
<td>(-2.55)</td>
<td>(-1.81)</td>
<td>(-2.02)</td>
<td></td>
</tr>
<tr>
<td>ln(n+g+G)</td>
<td>0.0582***</td>
<td>0.0565***</td>
<td>0.0612**</td>
<td>0.0791**</td>
<td>0.0769**</td>
<td>0.0799**</td>
</tr>
<tr>
<td></td>
<td>(5.93)</td>
<td>(5.78)</td>
<td>(6.28)</td>
<td>(6.08)</td>
<td>(6.31)</td>
<td></td>
</tr>
<tr>
<td>ln(y0)</td>
<td>-0.0280***</td>
<td>-0.0274***</td>
<td>-0.0281***</td>
<td>-0.0932***</td>
<td>-0.0961***</td>
<td>-0.0935***</td>
</tr>
<tr>
<td></td>
<td>(-6.00)</td>
<td>(-5.92)</td>
<td>(-5.99)</td>
<td>(-6.92)</td>
<td>(-6.71)</td>
<td></td>
</tr>
<tr>
<td>ln(A_{t-1})</td>
<td>0.0187***</td>
<td>0.0176***</td>
<td>0.0186***</td>
<td>0.0122*</td>
<td>0.0116</td>
<td>0.0122*</td>
</tr>
<tr>
<td></td>
<td>(5.93)</td>
<td>(5.62)</td>
<td>(5.92)</td>
<td>(1.96)</td>
<td>(1.88)</td>
<td>(1.97)</td>
</tr>
<tr>
<td>ln(edu_{t-1})</td>
<td>-0.0216***</td>
<td>-0.0215***</td>
<td>-0.0222***</td>
<td>-0.0723***</td>
<td>-0.0707***</td>
<td>-0.0732**</td>
</tr>
<tr>
<td></td>
<td>(-5.09)</td>
<td>(-5.06)</td>
<td>(-5.25)</td>
<td>(-3.18)</td>
<td>(-3.28)</td>
<td></td>
</tr>
<tr>
<td>conflict</td>
<td>-0.0115*</td>
<td>-0.0225**</td>
<td>-0.0083</td>
<td>0.0005</td>
<td>-0.0181</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>(-2.61)</td>
<td>(-3.17)</td>
<td>(-1.73)</td>
<td>(-1.70)</td>
<td>(0.11)</td>
<td></td>
</tr>
<tr>
<td>WpriConpri</td>
<td>-0.0106</td>
<td>-0.0162</td>
<td>-0.0167*</td>
<td>0.0115</td>
<td>-0.0074</td>
<td>-0.0114</td>
</tr>
<tr>
<td></td>
<td>(-1.68)</td>
<td>(-1.82)</td>
<td>(-2.48)</td>
<td>(-1.00)</td>
<td>(-0.45)</td>
<td>(-0.94)</td>
</tr>
<tr>
<td>year</td>
<td>0.0012***</td>
<td>0.0012***</td>
<td>0.0012***</td>
<td>0.0045***</td>
<td>0.0044***</td>
<td>0.0045***</td>
</tr>
<tr>
<td></td>
<td>(4.45)</td>
<td>(4.32)</td>
<td>(4.62)</td>
<td>(5.47)</td>
<td>(5.41)</td>
<td>(5.50)</td>
</tr>
<tr>
<td>cons</td>
<td>-1.927***</td>
<td>-1.872***</td>
<td>-2.011**</td>
<td>-7.834***</td>
<td>-7.752***</td>
<td>-7.892**</td>
</tr>
<tr>
<td></td>
<td>(-3.69)</td>
<td>(-3.58)</td>
<td>(-3.85)</td>
<td>(-5.09)</td>
<td>(-5.02)</td>
<td>(-5.12)</td>
</tr>
<tr>
<td>N</td>
<td>1764</td>
<td>1764</td>
<td>1764</td>
<td>752</td>
<td>752</td>
<td>752</td>
</tr>
</tbody>
</table>

The inclusion of the conflict dummy variable in regression (1-3) is mostly negative and significant. This is consistent with the expected significant, negative impact of conflict on economic growth. However, the results in regressions 4-6 of table 2 indicate that the dummies are no longer significant in a smaller sample. This indicates that for the reduced the sample the results are not as good a fit and are less well specified but maintain the desired sign.

Using the reduced sample, refugee flows are introduced to the investigation by including variables refugee by country of asylum and refugee by country of origin for the years 1990-2010.

For the 36 countries included in the sample, the average number of refugees by country of asylum is 91,620 and the mean of refugee by country of origin is 74,533. This indicates significant refugee flows within the period of 1990 to 2010. The inclusion of both variables allows for a measure of the number of refugees fleeing within the region within the period and the quantity of refugees being absorbed into the region as refugees for the period of 1990-2010. These variables provide a measure of spillover and allow for refugee flows to be disaggregated into two parts that consider both countries of origin and countries of asylum.

Table 3 tests the impact of both refugees by country of asylum and origin, refugees by country of asylum and refugees by country of origin on economic growth for each conflict dummy: conflict, intense and civil. Table 3 contains nine dynamic equations including the conflict dummy variables and their respective weights. Columns (1-3) indicate the results for regressions including the conflict dummy, (4-6) for the intense dummy and (7-9) for the civil conflict dummy. For all the regressions, the conflict dummies are not significant but largely report the desired sign which is consistent.
with the dynamic regression conducted for the reduced sample excluding refugee flows. For all regressions, the economic significance of investment in capital increased in comparison to the reduced sample in Table 2 however, none of the results were statistically significant. The impact of education on economic growth marginally increased in comparison to the reduced sample in Table 2. Finally, the lagged values of investment in capital and education notably increase in significance upon the inclusion of the refugee variables.

Refugee by country of asylum was both statistically and economically insignificant but maintained the expected sign. This can be attributed to the size of the sample, poor documentation of refugee inflows in Sub-Saharan Africa or may indicate that refugee spillover into host countries has almost a negligible effect on economic growth however more research is required as this result is unclear.

Refugee by country of asylum is significant in every specification. Despite being small, the variable takes the expected sign and is significant at the 1% level. This outcome is significant in understanding the economic cost of refugee spillover and can be understood in terms of the existing refugee literature. First, this result reaffirms the hypothesis that flight from a host country induces human capital loss that negatively effects economic growth within that country. This is arguably explained by the capital substitution effect of conflict (Collier, 1999) which explains significant economic loss, reduced rate of growth and reduced levels of growth within industry as a result of human capital loss. This is an interesting result as the conflict literature surrounding refugees focuses largely on refugees as a cost to the recipient country whereas these results show the negative impact that refugee flight has on the host country.

Additionally, this result can be explained through the increased probability of conflict as a result of refugees. Refugees settling in refugee camps in surrounding countries can lead to the militarization of the recipient country and the reinforcement of rebel troops thus increasing the cost of conflict to the country of origin which, as a result, faces greater and stronger opposition during conflict as is similar to the case of the armed Hutus in Zaire in 1996 (Akokpari, 1998). Alternatively this negative impact to economic growth of the country of origin can be explained through remittances sent back to the country of origin by refugees to fund opposition within the conflict (Koser and Van Hear).

### Table 3. Dynamic Regression Including Refugee Variables and Geographical Weights for 1990-2010

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δln(y)</td>
<td>0.0145</td>
<td>0.0110</td>
<td>0.0139</td>
<td>0.0130</td>
<td>0.01</td>
<td>0.0128</td>
<td>0.0144</td>
<td>0.0110</td>
<td>0.0138</td>
</tr>
<tr>
<td>Δln(A)</td>
<td>(1.51)</td>
<td>(1.14)</td>
<td>(1.48)</td>
<td>(1.36)</td>
<td>(1.04)</td>
<td>(1.36)</td>
<td>(1.50)</td>
<td>(1.14)</td>
<td>(1.47)</td>
</tr>
<tr>
<td>Δln(edu)</td>
<td>-0.293</td>
<td>-0.297</td>
<td>-0.277</td>
<td>-0.269</td>
<td>-0.265</td>
<td>-0.259</td>
<td>-0.275</td>
<td>-0.286</td>
<td>-0.264</td>
</tr>
<tr>
<td>ln(n+g+δ)</td>
<td>-2.25</td>
<td>-2.26</td>
<td>-2.19</td>
<td>-2.06</td>
<td>-2.01</td>
<td>-2.04</td>
<td>-2.12</td>
<td>-2.18</td>
<td>-2.09</td>
</tr>
<tr>
<td>ln(y0)</td>
<td>0.0622</td>
<td>0.0791</td>
<td>0.0613</td>
<td>0.0604</td>
<td>0.0761</td>
<td>0.0595</td>
<td>0.0599</td>
<td>0.0777</td>
<td>0.0593</td>
</tr>
<tr>
<td>ln(At-1)</td>
<td>(4.44)</td>
<td>(5.87)</td>
<td>(4.43)</td>
<td>(4.33)</td>
<td>(5.69)</td>
<td>(4.33)</td>
<td>(4.26)</td>
<td>(5.78)</td>
<td>(4.28)</td>
</tr>
<tr>
<td>ln(edut-1)</td>
<td>-0.0905</td>
<td>-0.0879</td>
<td>-0.0852</td>
<td>-0.0891</td>
<td>-0.0858</td>
<td>-0.0843</td>
<td>-0.0882</td>
<td>-0.0881</td>
<td>-0.0839</td>
</tr>
<tr>
<td>conflict</td>
<td>0.0227</td>
<td>0.0166</td>
<td>0.0199</td>
<td>0.0152</td>
<td>0.0170</td>
<td>0.0216</td>
<td>0.0162</td>
<td>0.0183</td>
<td></td>
</tr>
<tr>
<td>WpriConpri</td>
<td>(3.10)</td>
<td>(2.29)</td>
<td>(2.72)</td>
<td>(2.77)</td>
<td>(2.13)</td>
<td>(2.48)</td>
<td>(2.98)</td>
<td>(2.26)</td>
<td>(2.65)</td>
</tr>
<tr>
<td>refugee_a</td>
<td>-0.0905</td>
<td>-0.0879</td>
<td>-0.0852</td>
<td>-0.0891</td>
<td>-0.0858</td>
<td>-0.0843</td>
<td>-0.0882</td>
<td>-0.0881</td>
<td>-0.0839</td>
</tr>
<tr>
<td></td>
<td>(3.51)</td>
<td>(3.38)</td>
<td>(3.41)</td>
<td>(3.45)</td>
<td>(3.29)</td>
<td>(3.36)</td>
<td>(3.39)</td>
<td>(3.35)</td>
<td>(3.31)</td>
</tr>
<tr>
<td></td>
<td>0.0113</td>
<td>0.0039</td>
<td>0.0084</td>
<td>0.0094</td>
<td>0.0140</td>
<td>0.0099</td>
<td>0.0095</td>
<td>0.0033</td>
<td>0.0069</td>
</tr>
<tr>
<td></td>
<td>(1.40)</td>
<td>(0.50)</td>
<td>(1.07)</td>
<td>(0.84)</td>
<td>(1.24)</td>
<td>(0.89)</td>
<td>(1.13)</td>
<td>(0.39)</td>
<td>(0.86)</td>
</tr>
<tr>
<td></td>
<td>-0.0123</td>
<td>-0.0144</td>
<td>-0.0109</td>
<td>-0.0178</td>
<td>-0.0122</td>
<td>-0.0165</td>
<td>-0.0095</td>
<td>-0.0105</td>
<td>-0.0077</td>
</tr>
<tr>
<td></td>
<td>(0.99)</td>
<td>(1.15)</td>
<td>(0.90)</td>
<td>(1.03)</td>
<td>(0.70)</td>
<td>(0.97)</td>
<td>(0.73)</td>
<td>(0.79)</td>
<td>(0.60)</td>
</tr>
<tr>
<td></td>
<td>-1.06</td>
<td>-3.86</td>
<td>-1.57</td>
<td>-7.56</td>
<td>-1.26</td>
<td>-4.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.53)</td>
<td>(0.19)</td>
<td>(0.79)</td>
<td>(0.38)</td>
<td>(0.63)</td>
<td>(0.24)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Finally, the correlations between the two refugee variables, refugees by country of asylum and refugees by country of origin, and the other growth variables are investigated. This tests the relevance of the model used to conduct the analysis and provides a clear illustration of the relationships between refugee flows and other factors of growth.

Refugees by country of asylum is positively and significantly correlated with the conflict variables. While this correlation is small, it is to be expected as it is intuitive that conflict, especially intense conflict which incites fear, induces people to seek asylum elsewhere and become refugees. Refugee by country of origin has higher economic and statistical correlations with the conflict dummies indicating that the presence of conflict in sub-Saharan Africa and the outflow of refugees within the region are closely related.

Unsurprisingly, this relationship is largest for intense conflict, conflict in which there is more than 1000 battle-related deaths a year, which incites fear and thus flight amongst civilians. Refugees by country of origin have a negative, significant correlation coefficient which is intuitive as economic growth is not common in hosts of conflict and thus one would expect that the increase in economic growth is inversely related with an increase in refugees. The same inverse relationship holds true for lagged education and investment in capital which is unsurprising as high investment in capital and levels of education are not common in countries in which there is conflict.

6. Concluding Remarks
Conflict is found to have negative effects on the economic growth of both the host of the conflict and the neighboring countries of the host through the direct cost of conflict and less direct spillover effects. Refugee flows is a common cost of
conflict and spillover effect that can have negative effects on the economic growth of both the host of the conflict and the neighboring countries. By utilizing the theoretical framework established in cost of conflict and spillover literature, it is clear that refugee flows have a negative effect on economic growth for both the country of origin and the recipient country with the former effect being notably significant. This can be explained through by the drain of human capital as a result of the flight of asylum seekers from the host country that ultimately impairs production and growth within the host country and the additional security costs induced by refugees through remittances and rebel networks in refugee camps that fuel conflict in the country of origin. The implications of this is that refugee spillover poses a cost to both the host of the conflict and the recipients of the refugees suggesting that refugee costs should be considered as a multidimensional flow within cost of conflict literature. Additionally, in a broader sense, spillover also ought to be considered in a multidimensional sense as this paper highlights the multidimensional nature of the costs of conflict.

The narrative of refugee flows in Sub-Saharan Africa requires further investigation. One suggestion for further research is weighting the refugee flows according to the countries that they have left and the countries in which they have sought asylum based on distance and contiguity. Geographical distance and contiguity are significant to this as commonly refugees flee by foot or vehicle, methods that limit the distance that can be travelled. Additionally, it would be interesting to weight these variables according to political and economic similarities to test if that has any impact on the direction of refugees. Improved data quality and a larger sample size would further improve the accuracy of this investigation. Finally, it is necessary to establish the temporal effects of refugee flows on the economic growth of countries by origin and by asylum.
References
The Effect of Scandal on the Value of Political Connections in South Korean Chaebol

GRACE KIM
Georgetown University

ABSTRACT
Political connections define the relationship between Korean chaebol firms and the Korean government. What is the effect of corruption scandals on stock prices of firms, both chaebol and non-chaebol? I hypothesize that corruption scandals produce a negative price effect that is exacerbated for firms, such as chaebols, that derive their value from political connections. Using firm-level panel data from 2012 to 2016, I find that the relationship between scandal and stock price is inconclusive but leans negative depending on the specific scandal and news source reporting the scandal. Refuting my hypothesis, chaebol firms experience a positive and significant price effect in reaction to scandal, compared to that of non-chaebol firms. I conclude that chaebol firms are insulated from scandal because scandal breeds uncertainty. In times of uncertainty, investors move capital to safe haven investments like chaebol affiliates in Korea. This finding is contingent on the reporting and type of corruption scandal driving the stock price change.

---

I would like to thank Professor Marko Klašnja for his helpful suggestions and guidance. All remaining errors are my own.
1. Introduction

“I will leave the matters about my fate, including the shortening of my presidential term, to be decided by the National Assembly.” – Park Geun-Hye, November 2016

Most recently displayed by the impeachment of Park Geun-Hye by the Korean National Assembly over a major corruption scandal, corruption is embedded within Korean politics and business. The interconnectivity between politics and business in Korea is championed by the unique prominence of chaebol firms. Chaebol are family-run multinational conglomerates that receive priority treatment from the government as a remnant of the Korean War and state-led liberalization practices in the aftermath of the East Asian financial crisis. They were created for the purpose of promoting national unity and growth. This goal was ultimately met but simultaneously created a fertile breeding ground for corruption. Chaebol status is commoditized in Korean business culture as a prized political connection that provides benefits in the chaebol-government relationship. In this paper, I investigate the value of those political connections by comparing stock price changes between politically connected and non-politically connected firms in reaction to prominent corruption scandals in recent Korean history.

The central question of this study is as follows: what is the effect of Korean corruption scandals, as measured by corruption perception, on stock prices of Korean firms, both chaebol and non-chaebol? I begin with a brief literature review on political connections and chaebol firms. I proceed to outline the theoretical argument based on expectations theory, including episodes of major corruption scandals. Then, I state my hypothesis, which is that (1) scandal correlates negatively with stock prices and (2) chaebol stock prices decrease more than those of non-chaebol firms. I discuss my data and methodology. I highlight key findings in my regression results and sensitivity analysis. Finally, I conclude by discussing the implications of my results and outlining suggestions for further research.

2. Literature Review

2.1 Political Connections

The existing literature suggests that political connections correlate with and can potentially lead to both improved and worsened business prospects. Political connections have been measured by indices, campaign donations, and personal connections to the regime. Future business prospects have been measured by stock price, probability of bailout, loan access, and regulatory conditions.

In Estimating the Value of Political Connections, Fisman (2001) correlates rumor episodes of President Suharto’s adverse health with stock returns of politically dependent firms. The 25 firms studied comprise a substantial amount of economic activity in Indonesia, approximately 30% of total 1995 GNP in revenues. His results indicate that politically connected firms suffer more than those without political connections when the rumors of poor health are more serious. The death of President Suharto would terminate the priority treatment given to politically connected firms, rendering the political connection valueless. With this expectation, market participants re-valued politically connected firms at lower stock prices than those of non-politically connected firms. This reflects the increased sensitivity and potential downside that politically connected firms may bear as compensation for the value derived by political connections. Essentially, Fisman finds that Indonesian firms’ values are significantly dependent on political connections.

Further, in Political Connections and Corporate Bailouts, Facio et al. (2006) study the probability of government bailouts for politically connected firms. They conclude that politically connected firms are more likely to receive an IMF bailout or World Bank assistance. Interestingly, at the time of the bailouts, the politically connected firms studied demonstrated worse financial performance than non-politically connected firms. Political connectedness may hold value for firms undergoing economic distress by influencing the priority allocation of capital. The increased probability of a bailout is a potential upside of political connections.

Political connections also emerge when direct campaign contributions made before election results are “cashed in” by firms if and when the election results are favorable. In Political Connections and Preferential Access to Finance: The Role
of Campaign Contributions, Claessens et al. (2008) measure political connections using stock returns of Brazilian firms that contributed to successful federal candidates. Firms that provided contributions to winners saw higher returns than those who did not. Further, these politically connected firms increased their financing from banks after their candidates won their respective elections. This suggests that political connections work through bank financing in order to produce higher returns.

2.2 Chaebol
Several differences between chaebol and non-chaebol firms have been covered in the existing literature. In The Wealth Effects of Capital Investment Decisions: An Empirical Comparison of Korean Chaebol and Non-Chaebol Firms, Kim, Park, and Zychowicz (2005) argue that the incentive systems between chaebol and non-chaebol firms are different when making capital investment decisions. Investment decisions of chaebol firms did not increase shareholder wealth while the investment decisions of non-chaebol firms did: chaebol management structures promote non-value maximizing capital investment decisions. A potential structural driver for this result is that Korean chaebol firms value the maintenance and acquisition of market share more than returns.

In Financing Constraints and Internal Capital Markets: Evidence from Korean ‘chaebols, Shin et al. (1999) study the differences in investment cash flow sensitivity between chaebol and non-chaebol firms. They find that investment cash flow sensitivity is significantly lower for chaebol firms, relative to non-chaebol firms. Chaebol firms, within their structures, consist of internal capital markets that are legally separate. Compared to firms that do not have this internal system, chaebol firms are less financially constrained. This internal capital system matches the layered structure of chaebol conglomerates in that the conglomerates themselves are comprised of affiliates in multiple industries or sectors. Chaebol firms decidedly are a unique aspect of the Korean economy that produce interesting implications for the value of political connections in regards to corruption.

Myung Sub Choi (2017) of Queensland University of Technology draws inspiration from Fisman’s study in Indonesia to study the effect of interpersonal relations, such as marriage, on stock price movements of Korean firms, chaebol and non-chaebol. Choi measures interpersonal relations using a tier system of connections through blood relations, in-laws, friendship, and directorship with politicians such as the Korean president, ministers, members of parliament, and the heads of local governments. Politically connected firms, those designated as chaebol by the Korea Fair Trade Commission, on average were found to experience a larger stock price increase in response to the expansion of interpersonal relations than that of non-chaebols (Choi, 2017). This price increase correlates positively with increasingly family-oriented interpersonal relationships, maximized at blood relations and marriage. Marital connections and connections with top politicians have the strongest effect out of all connections studied. Further, this result on the affiliate level is consistent with that on the conglomerate level, suggesting a spillover effect between affiliates within the same business group.

2.3 Contribution
Using Fisman’s (2001) model of measuring the value of political connections in Indonesia, I study the effect of scandal on Korean chaebol in years 2012 to 2016 with monthly data. This adds to the existing literature because chaebol studies of recent years have not yet been published, excluding critical events such as the Sewol ferry incident of 2014 and the impeachment of Park Geun-Hye of 2016.

3. Background, Question, and Hypothesis

3.1 Background
Based on Fisman’s (2001) study of Indonesian firms in Estimating the Value of Political Connections, I estimate the value of political connections in Korean chaebol firms. Before proceeding to the central question of my study, I provide a brief background on the corruption landscape in Korea.

The Korean economy is considered to be one of the “largest, fastest growing and most advanced economies among those that are classified as emerging markets” and is dominated by “business groups of conglomerates, called chaebols, which
are owned and managed by a single family” (Kim et al., 2005). The chaebol firms first emerged after the Korean War and gained prominence through Korea’s liberalization process (Tejada, 2017). Liberalization was the Korean government’s reaction to the East Asian financial crisis that led Korea to eventually attain the status of an Asian Tiger. Examples of well-known chaebol firms are Samsung, Hyundai, and LG. Against the backdrop of this chaebol-government relationship, the norm of gift giving is alive in Korea at the intersection of business and politics. As reported by CNN, in Korea, “extravagant meals, gifts [and]… donations at birthdays and funerals are common and accepted as part of culture and business etiquette” (Kwan).

According to Transparency International in 2015, Korea ranks 37 out of 168 on the Corruption Perceptions Index with a score of 55 out of 100 (E.V., 2015). This is shown in Figure 1. As shown in Figure 2, in 2013 over 90% of respondents reported that corruption in the public sector is somewhat of a problem to a very serious problem. Corruption is very much at the forefront of the Korean psyche. Figure 3 highlights specific corruption scandals that occurred during this study’s time frame of January 2012 to December 2016. I expand on those specific scandals below.

Figure 1: Corruption Perceptions Index (TI) Ranking

![CPI in the Asia-Pacific Region, 2015](image)

Figure 2: Over 90% of Koreans Think Corruption is a Problem

![To what extent do you think that corruption is a problem in the public sector in this country? (South Korea, 2013)](image)
Figure 3: Timeline of Major Korean Scandals 2012-2016

In January 2012, the Korean Ministry of Foreign Affairs and Trade (MOFAT) was accused of engaging in bribery over the manipulation of stock prices (Park, 2009). The specific firm involved was a Korean company engaged in diamond mining in Cameroon (Park, 2009). After an investigation which included the raiding of eight offices for evidence and involvement by the Board of Audit and Investigation, Ambassador Kim Eun-Sok was found guilty of manipulating data in favor of CNK International’s bid to develop a diamond mine and in effect causing the company’s stock price to skyrocket. The MOFAT Ambassador was also accused of insider trading by leaking this information to family members prior to the announcement, allowing them to purchase CNK International shares. This insider trading scandal between the public and private sectors in Korea highlighted a “combination of all the corrupt practices in the public sector,” according to DUP Floor Leader Kim Jin-Pyo (Park, 2009).

The National Intelligence Service (NIS) public opinion manipulation scandal occurred during the 2012 presidential election, which Park Geun-Hye won. The scandal focused on the interference of the NIS in swaying public opinion towards Park Geun-Hye, alleging that it was both carried out by an NIS agent and ordered by the NIS Director. This claim was unverified prior to the election but then acted upon after the conclusion of the election, in the first half of 2012. The Seoul NIS Headquarters was raided by authorities who found evidence supporting allegations of election tampering. As a result, both the NIS Director Won Sei-Hoon and Seoul Metropolitan Police Agency Head Kim Yong-Pan were prosecuted for attempting to cover up the public opinion manipulation.

The Sewol ferry incident in April 2014 was particularly traumatic to the Korean psyche for two reasons. First, 200 young students on a school trip died in the sinking of the M.V. Sewol. Children and education are highly valued in Korea, emphasizing the catastrophic nature of this incident (South Korea Overview, 2017). Second, the accident was enabled by bribery that led to lax inspections of the ferry before allowing it to leave the dock (Hwang, 2015). Koreans were outraged that the loss of young, innocent life was made possible by corrupt bureaucratic behavior. The Sewol ferry incident was a showcase of lower level corruption rather than the grand corruption that Korean chaebol firms are linked to. Furthermore, this incident highlighted the fact that corruption can produce negative externalities for non-participants, such as the 200 schoolchildren who perished.

In reaction to this incident, the Kim Young-Ran Anti-Corruption Act went into effect on September 26, 2016. The law targets “public officials, i.e., all officers and employees of public service organizations under the Public Service Ethic Act including state owned enterprises, government institutions and other public institutions” (Ernst, 2016). According to the Korea Times, this law “aims to cut the chain of corruption by putting a cap on whatever goes to regulators and media employees as ‘gifts’ and is expected to force companies to change the way they do business” (Jung et al., 2016). This piece of anti-corruption legislation directly targeted the private-public corruption that caused the Sewol ferry incident.

Several months after the Sewol ferry incident in April 2014, Koreans saw another headline, this one exposing high-level corruption in the “nut rage” incident on a Korean Air flight (Fines, 2016). A Korean Air Vice-President and daughter of the CEO, Heather Cho reportedly went into a rage on a flight when her nuts were served in a bag rather than on a plate. She subsequently harassed the chief steward of the flight and forced the plane back to the gate in order to remove him.
After this incident was exposed, Koreans expressed outrage at the perceived “generation of spoilt and arrogant offspring of owners of the giant family-run conglomerates, or ‘chaebols,’ that dominate the national economy” (Fines, 2016). Forbes reported several months after the incident that although there was a difference between being “thoroughly corrupt… to manipulate records, to lie, cheat, and steal” and “causing mayhem inside a moving plane, […] a lot of Koreans aren’t making that distinction” (Kirk, 2016). Korean Air is owned by the notable chaebol group, the Hanjin Group. The nut rage incident exposed the perils of nepotism endorsed by Korean chaebols, fueling the pre-existing anti-corruption rhetoric.

Prime Minister Lee Wan-Koo, who continued the trend of corruption between public and private sectors, officially resigned in April 2015 after his admission to accepting illegal payments from a businessman (Choe, 2015). He was the second Prime Minister under Park Geun-Hye to resign. The first, Chung Hong-Won, resigned under scrutiny over the Sewol ferry incident in 2014. The businessman connected to this scandal, Sung Wan-Jong, committed suicide in light of corruption charges (Choe, 2015). Prior to his death, he disclosed other illegal payments he had made to members of the Park Geun-Hye administration, not including the Prime Minister (Choe, 2015). This scandal highlighted the pervasiveness and entrenchment of corruption in Korea between the public and private sectors.

The sudden impeachment of Park Geun-Hye in November 2016 struck a final traumatic blow to Korean corruption perception. Exposed through an iPad found in an unsecured building, Park Geun-Hye was accused of sharing classified information with a childhood confidant who did not possess a security clearance or any formal position, Choi Soon-Sil. This “carelessness,” as Park Geun-Hye characterized it, allowed Choi to exert considerable influence over domestic economic policy and relations with North Korea (McCurry, 2017). Choi was accused of exploiting her connection to Park by engaging in abuse of power, fraud, and coercion for the purpose of accumulating personal wealth through her foundations (McKirdy et al., 2016). This corruption scandal also implicated several of Park’s former aides and members of the Korean business community. Shortly after what is now referred to as “Choi-gate,” the acting head of notable chaebol firm Samsung, Lee Jae-Yong, was implicated for paying over $30 million in bribes to Choi Soon-Sil in order to secure support for a Samsung deal (McCurry, 2017). This implication supports the claims that Choi used her personal connection to President Park Geun-Hye to extract bribes out of the business community. Park Geun-Hye was officially impeached on March 9, 2017.

3.2 Question and Hypothesis
In this study, I utilize the effect of scandal to analyze the corruption landscape in Korea. I pose the following question: what is the value of political connections in Korea, as measured by change in stock price, in reaction to change in corruption scandal level, measured by change in corruption perception? I hypothesize that (1) scandal correlates negatively with stock prices and (2) this negative correlation is significantly intensified with chaebol firms because their value is tied to political connections.

The conceptual model behind my hypothesis rests on market expectations. Markets are people-driven. When a scandal occurs, I hypothesize that corruption perception and market uncertainty will increase. Increase in corruption perception will affect the decisions of market participants in the Korean Stock Exchange. Market participants will re-value Korean firms according to expectations of the effect of corruption scandal on future firm value. I hypothesize that this price effect both for chaebol and non-chaebol firms will be negative because (1) scandal increases uncertainty and (2) markets are known to react negatively to uncertainty (Segal et al., 2015). In cases of uncertainty, investors tend to move capital to safe havens. Stocks are considered one of the riskiest securities. Further, I hypothesize that market participants will re-value politically connected firms at a lower price than that of non-politically connected firms. Scandal may diminish the value of political connections in Korea. This difference in price decrease reflects the increased price sensitivity of firms that rely on political connections for value.

4. Data & Methodology

4.1 Independent Variable
The time period for this study is January 2012 to December 2016, during the presidency of Park Geun-Hye. I measure
corruption perception by using a LEXISNEXIS Academic search with the keyword “corruption” in the Korea Times. The Korea Times is unaffiliated with any chaebol firm and the sister newspaper of Hankook Ilbo, a major Korean newspaper. The articles are in English and translated directly from the original Korean version. I divide the number of articles with the keyword by the total number of articles per month to yield the percentage of articles referencing corruption. This measures corruption perception. I argue that corruption perception measured through this method is a valid and relevant proxy for general scandal. Market participants receive some information from news sources like the Korea Times and analyze that information to make trading decisions regarding specific firms.

\[
\text{Articles with keyword} \div \text{Total articles} = \% \text{Corruption Perception}
\]

Figure 4: Incidence of “Corruption” in the Korea Times

4.2 Dependent Variable
The stock price data is monthly from years 2012 to 2016. The indicator is calculated as an average of daily opening quotes for each month. The total number of observations is 948 for 16 firms listed below (7 chaebol and 9 non-chaebol). The firm data collected is on the affiliate level for chaebol firms because the multinational conglomerates as a whole do not trade.

Table 1: List of Firms

<table>
<thead>
<tr>
<th>Chaebol</th>
<th>Non-chaebol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daewoo Shipbuilding &amp; Marine Engineering</td>
<td>Dongbang Ship Machinery Co.</td>
</tr>
<tr>
<td>Dongkuk Refractories &amp; Steel Co.</td>
<td>Dastek Co.</td>
</tr>
<tr>
<td>Hyundai Mipo Dockyard</td>
<td>Kwang Jin Industry Co.</td>
</tr>
<tr>
<td>Doosan Engine Co.</td>
<td>Oriental Precision &amp; Engineering Co.</td>
</tr>
<tr>
<td>Samsung Heavy Industries</td>
<td>Daedong Co.</td>
</tr>
</tbody>
</table>
I measure stock prices of firms by using market data from Mergent Online. I limit my firm sample pool to manufacturing firms using Mergent Online’s primary NAICS classification in order to control for differences across industries. Within this classification, I use category 33 of manufacturing, which includes:

- Primary Metal Manufacturing
- Fabricated Metal Product Manufacturing
- Machinery Manufacturing
- Computer and Electronic Product Manufacturing
- Electrical Equipment, Appliance, and Component Manufacturing
- Transportation Equipment Manufacturing
- Furniture and Related Product Manufacturing
- Miscellaneous Manufacturing

Figure 5: Price Volatility of Firms (2012-2016)

In Figure 5, I calculate and plot the average change in price between months for both chaebol and non-chaebol firms. The graph indicates that, on balance, chaebol firms exhibit greater price volatility than that of non-chaebol firms.

4.3 Descriptive Statistics & Model
In my data analysis, I test the correlation between changes in corruption perception, as shown in Figure 4, and changes in stock price, as shown in Figure 5. Below are the descriptive statistics for the variables I use to conduct the data analysis and sensitivity analysis.
Table 2: Descriptive Statistics

<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>N</td>
<td>mean</td>
<td>sd</td>
<td>min</td>
<td>max</td>
</tr>
<tr>
<td>ktpercent (%)</td>
<td>948</td>
<td>2.981</td>
<td>2.364</td>
<td>1.170</td>
<td>15.74</td>
</tr>
<tr>
<td>khpercent (%)</td>
<td>948</td>
<td>2.578</td>
<td>1.945</td>
<td>0.390</td>
<td>12.82</td>
</tr>
<tr>
<td>kospiavg</td>
<td>948</td>
<td>1.976</td>
<td>61.22</td>
<td>1.827</td>
<td>2,115</td>
</tr>
<tr>
<td>unemployment</td>
<td>948</td>
<td>3.441</td>
<td>0.257</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>interest</td>
<td>918</td>
<td>0.999</td>
<td>0.307</td>
<td>0.500</td>
<td>1.500</td>
</tr>
<tr>
<td>usdkrw</td>
<td>948</td>
<td>1,113</td>
<td>49.71</td>
<td>1,012</td>
<td>1,238</td>
</tr>
<tr>
<td>gdpcont</td>
<td>948</td>
<td>3.333e+08</td>
<td>8.124e+07</td>
<td>3.600e+07</td>
<td>3.900e+08</td>
</tr>
<tr>
<td>price</td>
<td>941</td>
<td>3.440</td>
<td>26.852</td>
<td>35.51</td>
<td>179,095</td>
</tr>
<tr>
<td>chaebol</td>
<td>948</td>
<td>0.501</td>
<td>0.500</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>firmsize</td>
<td>772</td>
<td>1.077e+09</td>
<td>1.743e+09</td>
<td>-8.600e+09</td>
<td>5.800e+09</td>
</tr>
<tr>
<td>leverage</td>
<td>715</td>
<td>1.582</td>
<td>2.536</td>
<td>0</td>
<td>13.30</td>
</tr>
<tr>
<td>sewolcont</td>
<td>948</td>
<td>0.552</td>
<td>0.498</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>sewol</td>
<td>948</td>
<td>0.0169</td>
<td>0.129</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>impeachcont</td>
<td>948</td>
<td>0.0316</td>
<td>0.175</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>impeach</td>
<td>948</td>
<td>0.0158</td>
<td>0.125</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>pmcont</td>
<td>948</td>
<td>0.349</td>
<td>0.477</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>pm</td>
<td>948</td>
<td>0.0169</td>
<td>0.129</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>nis</td>
<td>948</td>
<td>0.0169</td>
<td>0.129</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>niscont</td>
<td>948</td>
<td>0.932</td>
<td>0.251</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>korair</td>
<td>948</td>
<td>0.0169</td>
<td>0.129</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>koraircont</td>
<td>948</td>
<td>0.417</td>
<td>0.493</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>placebo</td>
<td>948</td>
<td>0.0169</td>
<td>0.129</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>placebocont</td>
<td>948</td>
<td>0.164</td>
<td>0.370</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

I use Ordinary Least Squares (OLS) regression using panel data to test the effect of change in corruption perception on change in stock price for both chaebol and non-chaebol firms. Below is the expanded model that I use:

**Figure 6: Regression Model**

\[
\ln(\text{STOCK PRICE}) = B_0 + B_1 \text{KTPERCENT} + B_2 \text{CHAEBOL} + B_3 \text{KOSPIAVG} + B_4 \text{KTPERCENT} \times \text{CHAEBOL} + B_5 \text{KTPERCENT} \times \text{KOSPIAVG} + B_6 \text{FIRMCHARACTERISTICS} + B_7 \text{CONTROLS} + B_8 \text{DUMMY} + B_9 \text{KTPERCENT} \times \text{DUMMY} + B_{10} \text{KTPERCENT} \times \text{CHAEBOL} \times \text{DUMMY} + e
\]

I control for chaebol status and Korea Composite Stock Price Index (KOSPI) to minimize missing variable bias. Chaebol is coded ‘1’ for chaebol and ‘0’ for non-chaebol. KOSPI is a composite measure of the Korean stock market as a whole, included to provide context for the stock price movements of individual firms. KOSPI is calculated as a monthly average of daily KOSPI opening quotes.

I include two interactions with my measure of corruption perception. The first, interacting corruption perception and chaebol, isolates the additional effect of chaebol status on the base correlation between corruption perception and stock
price. The second, interacting corruption perception and KOSPI, isolates the additional effect of a bullish stock market on the base correlation between corruption perception and price. I include these interaction terms to observe whether chaebol status or the stock market as a whole has a significant influence on the effect of scandal on stock price. I also include two firm characteristics, firm size and leverage, in order to control for differences across firms. Firm size is calculated as total equity. Leverage is calculated as total debt over total equity. These two indicators are the most widely employed firm characteristic controls in the existing literature on chaebol firms. In my most expanded models, I include control variables for the Korean economy including exchange rate (KRW/USD), interest rate (10 year government bond), unemployment rate, and GDP. These variables provide a critical economic context and minimize missing variable bias.

I include several dummy variables for the specific scandals previously mentioned. I employ dummy variables using two methods. In the first, the scandal dummy variables are coded ‘1’ for the month and year they occurred and for every month after. Months prior to the occurrence of the scandal are coded ‘0’. This is the sustained version, measuring the long-term effect. In the second version, the scandal dummy variables are coded ‘1’ for only the month and year in which they occurred. Every other month and year is coded ‘0’. This is the un-sustained version, measuring the instantaneous effect. Both sustained and un-sustained dummy variables are included to test whether the correlation between corruption perception and stock price was significantly impacted by specific corruption scandals. The specific scandals that I convert in dummy variables are the NIS public opinion manipulation scandal, the Sewol ferry incident, the Korean Air nut rage incident, the resignation of Prime Minister Lee Wan-Koo, and the impeachment of President Park Geun-Hye. I exclude the MOFAT diamond scandal because it occurred at the very beginning of my study in January 2012, voiding any potential effect of using a sustained dummy variable.

Finally, I include triple interactions, along with constituent terms, in separate expanded models for each event dummy variable, chaebol status, and corruption perception. This interaction captures the additional price effect, if any, of chaebol status in reaction to a specific event while holding corruption perception constant. This will allow me to interpret whether different types of corruption scandals have different effects on the base correlation between corruption perception and stock price while also isolating the chaebol effect. The presence of this triple interaction serves as a secondary test of my second hypothesis.

5. Results
I present the results of the bare bones model, expanded model, and sensitivity analysis. All tables to which I refer may be found in the appendix at the end of this paper. In my tables, I highlight my main variables of interest.

5.1 Bare Bones Model
As shown in Table 3, the bare bones model—including stock price, corruption perception, chaebol status, and KOSPI—does not yield significant results. The second model with firm characteristics, third and fourth models with separate interactions, and fifth model with firm characteristics and both interactions all do not yield significant results. In all five models, the effect of leverage as a control variable is negative and significant, indicating that more leverage is correlated with a lower stock price. On balance, these results do not support either of my hypotheses.

5.2 Expanded Model
I expand the model in Table 4 to include economic control variables and the sustained dummy variables for specific scandals. I highlight several notable aspects of the regression results.

In all models, the coefficient on the interaction term ktchaebol (corruption perception and chaebol status) is positive and significant at the 99% confidence level. This indicates that chaebol firms experience a positive price effect in the relationship between corruption and stock price. This directly refutes my second hypothesis. The sixth model, controlling for the impeachment of November 2016, is dropped from the results table due to collinearity: the impeachment dummy variable is highly correlated with the corruption perception indicators. Both indicators spiked in November 2016 in reaction to the exposure of Park Geun-Hye’s corruption scandal.
In the third model, controlling for the Sewol ferry incident, the coefficient on the sustained dummy variable is -0.264 and significant at the 99% confidence level. The long-term impact of the Sewol ferry incident on stock price is negative. There is no special Sewol effect for chaebol firms, as shown by the insignificant triple interaction (sewolcontktchaebol).

In the fourth model, controlling for the Korean Air nut rage scandal, the coefficient on the sustained dummy variable is -0.329 and significant at the 99% confidence level. The long-term impact of the Korean Air nut rage scandal on stock price is negative. There is no special nut rage effect for chaebol firms, as shown by the insignificant triple interaction (korcontktchaebol).

In the fourth and fifth models, controlling for the Korean Air nut rage scandal and the Prime Minister’s resignation, the coefficient on corruption perception is negative and significant at the 95% confidence level. This indicates that for these models, more scandal is correlated with lower stock prices for non-chaebol firms. This supports my first hypothesis. As shown by the positive and significant interaction term on ktchaebol (corruption perception and chaebol status), chaebol firms experience a negative price effect as well but not as negative as that of non-chaebols.

- Model 4: a 1% increase in corruption perception correlates with a -0.838 percentage point decrease in the natural log of stock price for non-chaebols and a -0.722 percentage point decrease in the natural log of stock price for chaebols.
- Model 5: a 1% increase in corruption perception correlates with a -1.684 decrease in the natural log of stock price for non-chaebols and a -1.56 decrease in the natural log of stock price for chaebols.
- The interaction effects are graphed in Figures 7 and 8 in the appendix.

The coefficient on the interaction ktkospi (corruption perception and KOSPI) is positive and statistically significant but ultimately negligible because the value is close to zero.

<table>
<thead>
<tr>
<th>Table 4</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VARIABLES</td>
<td>Korean Air</td>
<td>Prime Minister</td>
</tr>
<tr>
<td>ktpercent</td>
<td>-0.838**</td>
<td>-1.684**</td>
</tr>
<tr>
<td></td>
<td>(0.333)</td>
<td>(0.669)</td>
</tr>
<tr>
<td>chaebol</td>
<td>0.637***</td>
<td>0.606***</td>
</tr>
<tr>
<td></td>
<td>(0.0816)</td>
<td>(0.0800)</td>
</tr>
<tr>
<td>kospiavg</td>
<td>-0.000789</td>
<td>-0.00203*</td>
</tr>
<tr>
<td></td>
<td>(0.000514)</td>
<td>(0.00108)</td>
</tr>
<tr>
<td>ktchaebol</td>
<td>0.116***</td>
<td>0.124***</td>
</tr>
<tr>
<td></td>
<td>(0.0283)</td>
<td>(0.0269)</td>
</tr>
<tr>
<td>ktkospi</td>
<td>0.000404**</td>
<td>0.000834**</td>
</tr>
<tr>
<td></td>
<td>(0.000166)</td>
<td>(0.000344)</td>
</tr>
</tbody>
</table>

I expand the model in Table 5 to include economic control variables and the un-sustained dummy variables for specific scandals. I compare the results with those of the sustained dummy variable models.

Again, the sixth model, controlling for the impeachment of November 2016, is dropped from the results table due to collinearity. The dummy variable is highly correlated with the corruption perception indicator. In all five models, the coefficients on the interactions between corruption perception and chaebol (ktchaebol) remain positive and statistically significant at the 95% confidence level, but decrease in coefficient value. The Sewol dummy variable’s statistical significance and Korean Air dummy variable’s statistical significance disappear.

The statistical significance of the main corruption perception coefficient in the fourth model (Korean Air nut rage scandal)
disappears but remains in the fifth model (Prime Minister resignation). In Model 5, the coefficient decreases in absolute value from −1.684 to −1.604. A 1% increase in corruption perception correlates with a −1.604 decrease in natural log of price for non-chaebols and a −1.494 decrease in natural log of price for chaebols.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VARIABLES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ktpercent</td>
<td>-0.554</td>
<td>-1.604**</td>
</tr>
<tr>
<td></td>
<td>(0.441)</td>
<td>(0.724)</td>
</tr>
<tr>
<td>chaebol</td>
<td>0.650***</td>
<td>0.607***</td>
</tr>
<tr>
<td></td>
<td>(0.0989)</td>
<td>(0.119)</td>
</tr>
<tr>
<td>kospiavg</td>
<td>-0.000716</td>
<td>-0.00206*</td>
</tr>
<tr>
<td></td>
<td>(0.000727)</td>
<td>(0.00107)</td>
</tr>
<tr>
<td>ktchaebol</td>
<td>0.0900**</td>
<td>0.110**</td>
</tr>
<tr>
<td></td>
<td>(0.0399)</td>
<td>(0.0492)</td>
</tr>
<tr>
<td>ktkospi</td>
<td>0.000255</td>
<td>0.000791**</td>
</tr>
<tr>
<td></td>
<td>(0.000222)</td>
<td>(0.000370)</td>
</tr>
</tbody>
</table>

I summarize the reported regression results below including the sign of the coefficient and significance level as denoted by stars:

**Table 9: Korea Times Regression Results**

<table>
<thead>
<tr>
<th></th>
<th>Sustained</th>
<th>Un-sustained</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bare Bones</td>
<td>Controls</td>
</tr>
<tr>
<td>corruption perception</td>
<td></td>
<td></td>
</tr>
<tr>
<td>corruption perception x chaebol</td>
<td>**</td>
<td>***</td>
</tr>
<tr>
<td>event dummy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>triple interaction</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1

5.3 Sensitivity Analysis

First, I account for any potential bias associated with the Korea Times by re-running my regressions with a different news
Both newspapers exhibit similar trends in corruption perception, but the Korea Herald exhibits sharper decreases, higher volatility, and greater variation. I replace the Korea Times variable with the Korea Herald variable in my bare bones table and expanded model tables. This change yields the following results:

In Table 6 (bare bones), the fourth model yields a negative and significant corruption perception coefficient. The fourth model includes firm characteristics and the interaction coefficient ktchaebol (corruption perception and chaebol status). This indicates that a 1% increase in corruption perception, measured by the Korea Herald, correlates with a −0.0271 decrease in natural log of price for non-chaebols. The interaction coefficient itself is not statistically significant. In this bare bones table, the Korea Herald measure is more supportive of my first hypothesis than the Korea Times measure.

For Tables 7 and 8 (expanded analysis), all models yield a positive and significant interaction term for ktchaebol (corruption perception and chaebol status). This reinforces the positive chaebol effect first noted in Tables 4 and 5 between corruption perception and stock price. Again, this directly refutes my second hypothesis. In Table 7, which includes economic control variables and the sustained scandal dummy variables, the coefficients on the second, third, and fifth models (NIS manipulation, Sewol ferry incident, and Prime Minister resignation, respectively) are negative and statistically significant. This is a net increase in statistical significance from the previous regression analyses using the Korea Times, in which coefficients on the second and third models were not statistically significant.

In Table 7’s third model, the Sewol ferry incident sustained dummy variable remains negative and statistically significant. This indicates that the long-term impact of the scandal on stock prices of non-chaebol firms is negative. The triple interaction between the Sewol ferry incident sustained dummy variable, corruption perception, and chaebol status is negative and statistically significant at the 99% confidence level. This indicates that, holding corruption perception constant, the long-
term impact of the Sewol ferry incident was more negative for chaebol firms by -0.0473.

In Table 7’s fourth model, the Korean Air nut rage sustained dummy variable remains negative and statistically significant. This indicates that the long-term impact of the Korean Air nut rage scandal on stock prices is negative. In Table 7’s fifth model, the triple interaction between the Prime Minister’s resignation sustained dummy variable, corruption perception, and chaebol status is negative and statistically significant at the 95% confidence level. This indicates that, holding corruption perception constant, the impact of the Prime Minister resignation scandal is more negative for chaebol firms. The sustained dummy variable itself is not statistically significant.

In Table 8, which includes economic control variables and the un-sustained dummy variables, the statistical significance of the main corruption perception coefficient disappears in the third model (Sewol ferry incident) but remains in the second and fifth models (NIS manipulation and Prime Minister resignation, respectively). In Table 8’s third model, the statistical significance of the Sewol ferry incident un-sustained dummy variable disappears but the triple interaction coefficient becomes positive and statistically significant at the 95% confidence interval. This indicates that the instantaneous impact of the scandal, holding corruption perception constant, is more positive for chaebol firms. All other triple interactions are not statistically significant.

Table 10: Korea Herald Regression Results

<table>
<thead>
<tr>
<th></th>
<th>Sustained</th>
<th>Un-sustained</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bare bones Controls NIS Sewol Korean Air PM</td>
<td>Bare bones Controls NIS Sewol Korean Air PM</td>
</tr>
<tr>
<td>corruption perception</td>
<td>-. * - **</td>
<td>- * - **</td>
</tr>
<tr>
<td>corruption perception x chaebol</td>
<td>+ ** + *** + *** + *** + ***</td>
<td>+ ** + *** + *** + *** + ***</td>
</tr>
<tr>
<td>event dummy</td>
<td>- ** - ***</td>
<td>- ** - ***</td>
</tr>
<tr>
<td>triple interaction</td>
<td>- *** - ***</td>
<td>- *** - ***</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1

Second, I conduct a placebo test by creating a dummy for a placebo scandal and re-running my expanded regression. The goal of the placebo test is to further confirm that the previously observed results are in fact driven by specific scandals. If the results are scandal-driven, the coefficients of the event dummy and triple interaction for the placebo scandal should not be statistically significant. I choose March 2016 as my placebo month and year. The results of the placebo test are below. The columns labeled “KT Controls” and “KH Controls” are the previous controls models, included for comparison purposes. I run the expanded model with both the Korea Times and the Korea Herald data. The full regression output is found in Tables 12 and 13 in the appendix.
Table 11: Placebo Test Regression Results

<table>
<thead>
<tr>
<th></th>
<th>KT Control</th>
<th>KT Un-sustained</th>
<th>KT Sustained</th>
<th>KH Control</th>
<th>KH Un-sustained</th>
<th>KH Sustained</th>
</tr>
</thead>
<tbody>
<tr>
<td>corruption perception</td>
<td></td>
<td>-*</td>
<td></td>
<td></td>
<td></td>
<td>-**</td>
</tr>
<tr>
<td>corruption perception * chaebol</td>
<td>++**</td>
<td>++**</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>event dummy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+*</td>
</tr>
<tr>
<td>triple interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1

Overall, the placebo test confirms that my previous results are driven by specific scandals. Using both the Korea Times and the Korea Herald, none of the triple interactions, which isolate the effect of the scandal on chaebols while holding corruption perception constant, are statistically significant. The only statistically significant event dummy variable is the sustained placebo using the Korea Herald at the 90% confidence interval. The sustained dummy variable is coded ‘1’ for March 2016 to December 2016, absorbing the effect of the Park Geun-Hye impeachment. The impeachment, rather than the placebo, is likely driving this result. This explanation is supported by the lack of statistical significance for the unsustained placebo dummy. Similarly, the main corruption perception coefficients are only statistically significant with the sustained placebo dummy variable. The statistical significance of the interaction between corruption perception and chaebol status is consistent with that of the controls model without any event dummy variables.

The inclusion of the November 2016 Park Geun-Hye impeachment scandal is likely driving the statistical significance of the main corruption perception variable. A potential solution to this problem is coding November and December 2016 as ‘0’ instead of ‘1’ to exclude the Park Geun-Hye impeachment. I run this model but exclude it from any further analysis because it yields similar results to Table 11. While this sows some doubt regarding the robustness of results, it does not provide a significant cause for concern. Neither the placebo dummy variable nor the triple interaction yield overwhelmingly statistically significant results.

6. Implications & Conclusion

On balance, the results of the bare bones model, expanded model, and sensitivity analysis models do not fully support my first hypothesis and refute my second hypothesis. The results indicate that (1) the correlation between corruption perception and stock price is small or leans negative depending on the model and (2) the effect of scandal on stock prices of chaebol firms is statistically significantly more positive than that of non-chaebol firms. This chaebol effect is stronger in the models with sustained dummy variables compared to those with unsustained dummy variables, indicating that the effect of scandal and chaebol status on stock prices is not a one-off event but one that carries long-term impacts. The sensitivity analysis using the Korea Herald corruption perception indicator yield stronger results than the original results using the Korea Times corruption perception indicator. This indicates a source of bias from the reporting of the scandals. I present several explanations for these results:

6.1 Safe Haven Effect

Returning to my conceptual model based on market expectations, market participants may price up chaebol firms in reaction to scandals. When scandals increase uncertainty, capital gravitates towards safer investments. In Korea, the safest equity investment is chaebol affiliates. Market participants base future expectations on past truths. Therefore, every corruption scandal that does produce real, systemic change reinforces the systemic power of chaebol firms. This results in a chaebol price insulation effect. Scandals between the public and private sector counterintuitively incentivize capital flows to the very type of firm that benefits most from the chaebol-government relationship. This effect demonstrates the cyclical nature of entrenched corruption in some societies.
6.2 Lost in Translation
The comparison between results using the Korea Times and results using the Korea Herald indicates that some of the corruption perception that I attempt to capture may be lost in translation. Because the Korea Times is the sister paper and direct translation of Hankook Ilbo, the keyword search for the English word “corruption” may not fully capture the actual incidence of “corruption” in Korean. Expanding the keyword search to include other corruption-related words such as “bribe” or doing the search in Korean using Hankook Ilbo could ameliorate this problem. The Korea Herald articles are originally written in English and therefore bypass the translation barrier encountered with the Korea Times. This potential translation error is reflected in the previous Figure 8, which demonstrates that the incidence of keyword “corruption” is more volatile for the Korea Herald compared to the Korea Times. A function of measurement error, this increase in variation may be driving the sensitivity analyses’ stronger results.

Beyond simple translation, the two newspapers may differ in the quantity and composition of their readerships. Given that both newspapers publish a comparable total number of articles online each month, the Korea Times may cater to a more Korea-focused audience than the international audience of the Korea Herald. The Korea Herald’s English publications imply that its domestic audience understands English as a second language. In Korea, students, especially those who are comparatively well endowed, are taught English as a part of their school curriculum (Park, 2009). The ability to speak English well serves as an elite social status signal (Park, 2009). Because of this, the Korea Herald may choose to focus on more internationally relevant issues for an internationally-minded audience.

6.3 Types of Scandal
Certain scandal characteristics may correlate more statistically significantly and negatively with stock price. A potential critique of this study is that it does not differentiate between different types of corruption such as petty corruption or grand corruption. I acknowledge that differences in event models may be driven by different types of corruption. At the same time, I argue that regardless of whether petty corruption or grand corruption drives the scandal, the common denominator is that the corruption transaction is between public and private entities. Of the scandals that were isolated through dummy variables, the scandal that had the strongest results across all models was the Prime Minister resignation of 2015. The Prime Minister resignation, a result of illegal monetary transfer between business and the highest level of government, captures that common denominator.

Furthermore, I argue that the price effect of scandals is driven by where the scandal blame lies. Scandal blame can lie with non-chaebol firms, chaebol firms, the government, or a combination. The Sewol ferry incident shifted the scandal blame to the bureaucratic level of government. The Korean Air nut rage incident shifted the scandal blame to the nepotism of the managing chaebol firm, the Hanjin Group. The NIS public opinion manipulation scandal shifted the scandal blame to mid-ranking individuals in the government. The Prime Minister resignation shifted the scandal blame to the highest level of government. Based on the results, scandals that both shift blame to higher levels of government and include clear connections to business yield greater price impacts. The market seems to be less impacted by scandals that focus on specific firms because the scandal scope is limited and easier for the market to dilute. Blaming the government, especially on higher institutional levels, produces a broader market effect that this study captures.

I was unable to reasonably isolate the Park Geun-Hye impeachment effect within my expanded models because it occurred in the second to last month of this study’s time frame. However, should this study be replicated with more recent 2017 data, I predict that the impeachment model will yield statistically significant results similar to those of the Prime Minister resignation. Both scandals embody clear ties between the public and private sector that produce negative externalities for non-participants in the corruption transaction.

6.4 Conclusion: Bad for Anti-Corruption
This study highlights one of the main barriers to anti-corruption efforts: the reaction to corruption scandals counter-intuitively reinforces the power of the firms who benefit the most from corruption. Corruption is persistent. The reason for this may be that the public sector is easier to punish than the private sector. While Koreans can express their distaste with corruption in the polls, Koreans have less flexibility in punishing the private sector because the economy is dominated by
chaebol firms. Chaebol firms wield immense influence in Korean market share as well as political outcomes (Tejada, 2017). After Korea’s transition to democracy in the 1980s, chaebol firms even became active in political campaigns and elections (Tejada, 2017). Demonstrating how dependent the Korean economy is on chaebol firms, heads of chaebol firms are given special treatment in order to preserve the stability of the Korean economy (Tejada, 2017). In corruption scandals centered on the chaebol-government relationship, public officials are duly punished by the people while the systemic power of chaebols as a group emerges relatively unscathed. However, these implications rely on data prior to the November 2016 Park Geun-Hye impeachment, which has the potential to be a turning point in the chaebol-government relationship. In the transition away from Korea’s corrupt history, Koreans must realize that punishing the public sector for the corruption between government and business is not enough. Chaebol firms must also bear the burden of corruption scandal backlash.

Outside of Korea, this study demonstrates a larger problem in regards to the establishment. Noting the rise of populism in both Europe and the United States, the Korean example demonstrates the status quo before this global turning point in which the establishment, represented by chaebol firms, is untouchable (Roth, 2017). The findings of this study may soon become outdated in a changing political and social atmosphere. The impeachment of Park Geun-Hye, the victory of Donald Trump, and the Brexit decision all broadly indicate a global trend away from the establishment insulation captured in this study. International investors, especially those looking at Korea, should be wary of basing future expectations on past correlations and trends. Those correlations and trends may soon weaken or even reverse.

7. Limitations & Further Research
I acknowledge the limitations of this study by virtue of its limits on time frame and industry. I also acknowledge potential bias introduced by the absence of time-fixed effects, which I could not apply because data for my economic control variables are not produced monthly.

Further research on this topic should investigate using alternate measures for corruption perception. Although I used newspapers, a modern approach could include social media measures. An expansion of this study should study the prevalence of certain hashtags on social media sites in reaction to scandals. Tracking social media traffic could produce a more organic, grassroots measure than one which relies on newspapers owned by corporations. Further research should also investigate the impact of calculating monthly stock price as a median of a given month’s stock prices or the opening quote of the first of each month, rather than as a monthly average. I acknowledge that averaging the stock price to produce a monthly indicator may dilute the actual price effect of scandal.

Regarding the placebo tests, further research should run a placebo model for every month not marked as a scandal month. Although sustained models closer to November 2016 may yield significance because of the Park Geun-Hye impeachment inclusion, further research should check that this trend weakens as dummy scandals are coded further away from November 2016. After running all the placebo tests, a percentage success rate should be calculated to determine even more robustly whether the scandal dummy variables yield meaningful results.

Regarding control variables, future research should include additional firm-specific data including chaebol revenue from government contracts, firm beta, firm dividends, coverage, and previous firm crisis management factors. I was unable to include these potentially meaningful factors in my regression analysis because of data constraints. Any observation limitations should be supplemented by increasing the number of firms in the study or expanding the study to other industries in addition to manufacturing.

Finally, further research should run the omitted sixth model controlling for the impeachment of Park Geun-Hye in November 2016 with 2017 data. Although I tried to capture the impeachment effect, the time constraint of real time and data limitations made it impossible.
Works Cited


“South Korea Overview.” NCEE, Center on International Education Benchmarking, ncee.org/what-we-do/center-on-international-education-benchmarking/top-performing-countries/south-korea-overview/.
Attribute-Based Regulations:  
The Case of Corporate Average Fuel Economy (CAFE) Standards

GRADY KILLEEN1  
Georgetown University

ABSTRACT

In 2011, the Department of Transportation modified Corporate Average Fuel Efficiency (CAFE) standards so that larger cars needed to meet a less stringent fuel economy target. Cars with a greater “footprint” – the area in square feet between the wheels of a vehicle – are now permitted to use more fuel per mile without paying fines. I use the EPA’s Fuel Economy Trends database, which was obtained through a FOIA request, to provide evidence that vehicle manufacturers expanded vehicle footprint to comply with the new CAFE standards. Since footprint and internal vehicle volume both measure vehicle size, the two vehicle specifications are expected to follow similar trends. However, I find a discontinuity between footprint and volume once footprint-based CAFE standards were adopted. Since 2011, footprint has grown by about one square foot relative to interior volume. Trends in vehicle performance and weight do not explain this growth. This growth in car footprints since 2011 offset around 17% of claimed fuel economy gains between 2012 and 2015.

1 I would like to thank Professor Arik Levinson for all of his help writing this paper and obtaining the data for it. I would also like to thank Professor Robert Cumby for useful advice.
Transparency and the Effect of Natural Resources on Economic Growth

OLIVIA BISEL
YUOU WU
Georgetown University

ABSTRACT

Natural resources represent a curse for a diverse array of economies. Yet in other countries, natural resource “rents” provide a lucrative source of economic growth. What explains the gross disparity in outcomes? We believe the answer lies in transparency. We hypothesize that transparency disproportionately benefits the GDP growth of resource-abundant countries, so that economically transparent countries may benefit from natural resources while economically opaque countries are hurt. Many scholars claim that transparency bolsters GDP growth by encouraging foreign investments. Given the higher risk of corruption for resource-abundant countries, transparency should have an especially profound impact for resource-abundant countries by enhancing resource revenues management. Analyzing panel data for 228 countries and territories from 1990 to 2010, we show that the effect of natural resource rents on economic growth depends on a country’s level of economic transparency. Our findings, which are robust to alternative specifications, suggest that countries with abundant resources should prioritize transparency to ensure resource dependence does not hinder GDP growth.
Oil wealth has been shown to shape the political economy of countries, but little research has analyzed how oil price fluctuations mediate its effects. I argue oil price fluctuations influence the clout of countries, whereby price increases magnify the clout of petrostates but diminish the influence of major oil importers and countries opposed to petrostates. Using United Nations General Assembly roll-call voting to measure changes in clout from 1960-2014, I find petrostates Russia and Saudi Arabia win a greater share of votes, in particular votes from oil importers, against the United States when oil prices are higher. My findings are robust to country and year fixed effects, alternative measures of vote agreement, and several robustness checks. To explain this finding, I examine how oil prices affect the domestic economics of states and alter the tradeoffs of seeking economic interdependence abroad.
The collapse of oil prices in early 2016 rattled the global economy. From upending financial markets (Sheppard and Hume, 2016) to threatening fiscal deficits for oil producers (Facuon et al., 2014) to providing gasoline savings to consumers (Wells, 2016), the nose-dive demonstrated the continued influence of petroleum, both in its great depth and breadth.

This influence has garnered much analysis from scholars, analyzing everything from how petroleum affects regime type to interstate conflict. Little research, however, has examined the role of oil prices. Thomas Friedman's (2009) "First Law of Petropolitics" notably theorizes high oil prices should magnify the impacts of oil wealth, while low prices should diminish those effects. Still, this theory has evaded close examination despite its potentially large impact. With this void in mind, I examine how oil price fluctuations affect the clout of countries, namely petrostates and major oil importers, as well as countries which tend to oppose petrostates, largely the West. In effect, I aim to develop a foundational relationship on how oil prices influence international relations and power.

Arguments that rising oil prices strengthen the clout of petrostates abound in the media but overwhelmingly go unexplained, not to mention untested. That said, there is good reason to believe this relationship holds. Building on Ross and Voeten (2016), I maintain two mechanisms connect price fluctuations to clout: domestic economics and economic interdependence. Domestically, petrostates are highly dependent on oil revenues. With high oil prices, these countries receive windfalls and can spread their influence abroad. But when prices fall, they must cut spending. Here, international aims lose importance and more emphasis is placed on appeasing the domestic population. For oil importers, high prices slow economic growth.

As for the second mechanism, petrostates typically do not seek economic interdependence, such as creating trade agreements or joining international organizations, because oil is always in demand and because they can easily attract foreign investment without doing so. When oil prices drop, however, oil producers must stabilize their economies. Consequently, they become more reliant on other countries, the views of other countries, and international norms. For oil importers, as oil prices fall they gain leverage over petrostates looking for economic interdependence and can more effectively impart their norms and policies, such as sanctions, against petrostates.

I argue oil prices influence the clout of major countries through these mechanisms. Here, "clout" signifies the ability to influence other countries, in turn affecting both hard and soft power (Nye, 2009). To operationalize changes in clout, I analyze vote agreement in the United Nations General Assembly (UNGA) from 1960-2014, where greater vote agreement with a country indicates greater clout for that country. This approach is strong for several reasons, notably because of the availability of voting data across time and the fact that UNGA voting has been established to show affinity among countries (Voeten, 2013).

I analyze voting with respect to three key countries: Russia, Saudi Arabia, and the United States. Russia and Saudi Arabia, while both major petrostates, capture different characteristics. Russia has long opposed the United States especially during the Cold War. Saudi Arabia has focused less on global influence but has long been the leader of the Organization of Petroleum Exporting Countries (OPEC) and most importantly is a Group of 77 (G77) original UN voting bloc member. Conversely, the United States has typically opposed the views of major petrostates and occupies a unique position in both international affairs and voting. These characteristics certainly generalize but ultimately should still capture the fundamental relationship I examine.

To preview my results, I find support for my argument: Russia and Saudi Arabia achieve greater vote agreement with higher oil prices, while the United States loses votes. When I restrict my sample to only votes on which the United States and Russia disagreed or the United States and Saudi Arabia disagreed, I find even stronger evidence that oil prices influence voting, especially swaying the votes of oil importers. These results hold under multiple specifications for voting agreement and after controlling for country and year fixed effects, the G77 voting bloc, and other factors. This finding adds to the vast field on the political economy of oil, filling a much needed void on how oil prices influence international relations. More generally, my finding illustrates how global economic factors significantly influence international affairs.
The rest of this paper proceeds as follows. Section two reviews the main literature related to oil and oil prices. In section three I present my theoretical argument. Section four introduces my data, and section five presents my results. Section six concludes and offers implications.

1. Oil and International Relations: Review of Literature
While much research has examined how oil wealth affects countries domestically, less has analyzed oil in terms of international relations. Common areas of focus have ranged from how oil wealth leads to corruption or poor institutions (Bulte et al., 2005; Isham et al., 2005; Beck and Laven, 2006; Knack, 2009; Anthonsen et al., 2012; Sala-i-Martin and Subramanian, 2013; Wiens, 2013), diminishes democracy (Ross, 2001; Mahdavy, 1970; Crystal, 1990; Lake, 1992; Andersen and Ross 2014), and triggers conflict (Gleditsch, 2012; Koubi et al., 2014). Other research has examined social issues, such as the status of women (Assaad, 2004; Ross, 2008; Do et al., 2011), and government transparency (Egorov et al., 2009; Kolstad and Wiig, 2009; Williams, 2011).

In terms of international affairs, two findings pervade the literature. First, petrostates tend to engage in more militarized disputes and be more aggressive abroad than other countries (Colgan, 2010; Kellogg, 2016). Specifically, Colgan (2010) finds that petrostates “engage in militarized interstate disputes at a rate more than 50 percent higher than nonpetrostates.” This finding is unsurprising given the spending priorities of petrostates: Since 2000, oil producers comprise six of the top ten countries in military expenditure (Hendrix and Nolan, 2014) and have been known to use oil revenue windfalls largely for military spending (Looney, 1988). The second main finding concerns international integration. Ross and Voeten (2016) and Hendrix and Nolan (2014) find that petrostates tend to be less integrated in global governance institutions. In particular, Ross and Voeten (2016) show that states which export more oil tend to join fewer international organizations and are less likely to accept compulsory jurisdiction of international judicial bodies. Their argument centers on the lack of incentive for petrostates to join these agreements because oil, which dominates their domestic economies, is always in demand.

Oil prices have not been ignored completely, however, as price is a direct factor of oil wealth. For instance, as Wright et al. (2015) and Smith (2004) find that greater oil wealth deters democratic transitions and increases regime durability, they imply that higher oil prices should strengthen this relationship. Nonetheless, specific focus on prices, disaggregated from oil production, remains largely unexamined despite Friedman’s (2009) theory on the matter. Friedman (2009) states that for petrostates decreases in the price of oil should be accompanied by increases in the pace of freedom and more sensitivity to how one’s country is viewed internationally. Generally, then, Friedman argues oil prices mediate the effects of oil endowments.

The only known study to directly analyze this theory comes from Hendrix (2014), who examines how oil prices predict militarized disputes between states. Using a sample of 153 countries from 1947-2002, Hendrix (2014) finds a drop in oil prices causes petrostates to behave less aggressively whereas a rise engenders aggression. No significant impact from prices is found for non-oil exporters.

How oil prices influence international affairs more generally is an unstudied area to the best of my knowledge. Yet, it is also an area of great importance, especially given the vast fluctuations in oil prices. It is also an area superficially discussed in the media. For example, a 2008 The Wall Street Journal article states, “The long oil-price boom is posing wrenching challenges for the world’s poorest nations, while enriching and emboldening producers in the Middle East, Russia and Venezuela. Their increasing muscle has a flip side: a decline of U.S. clout in many parts of the world” (King, 2008). Other examples abound, but few if any discuss thoroughly or actually examine how oil prices influence the clout of states. This reasoning is what I discuss next.

2. How Do Oil Prices Affect Clout?
Scholars of international relations have defined power as the “ability to influence another to act in ways in which the entity would not have acted otherwise” (Wilson, 2008) or the “ability to affect others to obtain the outcomes you want” (Nye, 2014). Nye (1990) in particular discusses the distinction between “hard” and “soft” power. Hard power typically refers to
Coercion, focusing on military actions and economic sanctions for example (Art, 1996; Campbell, 2007; Cooper, 2004; Wagner, 2005). In contrast, soft power is based on persuasion or attraction and thus relies on culture and values that ultimately shape the preferences of other states (Nye, 1990).

Clout, I argue, is more general. It is neither soft nor hard power but rather a general factor that influences both hard and soft power. In other words, countries with greater clout should have greater ability to coerce and should also be more effective at persuading and attracting countries without threat. I maintain clout is affected by oil prices through two avenues: domestic economics and economic interdependence. These mechanisms are stronger for petrostates but also apply to oil importers.

2.1 Domestic Economics
Petrostates have several key characteristics, including large military and foreign aid spending, low taxation, centralized leadership, and economies dominated by oil revenues due to the “Dutch disease” (Karl, 1997). The Dutch disease is used to describe how oil wealth leads to deindustrialization because oil production drives up real wages and the value of the local currency (Davis, 1995; Prebisch, 1950; Sachs and Warner, 1995; Corden and Neary, 1982; Neary and van Winbergen, 1986). In effect, the economies of oil-rich states tend to become highly dependent on oil revenues (Harding and Venables, 2013). This development especially matters because fluctuations in oil prices can therefore have serious effects on government revenues and a country’s entire economy. As Figure 1 shows for Saudi Arabia, Gross Domestic Product (GDP) per capita growth has been correlated with changes in oil prices for the last 25 years.

![Figure 1. Saudi Arabia GDP and Oil Price Changes (1990-2014)](image)

High oil prices therefore benefit the economy, give leaders greater stability, and in turn allow petrostates to focus on spreading their influence abroad, in effect magnifying their clout. Examples of this relationship are numerous. Venezuelan President Hugo Chavez ramped up military spending to support the Revolutionary Armed Forces of Colombia (FARC) when oil prices popped in the early 2000s (Alvarez and Hanson, 2009). Saudi Arabia, Libya, and Iran also increased foreign funding to strategic groups under high oil prices, supporting the Afghan mujahedeen, over 30 insurgencies around the world, and Hezbollah and Hamas, respectively (Colgan, 2014).

When oil prices fall, however, petrostates must prioritize spending. This scenario is especially problematic given the sudden onslaught of weak currencies, growing debts, trade deficits, rising food prices, and budget deficits which strike petrostates under falling oil prices (Politico Magazine, 2016). Even further, the fiscal deficits are often severe because governments tend to increase spending quite quickly when oil prices rise (Hertog, 2007; Ross, 2012). As a result, with the primary goal to remain in power, leaders tend to focus on appeasing their populace. In other words, under falling oil prices, foreign aid and military spending are the easiest areas for petrostates to cut (Powell, 1993). For instance, Hugo
Chavez was also known to have slashed foreign aid when oil prices fell (Luhnow et al., 2007).

Not only do petrostates need to cut foreign spending, but they also need to focus on revitalizing their domestic economies under cheap oil. Notably, the Saudi’s raised taxes, reduced gas subsidies, and began a campaign to diversify their economy and triple its nonoil revenue by 2020 after oil prices fell below $30/bbl in early 2016 (Anderson, 2016). With a record budget deficit of $98 billion in 2015, the Kingdom also launched plans to sell five percent of state-owned oil giant Saudi Arabia Oil Co., better known as Saudi Aramco (Carey et al., 2016). Other Arab Gulf states also began looking to develop new industries under low oil prices, for example, committing to a greater allocation of defense spending to domestic companies (Stancati and Parasie, 2017). Outside of the Middle East, Russia’s influence fell as well in the 1980s and 1990s when low oil prices crippled its domestic economy, especially its banking industry (Holeywell, 2015).

For oil importers, the impact is reversed. Cheap oil provides government, business, and consumers with a windfall. For example, the economies of major importers India, Japan, and Vietnam have thrived under cheap oil environments (Holeywell, 2015). Conversely, higher prices reduce economic growth and increase inflation, something that slowed the United Kingdom’s recovery (Rowley, 2011). Ultimately, for oil importers, the strength of their domestic economies affects their government’s ability to spread influence abroad.

2.2 Economic Interdependence

The second avenue concerns economic interdependence. As Keohane and Nye (1977) discuss, a country’s interconnectedness abroad is dependent on its reliance on commodities for which demand is inelastic. Ross and Voeten (2016) develop this relationship further arguing the reason for joining international agreements and organizations for countries is to help their exports and encourage foreign investment. Petrostates, however, need neither of these, because of the inelastic demand for oil, which is incidentally also the main good these states sell given the Dutch disease. Further, petrostates typically have little challenge gaining access to foreign markets (Voeten and Ross, 2015) and attracting foreign investment because of the upside in oil investments (Jensen and Johnson, 2011). Oil producers also easily attract capital from sovereign debt markets (Wellhausen, 2015) and already sell their products, largely oil, with the smallest duties (World Trade Organization, 2011). Several examples illustrate these points. As Voeten and Ross (2015) discuss, Sudan appeared to be an unattractive destination for Foreign Direct Investment (FDI) in the early 2000s, but investments seeking to capture oil rents poured in. Venezuela as well received significant oil investments from 1999-2013, even though the country presented risk, especially after ExxonMobil and ConocoPhillips both had investments expropriated (Ross and Voeten, 2016).

All these factors demonstrate the need for petrostates to seek economic interdependence is less than that of typical countries. Yet, the costs of doing so still exist. To seek more regional integration, for instance, petrostates would need to remove import barriers (Venables, 2011). Attempts to attract greater FDI would also entail costly commitments to international institutions with goals of lowering trade barriers and the ability to monitor, deter cheating, and employ dispute settlement (Haas, 1964; Keohane, 1982; Ruggie, 1982; Haggard and Simmons, 1987). This logic has been applied to several cases, ranging from why oil-rich Norway did not join the European Union (Moses and Jenssen, 1998; Mattli, 2000), to the lack of regional integration among Arab states (Hoeckman and Messerlin, 2002), to the finding that oil exporters have been less likely to seek World Trade Organization (WTO) membership (Davis and Wilf, 2011).

Under high oil prices, any incentive for seeking economic integration falls for petrostates. For example, as Voeten (2016) discusses, Bolivia, Ecuador, and Venezuela, all left the International Centre for Settlement and Investment Disputes (ICSID) and canceled bilateral investment treaties under high oil prices. When oil prices fall, however, petrostates become less stable and, faced with a weakening economy, often seek new revenue sources. Suddenly, economic interdependence begins to gain more importance. Stubbs (2000) argues falling oil prices pushed oil-rich Indonesia to seek deeper cooperation with the Association of South East Asian Nations (ASEAN). Mexico as well joined the General Agreement on Trade and Tariffs (GATT), the North Atlantic Free Trade Agreement (NAFTA), the WTO, and the Organization for Economic Cooperation and Development (OECD) after oil prices fell in the 1980s (Ross and Voeten, 2016). Even Russian President Vladimir Putin stressed how cheap oil was forcing Russia to diversify its economy (Reuters, 2014).
For oil importers, the ability to spread their influence via economic interdependence is also influenced by prices. Generally, oil importers, who need a constant flow of oil for their economies, rely on interdependence and stable relations (Keohane and Nye, 1977). Under low prices, importers gain more leverage against petrostates, who need greater integration and cooperation to stabilize their economies. In turn, importers have greater ability to punish transgressions and spread their influence abroad. For example, while high oil prices hurt the United States’ ability to challenge Iran’s nuclear program (King et al., 2008), the two reached an agreement when prices dropped in 2015. Former United States Secretary of State John Kerry also called the American tight oil (or “shale” oil) boom a “game-changer” in global diplomacy, as it gave the United States greater ability to influence prices and cut the profits of major petrostates (Wang, 2015).
References


Davis, C. and Wilf, M. (2011). Joining the Club: Accession to the GATT/WTO. Available at SSRN.


Dataset.
Voeten, E. (2013). Data and Analyses of Voting in the UN General Assembly. Available at SSRN.
Wells, N. (2016). Americans Spent as Savings on Dining, Shopping … and More Gas. CNBC.
Wiens, D. (2013). Natural Resources and Institutional Development. Journal Theoretical Politics,


The Economic Efficacy of Reintegration Assistance for Former Child Soldiers

JOHNATHAN KAUFMANN
American University

ABSTRACT
With over 300,000 child soldiers estimated to be in armed conflict today, international organizations are dedicating millions of dollars to disarming, demobilizing, and reintegrating these children. However, no consensus exists among scholars on the efficacy of these programs, including how their services affect reintegration outcomes. This research is the first statistical analysis of the economic impacts of reintegration assistance for former child soldiers. Several regression analyses were performed to determine the effect of reintegration assistance on earnings and social capital. The data is from the Survey for War-Affected Youth Phase One, a survey conducted in Northern Uganda in 2006 that includes 462 former male child soldiers. The results indicate that no statistically significant relationship exists between reintegration assistance and earnings or social capital. This disheartening finding has broad implications for aid organizations that have continuously implemented these ineffective services for decades in locales across the globe. Not only does this analysis demonstrate the necessity of improvements to reintegration assistance programs, it can also provide some insight into how they might be enhanced in the future.
1. Introduction

There are an estimated 300,000 child soldiers in conflict today, fighting for over 86 different groups—including governments—in 19 countries, despite international condemnation (Spitzer and Twikirize 2013; Betancourt et al, 2010). Even if they are demobilized, many will return to rebel groups if they are not successfully reintegrated (Chrobok, 2005). However, existing aid programs focus mainly on demobilizing children while paying little attention to reintegration (Güven, Kapit-Spitalny, and Burde, 2014). This phenomenon arises because child soldiering is often treated as a security issue by the state and international community, and the focus is on disarming militia groups rather than rebuilding society. Successful reintegration of child soldiers is essential to the long-term stability of a nation, and should be a primary concern for any state in civil conflict (Annan et al, 2011).

Disarmament, Demobilization, and Reintegration (DDR) programs aim to facilitate the structured reintegration of former combatants into society (Annan, Brier, and Aryemo 2009). They attempt to ease adjustment to normal life in social and economic spheres, thus minimizing the impacts of soldiering. Therefore, successful reintegration for former child soldiers means mitigating or eliminating the differences between these children and their noncombatant peers. Some studies have found that assistance is effective at improving mental health outcomes, but the authors ignored economic factors (Betancourt, Pochan, and De La Soudiere, 2006). This research proposes to answer the question: How effective is reintegration assistance at improving economic outcomes for former child soldiers? Specifically, we explore how assistance services affect two identified factors of successful reintegration—earnings and social capital.

We used regression analysis to examine the effects of assistance. We performed these tests on data from the Survey for War-Affected Youth (SWAY) Phase One, a survey of 462 male youth in the northern region of Uganda conducted in 2005. Not only is SWAY the sole large-scale dataset on former child soldiers that addresses the identified factors of successful reintegration, it also presents a unique opportunity for research due to the quasi-natural experiment created by the arbitrary abduction practices of the Lord’s Resistance Army (LRA) during the recent Ugandan civil war (Blattman and Annan, 2010). While tragic, these random abductions offer a rare possibility to explore the effects of reintegration assistance on former child soldiers whose only differences from their civilian peers stem from the abduction and its consequences.

No significant relationship was found between assistance and earnings or social capital, and we conclude that assistance services do not improve reintegration outcomes. The regression also shows that the mother’s education and displacement are critical components in the earnings per month of former child soldiers. Similarly, traditional cleansing ceremonies and parental understanding had significant positive effects on social capital. Based on this analysis and substantial supporting literature, aid programs should focus on these societal factors if they wish to improve post-conflict conditions.

2. Literature Review

Previously published literature provides a basis for understanding the essential factors of successful reintegration. The literature demonstrates that to improve reintegration outcomes, aid organizations should target three areas of life: economic opportunity, social inclusion, and health. Economic opportunity is the ability of the former child soldier to find employment that pays a livable wage, which can be difficult in conflict or post-conflict economies. Humphreys and Weinstein found that assistance did not improve economic outcomes for former adult combatants (Humphreys and Weinstein, 2007). However, they did not explore the relationship between assistance and economic outcomes among children. Social inclusion is participating and being accepted into the community, as many child soldiers face exclusion and stigma when they return home. Health can be split into two distinct areas: mental health and physical health. Studies have demonstrated that more violent abductions result in greater symptoms of mental distress, and that abductions often include debilitating physical harm or sexual assault that results in sexually transmitted infections (Blattman and Annan, 2010; Machel, 2001). While the health of the child must be addressed for successful reintegration to be achieved, this is not a factor examined in this study. Instead, we will examine the impacts of reintegration assistance on economic opportunity and social inclusion, as there is a clear gap in the literature in these areas.

2.1 Economic Opportunity

Several DDR programs make economic opportunity their primary goal, and more programs are shifting towards this view.
These programs generally focus on increasing schooling or providing vocational assistance. Some provide microfinance loans and business skills courses, coupled with follow up support. It should be noted that some programs that offer additional schooling are mainly using it to improve social skills, rather than human capital. However, psychological counseling and medical assistance is also implemented, and can have positive externalities on the economic opportunity of former child soldiers (Tonheim 2014). Thus, we can see that reintegration assistance programs attempt to influence economic opportunity both directly, through schooling and vocational training, but also indirectly, through medical and mental health programs.

Some studies have explored the differences in economic opportunity between child soldiers and their noncombatant peers, using earnings per month as a dependent variable (Blattman and Annan, 2010). However, no study has examined the differences between child soldiers who accept reintegration assistance and those who do not. Another study explored assistance for adult former combatants, once again using income as a dependent variable, and found that assistance did not improve earnings potential (Humphreys and Weinstein, 2007). According to the literature, earnings is a good measurement of economic opportunity, and will be explored in this study.

There is substantial scholarship on labor market economics and the determinants of income. Both age and education are good indicators of potential earnings, as they capture the accumulation of human capital over time, through either formal education or work experience (Mincer, 1974). Similarly, mother's education plays a strong role in the future earnings potential of their children, due to its impact on prenatal maternal behavior, as well as financial stability during childhood (Currie and Moretti, 2003; King and Hill, 1993). Marital status is also both a strong predictor of the decision to work, as well as earnings potential (Hundley 2000). These variables are essential to labor market economics literature, and should be included in any study of income.

However, certain factors are unique to the context of conflict and post-conflict reconstruction. Refugee camps have unique labor markets, and scholars have noted that they reduce the potential for earnings (Blattman and Annan, 2008). This is because work is both difficult to find, and jobs are low-skill, replacing workers daily. A construction site might hire the first 10 workers who show up for the day, and a new group for the next day, making finding work somewhat of a lottery in locations with a high population of displaced persons.

Abduction age and abduction length can also potentially diminish earnings by disrupting schooling or the occupational ladder. Annan and Blattman (2010) find that the greatest impact of child soldiering is a reduction in human capital, likely from interruptions to schooling (Blattman and Annan, 2010). Other studies have demonstrated the damaging effects of distress on earnings (Smith, Schnurr, and Rosenheck, 2005).

Both traditional economic variables and factors specific to the post-conflict context of reintegration are essential for understanding earnings in this environment. In our analysis, it will be important to include these variables as controls, to ensure that no confounding effect is causing the estimates to be biased. We can see that assistance programs attempt to influence economic opportunity through the aid offered, and that income is a good measure of economic opportunity. We therefore intend to examine the impact of assistance on economic opportunity through earnings, using the factors described above as control variables.

2.2 Social Inclusion
Some scholars have argued that social inclusion should be the primary goal of reintegration assistance, as it can have positive externalities for both mental health and economic opportunity (Tonheim, 2014). Social capital scores, which center on social networks and trust, have been used to evaluate social inclusion previously (Annan et al, 2011). Higher social capital has also been linked to higher earnings, and it is therefore an important indicator to examine the effects of assistance on economic outcomes (Aguilera, 2005).

Many NGOs focus specifically on social inclusion. In Uganda and Sierra Leone, massive sensitization efforts were implemented to persuade the populace that child soldiers were not responsible for their actions in the bush (Williamson,
2006). This effort succeeded in increasing community acceptance. Originally, communities rejected returnees, worried they would “contaminate” other children and cause them to join armed groups. Due to sensitization efforts, they became more ready to accept child soldiers. The mental health services that aid organizations provide may have positive effects on the sociality of children, and should be studied further.

Traditional cleansing ceremonies, such as *mato oput* in Uganda, are forms of reconciliatory justice that can encourage community healing and increase social inclusion. These ceremonies have been linked to greater community acceptance among participants (Wessells and Monteiro, 2006). In Uganda, *mato oput* is an important part of the social reintegration process, and is essential to a study of social inclusion.

Age is also a factor in social inclusion. Those who have been abducted at earlier ages, especially during critical development periods, are likely to be more hostile and have greater flight or fight responses to environmental stimuli due to abnormalities in amygdala development (Pechtel et al, 2014). Older children, as well as those who have been back from the bush for longer, will have had more time to readjust to community life, improving their level of social inclusion. As mentioned previously, older children have also had more time to gain educational or occupational experience, which can raise social capital. One study found that education was an important factor in social inclusion, while another found that those who participated in the labor market experienced increased acceptance (Betancourt et al, 2008; Veale et al, 2013).

Violence to the family and to the child can also play a role in social inclusion. One study found that children who had more violent acts inflicted upon them were more likely to participate in the community, particularly politically (Blattman, 2009). Another study found that households that had experienced more violence were more likely to be involved in community activities (Bellows and Miguel, 2009). While somewhat counterintuitive, this literature demonstrates that experiencing violence can make an individual more social, and is likely positively associated with social capital.

Reintegration assistance programs attempt to ease adjustment into society by increasing the community acceptance of former child soldiers. However, these children often experience stigma and exclusion, which can negatively impact their ability to participate in the community and the labor market. The scholarly literature on social inclusion has provided us with factors affecting social inclusion, and the method by which to measure it. We therefore examine the impact of assistance on social inclusion through social capital, while controlling for the factors, enumerated above, that might confound our results.

3. Hypotheses
The previous literature has demonstrated the importance of economic opportunity and social inclusion for former child soldiers, as well as several potential control variables that could affect the results. Additionally, it has demonstrated that earnings per month is a strong indicator of economic opportunity, and that social capital is a good measurement of social inclusion. NGOs aim to improve these outcomes, along with health. However, the effectiveness of these services has fallen into dispute. One study demonstrated that reintegration assistance improves mental health outcomes for former child soldiers, and another study found that services do not improve economic reintegration outcomes for adult former combatants (Betancourt, Pochan, and De La Soudiere, 2006; Humphreys and Weinstein, 2007). Despite this, no statistical analysis has ever examined the relationship between assistance and economic outcomes for children. The literature acknowledges that services can be improved, but there seems to be the broad assumption that reintegration assistance aids former child soldiers. However, Shepler, using an ethnographic methodology, anecdotally suggests that those who accept assistance are not better off than their counterparts (Shepler, 2014). While there is a dearth of statistical evidence supporting either claim, we formulate the following hypotheses about the impact of assistance on economic opportunity and social inclusion:

**Economic Opportunity**
- \( H_0 \): Assistance will not improve earnings per month for former child soldiers
- \( H_1 \): Assistance will improve earnings per month for former child soldiers

**Social Inclusion**
H₀: Assistance will not improve social capital for former child soldiers
H₁: Assistance will improve social capital for former child soldiers

4. Data

The data used to test these hypotheses comes from the Survey for War-Affected Youth Phase One (SWAY).¹ Phase One of SWAY is a simple random sample of households in two northern districts of Uganda—Kitgum and Pader—in 2005 and 2006, and includes 741 male youth. 462 respondents were abducted by the LRA, and represent the sample in this study. The unit of analysis is individuals, and so there are 462 distinct observations—or cases—in this study. SWAY includes information on experiences during the conflict involving the Lord’s Resistance Army (LRA) in Uganda, as well as reintegration experiences. This survey involved extensive interviews with the subjects, and is potentially limited by response bias due to the sensitive nature of the topics probed (Blattman and Annan, 2010). However, this data set is reliable due to the simplicity of the questions and the size of the survey sample. Numerous studies have used this data set to generalize to the entire population of soldiers, indicating both its accuracy as a metric for the population of child soldiers generally, but also the comprehensive nature of the survey (Annan et al, 2011; Blattman, 2009; Blattman and Annan, 2010). Of the 426 former male child soldiers in the sample, 203 received reintegration assistance services, or 43.9%. Five different types of assistance services were identified: medical assistance, psychological counseling, skills or vocational training, cash loan or grants, and reintegration assistance bundles, which are bundles of goods such as farming tools. Of those who accepted services, 34.5% accepted medical assistance, 60.6% accepted psychological counseling, 2.5% accepted vocational or skills training, 3% accepted cash loans or grants, and 49.8% accepted reintegration assistance bundles.

This has two implications for this study. Firstly, it is difficult to examine each service for its individual effects, because such a large proportion of former child soldiers accepted multiple forms of assistance. Therefore, only the effects of services generally can be examined. Secondly, the percentage of child soldiers who accepted skills training and cash loans is extremely small, making it difficult to generalize the findings to those who accept this form of assistance. This is especially problematic because vocational training is the form of assistance of greatest interest in a study of the impact of services on earnings and social capital. However, many of the recipients were given goods and other reintegration assistance bundles, which would be expected to have an economic impact.

5. Empirical Methods

Using this data, two ordinary least squares regressions using Huber-White standard errors were performed. The earnings model takes the form:

\[ EPM_i = \beta_0 + \beta_1 \text{SERVICES}_i + \beta_2 \text{AGE}_i + \beta_3 \text{ED}_i + \beta_4 \text{MTHRED}_i + \beta_5 \text{MAR}_i + \beta_6 \text{CAMP}_i + \beta_7 \text{RVIOL}_i + \beta_8 \text{PVIOL}_i + \beta_9 \text{ABDAGE}_i + \beta_{10} \text{ABDLGTH}_i + u_i \]

where EPM is earnings per month in Ugandan Shillings; SERVICES is a binary variable expressing if assistance services were received; ED is education in years; MTHRED is mother’s education in years; MAR is a binary variable expressing marriage status; CAMP is a binary variable expressing if the individual is displaced in a refugee camp; RVIOL is an additive index of 6 types of violent acts inflicted upon the child; PVIOL is an additive index of 8 types of violent acts the child inflicted on others; ABDAGE is the age in years at first abduction (some children were re-abducted); ABDLGTH is the total length in months of all abductions, and u is the error term. Services, age, education, mother’s education, marital status, and abduction age are expected to be positively correlated with earnings, based on the literature detailed above. Displacement in a refugee camp, violence received, violence perpetrated, and abduction length are expected to be negatively correlated with earnings.

The social capital model takes the form:

\[ SC_i = \beta_0 + \beta_1 \text{SERVICES}_i + \beta_2 \text{CLEANSE}_i + \beta_3 \text{AGE}_i + \beta_4 \text{ABDAGE}_i + \beta_5 \text{ED}_i + \beta_6 \text{RVIOL}_i + \beta_7 \text{FVIOL}_i + \beta_8 \text{NOMTHR}_i + \beta_9 \text{NOFTHR}_i + \beta_{10} \text{ORPHAN}_i + \beta_{11} \text{ESC}_i + \beta_{12} \text{RELEASE}_i + \beta_{13} \text{INDOC}_i + \beta_{14} \text{LOYAL}_i + u_i \]

¹ This survey was conducted by Jeannie Annan and Chris Blattman. (Blattman and Annan 2010)
where SC is an additive index of social support ranging from 0 to 13; SERVICES is a binary variable expressing if assistance services were received; CLEANSE is a binary variable expressing if the child underwent a traditional cleansing ceremony; AGE is age in years; ABDAGE is the age in years at first abduction; ED is years of education; RVIOL is an additive index of 6 types of violent acts inflicted upon the child; FVIOL is an additive index of 5 types of violent acts inflicted on the child’s family; NOMTHR is a binary dependent variable expressing if the mother is not present; NOFTHR is a binary dependent variable expressing if the father is present; ORPHAN is an interaction term between NOMTHR and NOFTHR; ESC is a binary variable expressing if the child escaped the LRA after abduction; RELEASE is a binary variable expressing if the child was released by the LRA; INDOC is a binary variable expressing if the child was indoctrinated by the LRA; LOYAL is a binary variable expressing if the child ever felt loyal to the LRA; and \( u \) is the error term. ESC and RELEASE are being compared against being rescued from the LRA, and NOMTHR, NOFTHR, and ORPHAN are being compared against both parents being present. Based on previous scholarship, services, cleansing ceremonies, age, education, violence inflicted upon the family, violence inflicted upon the child, and escaping the LRA are expected to be positively correlated with social capital. Conversely, not having a mother, not having a father, being an orphan, being released from the LRA, indoctrination, abduction age, and loyalty are expected to be negatively correlated with social capital. However, ORPHAN may have a positive coefficient, as not having a father or mother could lead to being sent to other relatives. It is the total interaction term \( (\beta_8 \text{NOMTHR}_i + \beta_9 \text{NOFTHR}_i + \beta_{10} \text{ORPHAN}_i) \) that is expected to be negative.

6. Analysis
6.1 Comparative Statistics
In order to understand the impacts of reintegration assistance, it is important to know who receives services. It is possible that there are differences between the population that received assistance and the population that did not. Those who accept assistance are, on average, significantly more likely to be displaced, take part in traditional cleansing ceremonies, have both parents alive, have inflicted more violence upon themselves and their families, have inflicted more violence upon others, have escaped the LRA or been rescued, and are more likely to have been indoctrinated. They are also more likely to have felt loyal to the LRA during their abduction and have longer abductions. Conversely, those who have not received assistance are more likely to be older, married, more educated, have returned from the bush for longer, have been abducted at an older age, and are more likely to have been released by their LRA captors. While this is clearly not a simple random sample, these variables can be controlled for. In the following regressions, we control for the variables that impact both the independent and dependent variable to minimize the expected value of the error term \( u \). We acknowledge the potential for unobservable, entity-fixed omitted variables in cross-sectional data, which will be discussed in more detail later.

6.2 Impact of Assistance on Economic Opportunity
It was hypothesized that receiving reintegration services would increase earnings per month in Ugandan Shillings (USh). $1 USD had equivalent purchasing power to USh 512.18 in 2006, according to the World Bank (2006). The mean earnings per month of all former child soldiers was USh 17,200.54, roughly equivalent to $33.58, well below the poverty line. The average earnings per month of abductees who received reintegration assistance services was USh 13,506.90, while the average for abductees who had not received services was USh 20,095.96.

In the first model (See Figure 1.1), violent acts perpetrated, age at abduction, and length of abduction were not significant. An \( f \)-test demonstrates that age at abduction and length of abduction were not jointly significant (\( p=0.48 \)), and so these variables were excluded from the final model despite Annan and Blattman’s theory that longer abductions lower human capital due to time spent away from education (Blattman and Annan, 2010). Additionally, violence received loses significance when violence perpetrated is removed as a variable. Together the two terms are jointly significant (\( p<0.1 \)), and were therefore not excluded from the final model.

In the second model, we can see that assistance has no significant correlation with earnings. We therefore cannot reject the null hypothesis. Additionally, interactions terms between services and violence received, as well as services and displacement, did not demonstrate any significant remedial effect. This is a disheartening finding, as it indicates that these services do not make former child soldiers economically better off than their peers, despite the millions of dollars spent
on assistance. However, it is noteworthy that reducing displacement and increasing access to education, including for the
mother of the child, significantly raised earnings, because these are factors which governments and aid organizations could
potentially improve upon. While this was not unexpected, given the literature discussed, it has important implications for
aid organizations that wish to improve the assistance they offer.

### Figure 1.1: Earnings per Month Regression

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings per Month</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received Services</td>
<td>-268.7</td>
<td>417.3</td>
</tr>
<tr>
<td></td>
<td>(3,152)</td>
<td>(2,937)</td>
</tr>
<tr>
<td>Age</td>
<td>1,679*</td>
<td>1,426**</td>
</tr>
<tr>
<td></td>
<td>(868.5)</td>
<td>(645.4)</td>
</tr>
<tr>
<td>Education</td>
<td>3,016**</td>
<td>2,920**</td>
</tr>
<tr>
<td></td>
<td>(1,464)</td>
<td>(1,399)</td>
</tr>
<tr>
<td>Mother’s Education</td>
<td>2,257**</td>
<td>2,248**</td>
</tr>
<tr>
<td></td>
<td>(897.1)</td>
<td>(897.1)</td>
</tr>
<tr>
<td>Married</td>
<td>15,809**</td>
<td>15,896**</td>
</tr>
<tr>
<td></td>
<td>(6,836)</td>
<td>(6,666)</td>
</tr>
<tr>
<td>In Refugee Camp</td>
<td>-6,847*</td>
<td>-7,072*</td>
</tr>
<tr>
<td></td>
<td>(3,684)</td>
<td>(3,687)</td>
</tr>
<tr>
<td>Violent Acts Received</td>
<td>-2,383**</td>
<td>-2,307**</td>
</tr>
<tr>
<td></td>
<td>(1,039)</td>
<td>(1,068)</td>
</tr>
<tr>
<td>Violent Acts Perpetrated</td>
<td>1,316</td>
<td>1,352</td>
</tr>
<tr>
<td></td>
<td>(1,710)</td>
<td>(1,704)</td>
</tr>
<tr>
<td>Age at Abduction</td>
<td>-350.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(616.1)</td>
<td></td>
</tr>
<tr>
<td>Length of Abduction</td>
<td>79.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(110.3)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-35,850**</td>
<td>-34,885**</td>
</tr>
<tr>
<td></td>
<td>(17,388)</td>
<td>(17,325)</td>
</tr>
</tbody>
</table>

Observations                | 458      | 458      |
Adjusted R-squared          | 0.108    | 0.110    |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

While accepting assistance did not lead to greater earnings over peers, there is the potential that it could lead to greater
social capital. Higher social capital now could lead to greater future earnings, according to some studies (Aguilera 2005,
585). While assistance may not improve earnings shortly after receiving it, it may raise the potential for greater future
earnings through increased social capital. We therefore turn to an examination of the impacts of assistance on social
capital, with the knowledge that it has the potential to raise future earnings.
6.3 Impact of Assistance on Social Inclusion

It was hypothesized that receiving reintegration assistance improved social capital. The mean score on the social capital index was 5.56. For those who had received assistance services, the mean score was 5.58, while the mean score for those who had not was 5.55.
In our model (See Figure 2.2), reintegration assistance has no significant effect on the social capital of former child soldiers. We therefore cannot reject the null hypothesis that assistance has no impact on social capital. Additionally, interactions terms between services, indoctrination, not having a mother, and abduction age demonstrated no significant remedial effect. Once again, assistance has failed to lead to improved outcomes over peers. However, it is noteworthy that cleansing ceremonies and education are correlated with greater social capital. As suggested previously, the effects of cleansing ceremonies have become a subject of debate in the scholarly community. However, this analysis demonstrates an extremely strong positive relationship between mato opit and social capital, significant at 99%. This study confirms previous findings that promoting reconciliatory justice processes and increasing access to education will improve social capital, and they are therefore methods by which aid organizations can improve assistance.

7. Discussion
This analysis reveals that reintegration assistance does not improve economic reintegration outcomes. No statistically significant relationship exists between assistance and earnings or social capital. This has dramatic implications for aid organizations, who have continuously implemented these services across contexts, without substantially altering them (Legrand, 1999). However, there are several ways that reintegration assistance services could improve upon the aid they offer. This study confirms the findings of King and Hill (1993), as well as Currie and Moretti (2003), that mother’s education has a significant effect on the future earnings of their children (King and Hill, 1993; Currie and Moretti, 2003). Similarly, we see a positive relationship between traditional cleansing ceremonies and the social relationships a child soldier can form, as described in Boothby et al (2006). In the future, aid organizations can focus on these societal aspects—increasing access to education for women and promoting traditional cleansing ceremonies—in order to improve reintegration outcomes. Many studies have highlighted the effectiveness of societal interventions as opposed to individual ones. Betancourt et al (2010) found that societal protective factors were superior to individual ones (Betancourt et al, 2010). Similarly, Boothby et al suggests that any assistance taking place in an interim center would likely be just as effective if implemented at home (Boothby, Crawford, and Halperin, 2006). Our findings support this research, which situates reintegration outcomes in the broader context of the community.

However, there are several potential critiques of these results that should be addressed. Firstly, the data is cross-sectional, presenting several potential limitations. The first of which is that no causality can be established. We merely can observe that no correlation exists, which is a suggestive result. Secondly, we cannot say that reintegration assistance has no positive effect, just that it does not improve outcomes over those who did not accept assistance. This critique, then, hinges on a self-selection or endogeneity argument, by claiming people who are worse off seek assistance and become just as well off as their peers.

First, we address self-selection. We know that there is self-selection into reintegration assistance, but this selection seems to be predicated, based on the descriptive statistics, on traumatic experience rather than earnings or social relations. Those with substantially greater traumatic experience are likely going to be those who seek assistance. Indeed, this fits with the forms of assistance offered. Former child soldiers are receiving medical assistance and psychological counseling, which fits this trauma narrative. Those with more violent abductions seek assistance for their trauma in order to fix their physical or mental problems. This also refutes the potential for endogeneity, that those with lower earnings seek assistance to raise their earnings. Especially given the types of assistance offered, this narrative of endogeneity seems high unlikely.

However, it is impossible to rule out entity-fixed unobservable effects using cross-sectional data. We recognize the possibility that unobservable variables, such as merit, could influence the decision to accept services and alter earnings or social capital. We cannot control for these variables, though we can say that the unobservable is not trauma. While trauma is likely to influence earnings or social capital, as well as accepting assistance, our models include experiences that would cause trauma. Given that we can control for traumatic experiences, it is exceedingly unlikely that trauma is this unobservable variable.

8. Conclusion
This research builds on existing literature that demonstrates the inefficacy of assistance services by demonstrating no
significant relationship between services and improved economic outcomes. However, this conclusion is limited in that we cannot address unobservable entity-fixed factors, such as merit. A logistic regression should be employed to demonstrate the determinants of seeking assistance, which we expect is trauma-driven. Additionally, future studies should examine this question with panel data that includes surveys before and after assistance, to control for entity-fixed effects.

We can easily generalize these results to the Ugandan context, due to the reliability and generalizability of STAWAY. Assistance has remained largely the same across conflicts and years, despite the criticisms of services (Legrand, 1999). Because of this, along with a lack of superior data, we tentatively generalize this finding to the larger population of former child soldiers, until superior data is available and a more conclusive finding can be reached. However, these findings are confirmed across contexts through qualitative studies, such as the one performed by Shepler, who observed no difference in former child soldiers who received assistance and those who did not in Sierra Leone (Shepler, 2014). Similarly, Humphreys and Weinstein found that assistance did not improve economic outcomes for former adult combatants (Humphreys and Weinstein, 2007).

This research is limited in that it only explores male former child soldiers, and future studies should focus on females, who have unique abduction and reintegration experiences. Additionally, this study is unable to explore the efficacy of vocational training or cash grants, because these services were not frequently accepted among the sample population. Exploring microcredit and vocational training specifically, in conjunction with panel data, would be the best avenue for future research.

We provide new insights into the effectiveness of reintegration assistance: namely, it has no correlation with economic outcomes. We also demonstrated the importance of access to education, particularly for women, and traditional cleansing ceremonies. Given the wealth of literature that support the importance of these factors, we can conclude that focusing on these elements should improve reintegration outcomes. Our findings confirm other studies that indicate societally focused interventions garner greater success than individually focused ones. The low participation in assistance among former child soldiers only exacerbates the issues with individual-focused interventions. In the future, aid organizations should work to enhance their services by focusing on societal interventions, with the hopes of improving outcomes for all.
References
Legrand, J.C. 1999. Lessons Learned from the UNICEF Field Programmes for the Prevention of Recruitment,
Demobilization, and Reintegration of Child Soldiers. New York: UNICEF.
Why Uncertainty Matters: 
An Empirical Approach for Spain

ROBERT PÉREZ GARCIA 
ALEXANDRE DÍAZ CODINA 
ADAM LAAZIZ ALAOUI DESCARREGA 
VICTOR M. QUINTAS MARTÍNEZ 
ADRIA ROCA VALLS 
Universitat Pompeu Fabra

ABSTRACT
The goal of this paper is to quantify and study the effects of uncertainty in the main economic variables. Acknowledging the close ties between economics and politics, we have widened the scope of our research and introduced an innovative index that grasps economic and political uncertainty. The result is a comprehensible and coherent framework that can be used to better understand the lion’s share of the uncertainty effects on the economy of Spain.
1. Introduction
Being fully aware of the globalized world in which we are living; noting with deep concern the international role that recessions and political changes are playing; and deeply regretting the vast scope of the current unexpected major events such as terrorist attacks; we would like to highlight the uncertainty scenario generated by this framework.

In our hypothesis we affirm that economic agents are not indifferent towards this uncertainty, and here is where it comes our objective. The goal of this research project is to understand, measure and quantify uncertainty to be able to analyze its effects on the economy, and further predict future reactions to its variations. Guided by the established methodology, we build an uncertainty index, we analyze its relation to reality in order to prove its representativeness, and we finally calculate regressions of some economic indicators on it, to determine the impact that uncertainty has in our economy.

1.1 Review of Literature
The Discussion about policy uncertainty has been intensified during the Global Financial Crisis. It has received substantial attention as a determinant and important factor shaping the depth and the duration of the Recession. For instance, it has been stated that the decrease in the employment and output level during the time horizon from 2007–2009 comes basically from financial and uncertainty shocks (Stock and Watson, 2012). Moreover, the Federal Open Market Committee (2009) as well as the IMF hinted that uncertainty about both the U.S. and the European fiscal, regulatory and monetary policies, was transmitted to consumers and policy makers that finally caused a huge decline in the economic situation and a slow recovery afterwards.

We must consider three kinds of literature in this paper: the first is the research about the impact of uncertainty on growth and investment. Bernanke (1983) provides some theoretical background about the fact that during periods of high uncertainty firms use to delay investment and hiring. E. Schaal (2015) in his paper focuses on the mutual relation between unemployment and uncertainty. His findings suggest that volatility shocks intensify the reallocation process, provoking larger gross labor market flows and higher unemployment.

On the other hand, it is essential to point out that the concern about the study of policy uncertainty is something relatively new. For this reason there is some literature against the fact that shocks to firm-level risks are likely to be major drivers of the business cycle and hence the global economy (Bachmann and Bayer, 2013). The model they used features a “wait-and-see” effect—this concept is related with the fact that firms postpone investment in periods of high frequency volatility—for investment after a surprise increase in firm-level risk. They particularly state that risk fluctuations are too small to make the “wait-and-see” effect matter so as to change the business cycle fluctuations significantly. However one of the main problems that we observe in this paper is that they exclude other important agents that interact in global markets and that the model they used can be only applied assuming aggregate productivity shocks.

A second important concern is the fact that uncertainty is difficult to measure as it is not directly observable. In response to this issue some economists (Bloom et al.) have attempted to design a way of measurement in order to analyze and quantify the effect that uncertainty has on the economy— both in theory and in practice.

In the research done by S. R. Baker, N. Bloom and S. J. Davis (2016) they developed an Economic Policy Uncertainty index (EPU) for the U.S. since 1985 as a newspapers-based study. There is a growing literature about using text search methods, such as the business information tool Factiva and others. The research in newspapers’ articles widens the scope of research itself. Moreover, one of the main issues about uncertainty is the lack of several data sources as Bloom (2016) supports.

R. Casarin and F. Squazzoni (2013) in their paper used the front page banner headlines of three financial newspapers in order to examine the influence of bad news in stock market volatility as well as the dynamic correlation between them. They show that in reality there is a retroactive effect between press and markets as they influence each other.

Finally there is also a large literature about policy uncertainty. M. Friedman (1968) in “The Role of Monetary Policy” states the concern about policy uncertainty and disturbances that cause instability. He points out that monetary
policy must contribute to offset those disturbances and to generate monetary, regulatory and fiscal guarantees as well.

2. Theoretical Background

2.1 What is uncertainty?

When trying to define the concept itself, we could just say that uncertainty is lack of certainty, and that certainty is a sure and clear knowledge about something. But this definition is kind of vague and abstract, and we should try to bring it into our field so we can end up analyzing the real impact that uncertainty has in our economy. Aiming to be more precise, we now say that uncertainty is a situation in which the probability of an event occurring is not completely known. It is important to notice that we are giving the uncertain character to those scenarios in which not all the variables that affect the calculus of the probability are known.

Illustrating this argument with an example, it is easy to see that, on the one hand, the toss of a coin assuming fairness is facing a binomial distribution with probability equal to 0.5 known for sure, we are certain about it. But, on the other hand, imagine a country facing elections and the possible consequences of this political change in economic indicators. The point here is not if there is a high or low probability of this indicators changing. The point is that there are lots of political, social and economic variables affecting first the result of the elections and then its impact on the economy, making it quasi impossible to calculate a certain probability, so turning it into an uncertainty scenario.

The situation explained within the last paragraph was just a mere example to clarify the concept. But it can also be useful to understand that, according to the definition and pursuing its practical implications, uncertainty is translated in reality to unpredictability and imperfect forecast. And again, this is due to the fact that the probability of the outcome forecasted is not completely known because of the complexity and the enormous amount of variables affecting the event.

Finally, it is also important to remark that apart from political issues, there are other parameters such as economic recessions or unexpected major events that can lead to an uncertainty situation, being them its cause. Those will be explained in the following section.

2.2 Causes of Uncertainty

Following with the definition of the previous section, the inability to forecast the future outcomes depends mainly on three main parameters: recessions, political changes and unexpected major events.

2.2.1 Recessions

The existence of a generalized slowdown in the economy modifies greatly all the expected future outcomes. The change in the expectations of all the economic agents is especially relevant regarding the government. Literature suggests that incumbent politicians have strong incentives to pursue economic policies that guarantee low inflation, high growth and, in general terms, do not expand free trade (Chappell and Gonçalves, 2000; Conconi, Facchini and Zanardi, 2014). Nonetheless, the means to achieve those goals is the main uncertainty factor as new macroeconomic scenarios lead to new and unknown policies applied (i.e. TARP) (Pastor and Veronesi, 2012).

2.2.2 Political Changes

The fact that a country is facing elections clearly represents an increase of uncertainty per se. The main driver of this situation is the potential shift in power, which can modify greatly the economic context and the regulations affecting business investment and agent's forecasts. It is specially remarkable the increase of the uncertainty in close polls (i.e. Al-Gore vs. Bush).

2.2.3 Unexpected Major Events

Unexpected events of a major importance have proven key factors to understand variances in uncertainty over time. Literature has proven that events such as terrorist attacks have two sided effects. On the one hand they produce a short term shock to the economy. On the other hand, the uncertainty arising from those facts can lead to an increase in uncertainty and investment priorities shift prolonged in the mid-long term (see Bloom, 2014). To exemplify both claims, the existing
literature proves that the effects of 9/11 were mainly felt years after the attack in the form of a significant shift in resources expenditure (i.e. war against terror, homeland security... ) and modifications in fiscal and monetary policy which could have been the base of the current crisis (see Blomberg, 2010, or Makinen, 2002).

2.3 Theoretical Effects of Uncertainty
In this section we will briefly explain which are the theoretical effects of uncertainty in the economy. We will base our analysis on different economic models such as the Keynesian one or the Orthodox one. Furthermore, we will define the possible future effects over time differentiating from the short and the long run.

2.3.1 Negative Correlation with Uncertainty
On the one hand, we expect uncertainty to be negatively correlated with the majority of economic indicators. For example, we expect GDP and its components — investment, private spending, exports and imports — to have a negative relation with uncertainty. The theoretical reasons are that an increase in uncertainty may lead to: a decrease in investment because the returns of it might be unpredictable—based on our definition of uncertainty — a decrease in private spending due to the fear and lack of confidence on the future economic situation—one postpone his consumption today in order to secure his economic stability in an uncertain future — a decrease in exports caused by a lack of confidence of foreign buyers in domestic economy and a decrease in imports due to the reduction of domestic spending and investment. We expect all of the above to cause a general reduction of GDP. In this scenario, we expect central banks to release expansionary monetary policies, which will push down the short run interest rates in order to incentivize investment and consumption. Furthermore, an increase in uncertainty may lead investors to adopt less risky financial positions by changing their financial portfolio. Consequently, the demand for those more secure financial assets—like government bonds will increase and so their prices. This could mean that the interest rates — specially the ones related with long term assets — will decrease. Finally, we also expect the industrial production index and the consumer confidence to be negatively affected by uncertainty for all what stated before.

2.3.2 Positive Correlation with Uncertainty
On the other hand, there are some economic indicators which might be positively correlated with uncertainty. One of them is unemployment. In more uncertain periods, companies may freeze or cut down its budgets and postpone investments until economic forecasts become reliable again. Due to this reduction, we expect unemployment to increase. Also, the expansionary monetary policies mentioned before may go together with fiscal ones. In this case, expansionary fiscal policies may be released in order to stabilize the economic indicators and incentivize economic growth. In this sense, we expect the public spending to increase too. Furthermore, if uncertainty comes from domestic events, the domestic interest rates may increase to compensate foreign—and domestic—investors for its probably more risky assets. This might deviate our previous expectations about uncertainty effects on long term interest rates.

2.3.3 Unpredictable Effects of Uncertainty
There are still some economic indicators which remain unexplained because there is not a feasible explanation of the effects on them. One of them is inflation. Assuming that everything of what we have expected before happen, the effects on inflation are unpredictable. The expansionary monetary and fiscal policies might increase inflation, but at the same time the reduction in investment and the slowdown of the economic activity may reduce it. Another variable with unpredictable effects is stock market prices. Although the decrease in interest rates will make the stock market more attractive for investors, the lack of confidence and the slowdown of companies and economic activity might decrease stock market prices.

3. Building the index
Before the financial crisis, economists used to think of uncertainty as an abstract, subjective concept, too intangible to be measured. However, the recession brought about a great interest on the topic, so in recent years substantial advances have been made in quantifying and finding proxies for uncertainty. Following a paper by A. Moore (2016), in which he creates an uncertainty index for Australia, most of these measures of uncertainty fall in three broad categories: financial-based measures, newspaper-based measures and disagreement among forecasters. Since we try to construct an economic
uncertainty index as general as possible, that can be used in many different areas of economic study, we have selected one measure of each type:

1. **Financial-based**: Financial-based measures of uncertainty account mostly for the volatility of different financial variables, which can include stock market returns, prices, exchange rates, interest rates, etc. The advantages of these financial-based measures are that they are available with high frequency and that they usually have equivalents in other countries to compare them internationally. For the case of Spain, we have the I3E Index, elaborated by the International Center for Decision Making (ICDM) and IESE. I3E incorporates daily growth of four variables: stock market prices (IBEX 35), euro-dollar exchange rate, the price of Brent crude oil, as it is quoted in the Electronic Continental Exchange (ICE) and the price of 10-years Spanish government bond. The index is available monthly since January 2000.

2. **Newspaper-based**: A very recent and innovative approach in measuring economic uncertainty consists in performing an extensive search of the term “uncertainty” in newspapers, followed by other economic-related words. This method was introduced by Baker, Bloom and Davis (2016), and it has been proved to be a very good proxy for uncertainty. The same authors, Baker, Bloom and Davis (from Northwestern, Stanford and Chicago universities respectively), publish a monthly newspaper-based index, EPU, of economic policy uncertainty in several countries. While financial volatility is perhaps more tangible and objective, the subjective component of uncertainty that newspaper-based indices measure is also important, as it can have a great impact on economic agents when taking decisions. Moreover, the authors showed that, at least for the USA, EPU is very correlated with other measures of uncertainty (for example, VIX index). In our own uncertainty index we have incorporated the Spanish EPU, which includes *El País* and *El Mundo* as reference newspapers and is available since January 2001.

3. **Disagreement among forecasters**: A third proxy for uncertainty is forecast dispersion, that is, how far the predictions of professional forecasters regarding economic variables (GDP, employment, etc.) are from one another. Baker, Bloom and Davis (2016) used Federal Reserve Bank of Philadelphia’s Survey of Professional Forecasters (difference between 75th and 25th percentile of forecasts for several variables) in their EPU for the USA, and Moore (2016) calculated standard deviations between forecasts and incorporated them in his uncertainty index for Australia. Since we did not have a comparable index of forecast dispersion for Spain, we calculated it ourselves based on the *Panel de Previsiones de FUNCAS*, a forecaster survey that is published bimonthly since 1999. It includes forecasts of around 25 public and private organizations (including IMF, OECD, Spanish Government, BBVA, Santander, La Caixa, Repsol, etc.), and regarding around 16 variables of Spanish economy (GDP, components of aggregate demand, prices, unemployment, etc.). Using this data, we computed the variation coefficient for each variable in the survey (the standard deviation put in proportion to the mean of the forecasts), and then we aggregated them using geometric means (not arithmetic because they were too sensible to a single variable having huge dispersion). We will refer to this bimonthly index of forecast dispersion as “DISP”.

94
After deciding which financial-based, newspaper-based and disagreement among forecasters measures (I3E, EPU and DISP respectively) our index will incorporate, we have scaled them so that they take values from 0 to 100 (not included) with the following formula:

$$
X_{scaled} = \frac{X_t - 0.95 \times \text{min}_t X_t}{1.05 \times \text{max}_t X_t - 0.95 \times \text{min}_t X_t}.
$$

Then, we have combined them using 10 different specifications of weights, obtaining thus 13 different uncertainty indexes. The indexes are bimonthly (since the Panel FUNCAS is only available with that periodicity) and they begin in January 2001 (the first data we have for EPU). We present their specifications in Appendix A.

4. Index Analysis
In the previous section, we constructed 10 different uncertainty indices for Spain, that add up to the 3 raw measures of uncertainty. In Section 5, we try the explanatory power of these 13 indices and measures on several variables of the Spanish economy in order to select the one that best fits our data. We conclude, based on the significance of the models, that the most explanatory uncertainty index is C31 (see further explanations in Section 5). In the following section we will analyze several characteristics of our chosen uncertainty index (C31), including its behavior in comparison to major events occurred, and its relation to other uncertainty measures (VIX) or uncertainty measures for other countries.
4.1 Adjustment to Reality

Mark Twain once said: “Facts are stubborn, but statistics are more pliable”. As thoughtful economists wishing to adjust our model to the reality in the most plausible way, we have included this section to assure that our statistical model has economical meaning when compared with reality. Recovering the identification of the causes of uncertainty in Section 2.2., we now have the tools to analyze the overall background. Thus, arising from the three origins of uncertainty, we can clearly identify two periods.

**Figure 2:** Our chosen uncertainty index, C31.

**Figure 3:** Impact of major events in our index, C31.

A: Irak War
B: 11M + Elections [change in political make-up]
C: Lehman Brothers – First stages of the Global Meltdown
D: European Debt Crisis
E: Greece Debt Crisis (II): Syriza vs. EU + Nationwide autonomic and local elections (15M)
F: Spain General Elections: The rise of the outsiders
4.1.1 Pre-Meltdown Époque
Between 2000 and 2007, Spain enjoyed a buoyant economic growth. Nonetheless, the political landscape and the third factor: major unexpected events, appeared to be the most important uncertainty drivers. The importance of this third factor in the Spanish uncertainty levels appeared first as a result of the convergence of different violent events.

On the one hand, it must be taken into account that during all these period, ETA, a terrorist organization, maintained operations and continued to commit attacks in Spanish territory. On the other hand, there was a branch of violence originated from international situations. The initial point could be considered the 11th September terrorist attack in the USA. As retaliation, the USA government declared the war to Afghanistan in 2001. Shortly after, the Spanish government decided to join the international coalition. Furthermore, following with the war against terror, in 2003, USA, with the support of Spain, launched the invasion with Iraq. It is remarkable that both decisions were taken without consensus and thus created a widespread refuse of the citizenship that worsened with time.

The year 2004 was a turning point in terms of violence and political landscape in Spain. In 11th March, three days prior to elections, an Islamist terrorist group perpetrated a large-scale terrorist attack in Madrid. It is especially remarkable the reaction of the stock exchange to the attack.

As seen in the graph, the market experienced a serious fall in the aftermath of the attacks. Nonetheless, the successive appearances of government officials explaining the situation allowed the market to recover enough confidence to reduce the gap significantly.

It is essential to study the management of crisis by the government in order to understand the deep political shift that occurred. Due to the mismanagement of the government's communication, people's confidence in their government decreased substantially. Furthermore, the attacks put the spot on the Iraq war, as people felt that Spain had no need to be in that war and that it was the cause of all the suffering. Following the elections in which Zapatero won the presidency of Spain, the country did not suffer any other major terrorist attacks.

4.1.2 Global Crisis (2008-2015)
Inception: As in the previous époque, this period of Spain's history began in the other part of the Atlantic: USA. Due to the expansive monetary policies in the first decade of the 2000's, cheap credit flew throughout the global economy. The most prominent financial institutions all across the world began the commercialisation of complex financial products called CDO, which, in essence, were lot more risky than what rating agencies believed. Once the confidence in those products broke and the music stopped playing, the financial system collapsed. The most prominent example was the fall of Lehman Brothers in September 2008.

Focusing in Spain, the effects of the cheap credit burst were felt through two main drivers. Firstly, due to the sudden discovery that the housing market had a massive oversupply, the housing prices plummeted. This occasioned the loss of huge amounts of family's wealth and, in addition, the destruction of the most worker intensive industry in Spain: construction. Secondly, the exposure of the banking sector to the drastic shutdown of the financial liquidity international markets while facing increasing demands from regulators to cover losses from real state loans.

The whole impact of the virulence of the crisis can be seen in the graph. Beginning in 2007, we can clearly see a fast growth of the uncertainty index reaching maximum levels during late 2008 and beginning 2009.

Mirage: In late 2009 and 2010, the macroeconomic variables seemed to recover and thus the confidence in the economy. Specifically, in the first trimester of 2010, the Spanish GDP grew 0.1%.

Back to Reality: Elections and Recession: In July 2011, Zapatero's administration calls on early elections that will give Mariano Rajoy the presidency. Furthermore, the anemic growth of 2010 disappears and the macroeconomic environment worsens again.
Precisely in 2011, Greece’s looming default deployed fear all over Europe about a potential spread to other countries such as Spain. This is clearly portrayed in our graph when we see that 2011 and 2012 are critical points that represent the maximum of our series. 2012 was a year in which the social upheaval and unconventional fiscal policies were the main channels for uncertainty. Furthermore, constant rumors that questioned the capability of the Spanish banking system eroded even more the confidence of the international agents in Spain. In fact, in May 2012, the Spanish Government is forced by Europe to bail out Bankia to avoid a structural collapse of the economy.

Recovery: The époque between 2013 and 2016 comprises mainly the recovery time.

4.2 Comparison to VIX

Before starting with a proper comparison, we first need to agree on what is VIX and how is it related with uncertainty. VIX is a volatility index calculated by the Chicago Board Options Exchange (CBOE). Its calculation is based on the implicit volatility of the S&P 500 options, which leads to an inversely proportional relation between both index. This basically means that when volatility is high, VIX is high too which reflect the pessimist feeling of the market and is directly correlated with falls in S&P 500.

But how do we relate it with uncertainty? As stated above, VIX reflects the volatility of the market and, at the same time, volatility reflects the level of dispersion between forecasted possible outcomes by the investors. We already implicitly stated within the description of indicators (see Section 3) that, and according to our definition of uncertainty, the dispersion between forecasts is a clear gauge about the ignorance to the possible outcomes of a specific situation, which clearly describes an uncertainty scenario.

Looking now at the graph, we can divide it in two stages. The first one would go from 2001 to the beginning of 2009, and it would be characterized by the strong correlation between our index and VIX. This basically points out that, despite giving a 33.3% weight to the financial indicator, which is a relatively low percentage, the financial stability still have a huge importance regarding the uncertainty.

The second stage would go from 2009, more specifically from the most atrocious point of the financial crisis, until now. It is important to notice that the correlation in this phase decreases, which not only reflects the differences between the measures applied by the Federal Reserve, versus the ones taken by the Spanish Government and the European Central
Bank, but also the political, social and economic consequences that those measures caused. In other words, we can see in the graph that while VIX immediately started decreasing in 2009, our index remained constant in such high uncertainty points (the financial stability seems not to be that important anymore). This is not just due to the much slower recovery of the Spanish financial market, but also because of the weight that the economic and social-political indicators gained in this stage.

The reasons and causes behind this increase have already been explained in the previous section, “Adjustment to Reality”. But to conclude, what we can subtract from the comparison between a financial volatility index and our index, is that uncertainty can be caused by many diverse variables depending on the national and international contemporary framework. So the financial situation is not the only one that can cause uncertainty, and it is important to be careful and fully aware of the possible consequences of incurring in some political acts.

4.3 Foreign Component

![Figure 5: Foreign and domestic components of our index, C31.](image)

We know that Spain is one of the most important countries inside the European Union but we cannot consider it as a big economy. Even though there is empirical evidence—as stated by the CES (2015)—the level of openness (a measure as a percentage of GDP) has changed significantly from 37.6 in 1986 to 61.7 in 2014, particularly thanks to the integration to the EMU. However it intrinsically implies that the economy of Spain is more vulnerable to foreign shocks. In this part of the paper we try to analyze these effects and to understand how vulnerable it is.

We tried to disaggregate our uncertainty index C31 in a domestic and a foreign component using the same method as A. Moore (2016) did for Australia. We have regressed our Spanish index, C31, on the EPU for USA, Germany, Italy, United Kingdom, France, Japan, China and Russia. We have used only the EPU for other countries because it is the most uniform and widely available measure of uncertainty. The only significant coefficient was the one on the American EPU, with a p-value < 0.001. That might be due to the fact that the EPU for the USA includes not only a newspaper-based measure of uncertainty, but also forecast dispersion and financial measures.

With these results, we treated the fitted values as the “foreign component” of uncertainty in Spain (green line), and the residuals of the regression as the “domestic component” (red line). That is, the foreign component represents the part that can be explained linearly by the evolution of uncertainty in other countries, while the domestic component is the part that
cannot be explained (the real value of the Spanish index minus the explained part). The formulas used for the foreign and
domestic component are, respectively:

\[ \text{FOR}_t = \beta_0 + \beta_1 X_t \]
\[ \text{DOM}_t = C_{31t} - \text{FOR}_t \]

where \( X_t \) is a vector containing the EPU for USA, Germany, Italy, United Kingdom, France, Japan, China and Russia, and
the coefficients of the upper regression are estimated using OLS. We present the results of the regression and more data
about foreign indices in Appendix B.

Firstly we can observe that during the period 2001–2008 the foreign component lines up several times with our index
while other times it is even larger, and as a consequence the domestic component goes between −20 and 5. It means that
during this period foreign uncertainty seems to have been larger than domestic uncertainty. However as we have said
on average it lines up accurately. When the financial crisis began we can appreciate a slight change in its behaviour. The
domestic component makes a peak between 2008 and 2009 and it continues positive until at halfway of 2010. It means that
uncertainty in Spain was larger than in other countries. Nowadays we can say that Spain is behaving more or less as other
countries even though it seems to be some correction.

All of before can be explained by the location of events that caused these shocks in uncertainty and how their economic
consequences affected Spain. In 2001, we find a peak in foreign component probably caused by 9/11 terrorist attack. The
following periods of high foreign uncertainty may be explained by Afghanistan and Iraq wars. Another huge peak in
foreign component appears in 2008 with the fall of Lehman Brothers and its consequences. However, the effects of the
crisis since this point were more heavily suffered in Spain than in USA, whose recovery was quicker. This fact explains the
higher weight of the domestic component in the following periods.

5 Empirical Study

5.1 Data and Methodology

The main purpose of constructing a measure of uncertainty is to estimate quantitatively the effects that uncertainty has
on the economy. In order to do so, we have compiled a dataset of economic variables, which we will try to explain with our
uncertainty index. In Appendix C, a table can be found explaining the meaning, the provider and the availability of each
variable in the dataset.

Once we have the data, a first approach to calculating the effect of uncertainty on these economic variables would be to
estimate the following model:

\[ Y_{t+k} = \beta_0 + \beta_1 X_t + u_{t+k} \]

where \( Y_{t+k} \) is the value of one of our dependent variables, \( k \) periods ahead from \( X_t \), the value of one of our uncertainty
indices, and where \( u \) represents an error term under the usual assumptions.

However, the last model would not yield consistent estimators, because we are dealing with time series data, and the OLS
assumption that the values of \( X_t \) and of \( Y_t \) are independent and identically distributed is not satisfied. In particular, they
are not independent (the value of \( X \) or \( Y \) in moment \( t \) depends on past values of \( X \) or \( Y \) ). One basic solution to that problem
would be to regress not the values of the variables, but their growth rates (which we will denote with the corresponding
lowercase letters, \( x_t, y_t \)), defined as:

\[ y_t \equiv \frac{Y_t - Y_{t-1}}{Y_t} \times 100, \quad x_t \equiv \frac{X_t - X_{t-1}}{X_t} \times 100. \]
Also, to make sure that the effect is related to the uncertainty index and not to past values of the same variable (the most likely cause of bias if omitted), we have included the growth rate of the dependent variable in the same moment as the one of the uncertainty index as a regressor.

The first model we estimated is, therefore:

\[ y_{t+k} = \beta_0 + \beta_1 x_t + \beta_2 y_{t-1} + u_{t+k}. \]  \hspace{1cm} (1)

For the dependent variables IPI, UNEMP, INFL, STCK, INT1, INT12, INT10 and CONF, of which we had monthly data, \( t \) is measured as bimonthly periods (the same as our index), and \( k = 1, 2, 3, 4, 5, 6, 9, 12, 15, 18 \). For the dependent variables GDP, PRV, PUB, GFCF, IMP and EXP, of which we only had quarterly data, \( t \) is measured quarterly and \( k = 1, 2, 3, 4, 5, 6, 7, 8 \). To make sure that we did not omit other variables, we estimated a second model:

\[ y_{t+k} = \delta_0 + \delta_1 x_t + \delta_2 y_{t-1} + \delta_3 c_t + \varepsilon_{t+k}, \]  \hspace{1cm} (2)

where everything is the same except for the fact that we added a vector of controls from our dataset, \( c_t \), in the same moment \( t \) as the uncertainty index and the lagged value of the dependent variable, and also in growth rates. That yields a total of \((8 \text{ monthly variables} \times 10 \text{ time lags} + 6 \text{ quarterly variables} \times 8 \text{ time lags}) \times 13 \text{ uncertainty indices} \times 2 \text{ models} = 3328 \) regressions.

5.2 Interpretation

Based on the \( F \)-statistics that we obtained for these regressions, we concluded that the uncertainty index with a highest explanatory power in our set is C31. We have noticed that our index explains significantly some of our dependent variables and that some of them meet our expectations about possible effects. We present the regressions corresponding to the index C31 in the Appendix D, the others you can find compiled in the Online Appendix.\(^1\)

To ease and make more visual the results of our regressions we have plotted the different coefficients in a time horizon for each of the dependent variables. Each plot includes the coefficients and a 90% confidence interval for each of them.

5.2.1 Industrial production index (IPI)

![Figure 6: Coefficients of the regression of IPI on C31.](https://drive.google.com/file/d/0B84wZwwQNe3hWTRMTG1CazBMeIE/view?usp=sharing)
These results show that our initial expectations were wrong. Our prediction was that uncertainty may have a negative effect on IPI. We can explain these results thinking that industrial production levels are not very flexible and it is not easy for companies to adjust its production in the short term. We can observe in the graph that the significant coefficients in $t+9$ and $t+15$ are negative.

5.2.2 Unemployment (UNEMP)

![Figure 7: Coefficients of the regression of UNEMP on C31.](image)

In this case, the results validate our original predictions which were that unemployment and uncertainty are positively correlated. Although some of the coefficients are not significant, almost all of them are positive and in particular, the significant ones of $t+4$ and $t+9$. This ones show that an increase of a 1% in our uncertainty index will lead to a 0.03% increase in unemployment in the following 3 to 9 months. Based on the data from the INE for the first trimester of 2016, this change will mean an increase of approximately 1,437 unemployed people.

5.2.3 Inflation (INFL)

![Figure 8: Coefficients of the regression of INFL on C31.](image)

These results match perfectly with our predictions given that just one of the coefficients is significant, in this case positive. This means that we can not refuse the hypothesis that uncertainty has no effect on inflation, or equivalently, that the effects
are neither positive or negative.

5.2.4 Stock prices (STCK)

![Figure 9: Coefficients of the regression of STCK on C31.](image)

Although the tendency is that an increase in uncertainty will lead to a decrease in stock prices as of period \( t + 2 \), none of the coefficients is significant. This could be explained by the irrationality of stock markets, cause as Keynes once said (1930): “Markets can stay irrational longer than you can stay solvent.”

5.2.5 One Month Euribor (INT1)

![Figure 10: Coefficients of the regression of INT1 on C31.](image)

Euribor indicates the average interest rate at which Eurozone banks offer to lend money to other banks. As its calculation is based on the 24 main European banks, it is a market interest rate and not one fixed by the ECB, and this is exactly why we think it is interesting to study this economic variable.

As we can see in the graph, its behaviour is the expected. A percentual increase of uncertainty, initially \( (t + 2) \) cause an increase on the “1-month Euribor”. We think that this responds to the principle “risk needs to be paid”, as an uncertain situation decreases the capacity of risk controlling, thus turning the scenario in a more risky one.

The following decreasing tendency just explain how current uncertainty loses importance over time.
5.2.6 Twelve Month Euribor (INT 12)

![Figure 11: Coefficients of the regression of INT12 on C31.](image)

The behaviour of the “12-month Euribor” is very similar to the previous one. The shape of the graph basically responds to the same principle of the “1-month Euribor”, but this time it takes a little bit longer to notice the positive correlation ($t + 3$). This is due to the fact that the INT12 is much less sensitive than the INT1 because of the temporal horizon embraced.

5.2.7 Interest Rate on 10-years Spanish Government Bonds (INT 10)

![Figure 12: Coefficients of the regression of INT10 on C31.](image)

The interest rate on 10-years Spanish Government Bonds also responds approximately as expected. As we can see in the graph, a 1% increase on the current uncertainty is translated to a 0.1% – 0.15% interest rate increase at $t + 3$ (bimonthly regression), which periodically decreases until $t + 6 - t + 9$ where the current uncertainty starts losing importance.

The interest rate basically reflects the likelihood that the debt (issued as a bond) will be repaid. Taking into account that our index includes economic and financial but also social-political variables, it is normal to perceive an increase of uncertainty as a decrease of the likelihood to repay by the government. In other words, an increase of uncertainty leads to a decrease of the investors (bond-buyers) confidence on the government, which is reflected by an increase of the bond interest rate.
5.2.8 Consumer confidence (CONF)

We can see in the graph that the negative effect of uncertainty on consumer confidence that we predicted seems to take place in periods $t + 2$ and $t + 3$, that is, between 4 and 6 months after an 1% increase in uncertainty, consumer confidence tends to decrease a 2% and then again a 1.5%. The good news is that the negative effect takes place only in the short term, and in a little more than half a year, consumer confidence begins to recover from a negative shock in uncertainty. One explanation to such short persistence of consumer pessimism might be that, according to Milton Friedman’s (1956) permanent income theory, consumers might try to smooth consumption over time, and save during the good times to avoid decreasing their consumption by much in hard times (which tend to be also the most uncertain).

5.2.9 Real GDP (GDP)

The effect of uncertainty on GDP seems to be negative, and it gets worse until the periods $t + 5$ and $t + 6$ (now measured in quarters, that means a year and one or two quarters ahead respectively), while recovery or slowness of GDP fall begins in period $t + 7$ after an uncertainty shock. However, we see also that most of the coefficients are not significant at the 10% level. The ones that are seem to be ridiculously small, yet a decrease of 0.05% of GDP means e58 thousands of millions. Another characteristic that comes to our attention is that GDP is one of the variables where the negative effect of the uncertainty shock arrives with a longer delay. This might be because GDP represents the aggregate output of the whole economy, and, even if there is an increase in uncertainty, it takes a lot of time for the whole economic engine to adapt and respond to the new situation.
5.2.10 Private Expenditure (PRV)

![Figure 15: Coefficients of the regression of PRV on C31.](image)

The previous graph is totally consistent with the outcome that could be expected from the private agents when facing an economic downturn. The worsening of the economic forecasts incentivizes agents to postpone consumption and accumulate savings. Furthermore, the real net worth of families tend to decrease significantly during crisis (i.e. falling housing prices) and the lending opportunities become more scarce and more expensive. Nonetheless, the effects aforementioned vanish gradually until $t+7$ where private expenditure grows again supported and supporting an economic turnaround.

5.2.11 Public Expenditure (PUB)

![Figure 16: Coefficients of the regression of PUB on C31.](image)

The evolution of the public expenditure in the figure represented above presents a mixed but coherent picture with the predicted situation. The unleashment of an economic crisis activates the automatic stabilizers; mainly unemployment subsidies and transfers to social programs. Arising from this fact, we can see that between $t+1$ and $t+6$, the government increased or maintained the public expenditure. Moreover, it must be remarked that, as stated in WB report on public expenditure in the aftermath in crisis, the critical point is when governments have to pursue a fiscal consolidation (namely $t+7$).
5.2.12 Gross Fixed Capital Formation (GFCF)

![Figure 17: Coefficients of the regression of GFCF on C31.](image)

Given that all investment is backed up by its returns and this returns become unpredictable as uncertainty increase; our predictions match with our results. We can see that a 1% increase in uncertainty today is translated to a constant decrease in investment as of \( t + 3 \) (quarterly variable, 9 months later) to \( t + 6 \) of almost \(-0.02\%\) each period. This delayed effect may be explained by the difficulty to adjust the fixed capital amount given its nature and its cost.

5.2.13 Exports (EXP) and Imports (IMP)

![Figure 18: Coefficients of the regression of EXP on C31.](image)

![Figure 19: Coefficients of the regression of IMP on C31.](image)
As we can observe exports and imports respond extremely well according to our initial expectations. From $t + 1$ to $t + 6$ the coefficients are negative for both and they behave in a similar way. As we had initially predicted the negative relation is due to the fact that an increase in uncertainty also implies that foreign buyers are less confident about the domestic economy as well as a decrease in imports because of the contraction in domestic and foreign expenditure. Moreover, foreign uncertainty shocks also harms exports as discussed by D. Taglioni and V. Zavacka (2012). On the other hand it is not until $t + 7$ and $t + 8$ (a year and three quarters and two years respectively) that both variables start the recovery. One possible explanation is that as well as in the case of real GDP, imports and exports depend on the recovery of the economy as a whole. An important appreciation is that in the exports’ graph only $t + 3$, $t + 4$, $t + 6$ are significant at 10% level while in the imports’ graph only $t + 3$, $t + 4$ are significant. We can quantify this effect as a 1% increase in the uncertainty index leads to a decrease in exports of approximately 0.02% (51 M.e) and a decrease in imports of approximately 0.015% (42 M.e) in period $t + 4$.

6. Conclusions
Along this paper we try to understand which are the effects of uncertainty in the Spanish Economy. To answer this question, we construct an index based on 3 uncertainty measures, which grasp the effects of political and economic uncertainty, and is capable of predicting the effects of the uncertainty on the main economic variables of Spain.

The efficiency of our index is due to the fact that we have included the three main estimators of the causes of uncertainty: financial-based measures for recessions, newspaper-based measures for political turmoil and unpredicted events and disagreement among forecast for recessions and unpredicted events. The combination of the three elements is able to grasp the effect of uncertainty in the economy so we can estimate the effects over different timeframes.

After careful examination, the predicted effects of uncertainty have proven, in general, to be a good indicator of the overall behaviour of the economy when uncertainty strikes. Thus, we observe that there is a negative correlation, specially in the first stages of the shock, with GDP, consumer confidence, private investment and Gross Fixed Capital Formation. On the other hand, we see a clear positive correlation with Unemployment and Public Expenditure.

To further illustrate the practical applications of our work, one remarkable example of our investigative efforts is that we can now quantify approximately the economic costs associated with current topics such as the call for new elections. Hence, quantifying the costs and providing better information to the relevant political actors.

Needless to say, however, much efforts remain to be dedicated to this topic before economists reach the model which can predict the uncertainty impacts with consistency, univocity and precision. Our approach has been constrained by technical limitations (mainly our lack of expertise in the statistical method) and by data gathering constraints (mainly number of observations and periodicity of data). To sum up with, we hope that this paper has shed some light on the interrelations between uncertainty and economic activity, and that it is a first step towards a more integrated decision making framework in which all this costs are taken into account.
References
(Consejo Económico y Social De España), CES (2015). “Informe sobre la competitividad de la empresa española en el mercado doméstico”. In: Colección Informes 1.
Why don't countries sue more often at the WTO?  
An Analysis of Specific Trade Concerns in the SPS Committee

MATHISON S. CLORE  
Georgetown University

ABSTRACT

Why don't countries sue each other more often at the World Trade Organization (WTO)? Despite increased complexity in the international trade regime and the proliferation of non-tariff barriers, only 524 trade disputes have been filed since 1995 at the WTO. Other organizations, such as UNCTAD or ICSID have had over 1,280 disputes, representing a 240% greater caseload than the WTO. Why does the WTO have fewer disputes? International trade literature finds the WTO is incredibly successful at preventing disputes through its formal consultation mechanism (Busch and Reinhardt 2000). However, little attention has been paid to the specific trade concern (STC) mechanism, which may be even more influential in preventing WTO disputes (Horn, Mavroidis, and Wijkstrom 2013; Wolfe 2013). To the author’s knowledge, no paper has attempted to answer why some STCs resolve disputes and why others fail. Thus, this paper seeks to explain the effectiveness of the STC mechanism at the World Trade Organization. This paper tests three emerging theories regarding the resolution of specific trade concerns within the Sanitary and Phytosanitary Measures (SPS) committee. First, countries with strong economic interdependence are expected to resolve disputes to prevent spillover into the broader economic relationship. Second, ideologically-aligned countries may resolve disputes as political favors or perhaps richer countries provide technical assistance to poorer countries in exchange for being allowed to violate their international trade obligations. The last theory argues that STC-specific characteristics, such as its trade value, primary keyword, or sentiment, drive dispute resolution. Ultimately, this paper finds the most support for the second theory that political factors, whether ideological similarity or economic status, drive STC resolution. It also finds that narrower STCs, measured by fewer keywords, are more likely to be resolved. Thus, while the STC mechanism does not resolve fundamental disagreements over trade, it does ensure that free trade remains open. In other words, STCs are successful resolving specific, compact concerns, which is exactly what the STC mechanism was designed to do.
Do Language and Cultural Barriers Enable Firms to Price Discriminate Between Countries in the European Union?

TIM MAECKER

University of Warwick

Supervisor: Dr. Amira Elasra

ABSTRACT

This paper analyses price discrimination that emanates from language and cultural barriers in the Eurozone, that is, companies’ ability to segregate markets for flight tickets. The analysis exploits a novel set of price data on over 10,000 unique flights across 8 different countries in the Eurozone, which was collected repeatedly over a period of two weeks in January 2017. As a single-market with a common regulatory framework and currency, but different languages and cultures, the Eurozone provides an ideal setting to test for price discrimination. Combined with these features of the Eurozone, analysing prices of identical virtual goods allows to abstract from distorting effects such as trade costs, nominal exchange rate pass-through, and product heterogeneity. Beyond these effects the model controls for distance and income effects. This allows to test whether income and distance effects found by other studies in the international trade literature still persists when accounting for language and cultural effects. The analysis provides evidence that language barriers significantly increase the ability of firms to price discriminate across countries. By empirically testing language and cultural mechanisms, this paper contributes to the existing literature on firm level international price discrimination.

1 The author is very grateful to Dr. Amira Elasra, Dr. Nathalie Chen, Dr. Andrew Harkins and Prof. Michael Waterson for excellent guidance and encouragement.
1. Introduction & Literature Review

1.1 Introduction
Whilst it seemed that the widespread adoption of the internet would mark the end of price differences across regional markets, an abundance of outcries in the media about online price discrimination suggests otherwise. Therefore, in November 2016 the EU trade ministers decided to take action against the practice of “geoblocking”, in which customers based in one country are deterred from purchasing articles on the regional website from another country. However, curiously little attention has been given to the question to what extent firms are able to price discriminate between regional EU markets in the absence of geoblocking and otherwise costless arbitrage, for example due to language barriers. Therefore, this paper explores empirically the factors that impact dynamic price discrimination of virtual and homogenous goods between countries in markets without geoblocking. In particular, the segmenting effect of language and cultural barriers in the EU flight market are at the centre of this analysis. As a single-market with a common regulatory framework and currency, but different languages and cultures, the Eurozone provides an ideal setting to test for price discrimination and allows for distorting effects commonly found in the literature, such as shipping cost, product heterogeneities or exchange rate pass through (Cabolis et al., 2007). Therefore, the chosen setting allows isolating the effect of language and culture to explicitly answer the research question - Do language and cultural barriers allow firms to price discriminate between countries in the European Union? In order to investigate this hypothesis, this paper exploits a novel set of price data on over 10,000 identical flights across 8 different countries in the Eurozone, which was collected repeatedly over a period of two weeks in January 2017.

In line with the research hypothesis, the analysis provides strong evidence that language barriers are associated with higher price discrimination. Notably, sharing the same website language is associated with a 1% reduction of relative price differences between two countries on expedia and 0.6% on skyscanner. Language therefore explains 27% and 32% of the variation compared to the average level of price dispersion, respectively. Following the interpretation of Engel and Rogers (1996), the variation induced by sharing the same language is equivalent to a 1694km change in distance between the two countries. Importantly, the above-mentioned results hold whilst controlling for country specific unobserved heterogeneities. Thus, language effects are not simply driven by the fact that countries that share the same language are more similar. The results are robust to controlling for strategic ordering of firms and only including commonly bought flights. In line with the literature, the model finds that there is no economically important effect of distance and income on price differences of virtual homogenous goods (Blum and Goldfarb, 2006; Rauch, 1996).

This paper contributes to the literature in two ways; it analysis a large novel set of price data, and it combines methods from two separate streams in the literature. In the first stream, the literature defines and empirically tests firms’ willingness and ability to price discriminate between countries on a firm level. However, there has not been sufficient consideration of language and cultural effects for virtual goods (Verboven, 1996; Cabolis et al., 2007). Conversely, in the second stream, the international trade literature on virtual goods, these effects have rather been explored on the national level in the form of aggregate price differences or trade flows (Egger and Lassmann, 2012; Melitz and Toubal, 2014). Accordingly, this paper bridges the gap between these two branches in the literature by empirically establishing the aforementioned language effects for virtual goods on firm level price discrimination. It does so by using the theoretical framework from the literature on firm level price discrimination and the empirical framework from the international trade literature.

1.2 Literature review
To elaborate on this literature gap, in the following section both mentioned branches of the literature are discussed in detail. First, the analytical framework to establish price discrimination is discussed, second the literature that aims to empirically establish price discrimination is evaluated and last the literature relevant to analysing language effects on homogenous and virtual goods is investigated.

---

1 To provide an example: www.theguardian.com/money/2011/jan/23/booking-flight-airline-websites
In the literature, price discrimination is commonly defined as price differences of similar goods that are not justified by differences in costs (Stigler, 1987). This paper will analyse the ability of firms to group consumers by country (third-degree price discrimination) as described by Stole (2007). In the theoretical price discrimination literature, the segregation in regional markets has mainly materialised through the spatial model (Holahan, 1975). In the context of the spatial model, language differences could be interpreted as the transport cost for consumers that provide local monopoly power to firms and thus allows them to price discriminate. Consequently, the spatial model provides a theoretical foundation for why differences in languages could enable firms to price discriminate.

The studies of price discrimination that are most related to this paper focus on the analysis of virtual goods to avoid the bias of different production and shipping costs - flight tickets, (Borenstein and Rose, 1994; Stavins, 2001; Giaume and Guillou, 2004). In her widely reknown analysis, Stavins (2001) finds evidence for price discrimination whereby airlines use ticket restrictions to segment the markets. Nevertheless, the analysis does not directly control for demand heterogeneities arising from differences between the airlines. This shortcoming motivated the paper to analyse price differences for the same good across markets.

A similar approach has been taken by Mikians et al. (2012) in their wide-ranging empirical analysis of online (first-degree) price discrimination. In their publication, the authors find online price discrimination for identical products based on consumers’ IP address and hardware. However, whilst also testing for online price discrimination across countries, unlike Mikians et al. (2012) this paper tests for price discrimination in which the “consumer addressability” (Chen and Iyer, 2002) is only indirect as it is based on language and culture.

In the second literature stream, the international trade literature, the possibility of price deviations for virtual goods has also been analysed, albeit on a national level. Whilst the widespread introduction of the internet was first deemed to eradicate all unexplainable price differentials (Cairncross, 2001), the literature has acknowledged that price differentials still persist (Hortasu et al., 2009). Several scholars in the trade literature argue that information friction generally account for price differentials (Head and Mayer, 2014). On the one hand, Rauch (1996) argues that price dispersions are driven by an increase in search costs over distance. On the other hand, Blum and Goldfarb (2006) use consumer search data to find that changes of preferences over distance are the key driver of price differentials. Notably, both authors argue that there exists no distance effect for homogenous virtual goods. Consequently, this paper controls for distance and income to contrast with and contribute to the existing literature.

Motivated by the conclusion that distance is not the main driver of price deviations, this paper aims to establish the impact of language and cultural effects on price deviations and thus price discrimination. In the international trade literature, this effect has been established on the national level but not on the firm level. Egger and Lassmann (2012) identify an increase of 44% on trade flows across the literature if countries share an official language. Melitz (2008) and Melitz and Toubal (2014) find that language and cultural similarity are two distinct effects, also for homogenous goods for which information frictions might be less relevant. Besides naming cultural proximity as a necessary control (Felbermayr and Toubal, 2010), several authors have also established trust as a mechanism that could drive price differences (Guiso et al., 2009). This paper therefore includes trust and cultural proximity as additional mechanisms in order to empirically test their interaction with language effects.

2. Data
For the analysis of this paper, prices were collected from 8 (Eurozone) country-specific sub-pages of Expedia and Skyscanner, e.g. expedia.fr and expedia.de. An example of the websites and the analysed price differences is provided in Figure 1. Prices of flights departing on 4 different days were collected repeatedly over a period of 17 days in January and February 2017. These flights have been matched over time and relative prices of 28 country pairs have been calculated. The final dataset includes more than 1.5 million price pairs from over 10,000 flights.

2.1 Data collection
To create a comprehensive and representative dataset, several strategies were applied. These strategies are based on previous
literature on the pricing dynamics and price discrimination of flight tickets. This allows to contrast the results of this paper with the findings of the recent literature (Salanti et al., 2012; Moreno-Izquierdo et al., 2015). The different selection criteria are summarised in Figure 2 in Appendix A. The main variables of interest for this analysis capture language and cultural similarities. The data of the proportions and level of foreign languages spoken in each country is taken from official EU-reports such as Special Eurobarometer 386, “Europeans and their Languages”. Additionally, national level statistics about the countries in the sample (e.g. real GDP per capita and popularity of flights in the EU) are taken from Eurostat and considered constant throughout the sample. The main variables are constructed closely following the related literature (Melitz and Toubal, 2014) for measures of language and culture and (Guiso et al., 2009) for a measure of trust. The construction of the main variables is summarised in figure 8 (Appendix C) and further discussed in section 3.

![Figure 1. Example of two Expedia search queues](image)

The prices of the flights on the regional pages (e.g. DE-FR) were uniquely identified and matched into 28 pairs. The dependent variable was created to be the absolute of the natural logarithm of the relative price of the pair. This allows measuring the magnitude rather than the direction of price discrimination and reduces the impact of extreme outliers in the sample. Using the 547,011 prices taken from the 8 regional pages, the main sample of 1.52 million price pairs was constructed.

To provide an example of the construction of the dependent variable, the relative price deviation for flight i at time t on the German and the French website is calculated as:

$$\text{Abs}[\ln(PR_{i,t,DE&FR})] = \text{Abs}[\ln((PR_{i,t,DE}) / (PR_{i,t,FR}))]$$

(1)

2.2 Data Analaysis

The unconditional average of absolute relative price deviation is 2.5% on Expedia and 1.7% on Skyscanner. In both cases, the averages are significantly different from zero at the 1% level. The standard deviations are 5.4% and 3.1%, respectively. Together with a median of 0.6% on both websites, these findings suggest fat tails or in other words the existence of extreme cases of price deviations in both samples. To avoid results driven by extreme outliers, only relative prices lower than the 99th percentile were considered in the main model.

In line with the research hypothesis, if the websites share the same language, the average price deviation falls from 2.6% to 1.6% on Expedia and 1.7% to 1.5% Skyscanner. Both decreases are significant at the 1% level. Average price deviations also vary depending on which country is included in the pair. This suggests the inclusion of country fixed effects in the main model.

---

4 Eurostat Database nama_10_aux and avia_par, Data from 2015
analysis. On Expedia, pairs that include the Austrian regional page are associated with the lowest average price deviation of 2.07%, whereas pairs that include Belgium are associated with higher price deviations of 3.28%. This is driven by the fact that the prices on the Belgium regional page are lowest and prices on the Austrian regional page are close to the average of the prices on the other websites. On Skyscanner the highest and lowest average price deviation is between pairs including Germany and France respectively.

The average price deviation also varies significantly between departure dates. On Expedia, the average price deviation is the highest for flights closest to departure (2 weeks at the start of the sampling period) and falls the later the flight is. On the contrary, price deviations on Skyscanner are higher for flights that are further in the future. Interestingly, the time trends for the flights are distinctively different depending on their departure dates. On Expedia the search date for the short and medium term flights follows an inverted U-shape relation. Conversely, long term flights on Expedia and flights on all dates on Skyscanner have a constant positive time trend. The finding of distinct long run and short run trends is supported by the related literature (Salanti et al., 2012) and motivates the inclusion of time fixed effects rather than the estimation of a functional trend in the main model.

3. Empirical Strategy

The first part of the empirical strategy of this model is to focus the analysis on virtual goods sold on regional websites of Eurozone countries. As a single-market with a common regulatory framework and currency, but different languages and cultures, the Eurozone provides an ideal environment to test for price discrimination. Combined with these features of the Eurozone, analysing prices of identical virtual goods supplied by one firm allows to abstract from distorting effects such as shipping costs, nominal exchange rate pass-through, and product heterogeneity. Hence, this analysis can therefore abstract from these factors that are found to distort findings of price discrimination in the literature (Shepard, 1991; Cabolis et al., 2007). Notably, the ideal setting permits to conclude price discrimination when price differences are found. This allows to apply the empirical framework of the international trade literature to test for price discrimination.

3.1 Preliminary Analysis: Market & Flight Characteristics

A preliminary model is applied to explore the impact of market and flight characteristics on price discrimination that have been explored in the related industrial organisation literature (Verboven, 1996). The findings of this preliminary analysis then motivate the choice of fixed effects dummy variables in the main model. The preliminary analysis also allows to contrast the results of the data with the literature, e.g. the impact of concentration (Stavins, 2001), and in this way, provides a possibility to test for external validity of the analysis.

For this preliminary analysis, a multidimensional fixed effects model is applied with relative price deviations as dependent variable, and flight and market characteristics as independent variables. Whilst this analysis estimates flight and market characteristics (e.g. concentration) and absorbs any variables that are constant between pairs (e.g. language effects), the main model estimates pair-based variables and absorbs flight and market characteristics. The exact specification is outlined in Appendix B.

3.2 Main Analysis: Language & Cultural Effects

The main model focuses on estimating the impact of language and cultural effects on price deviations between country-pairs. The methodology is inspired by Shepard's (1991) analysis of price discrimination in the gasoline market and Engel and Rogers' (1996) city-pair analysis on the effect of borders on trade. In particular, a panel data model is applied, where one dimension represents different flights, one dimension the 28 country pair combinations and one dimension the day of search. The main model aims to explain $\text{Abs}[\ln(RP_{ijt})]$, the absolute of the natural logarithm of relative prices of flight i between pair j at time t, whilst grouping the observations by flight, country and time. The dependent variable is measured in logs as this mitigates the impact of extreme outliers in the sample. Lastly, the model measures the absolute of relative price deviations as this analysis focuses on the magnitude rather than the direction of price discrimination. The model estimates

$$ FE: \text{Abs} [\ln(RP_{ijt})] = \beta_1 \text{SameLanguage} + \beta_2 \text{Latin}_j + \beta_3 \text{Trust}_j + \beta_4 \text{InternetUse}_j + \theta_1 \ln(\text{Distance}_j) + \theta_2 \ln(\text{GDPr}_j) + \varepsilon_{ijt} $$

(2)
The main mechanism is Same Language, which indicates whether the website of the two countries in pair $j$ have the same language and thus aims to capture the effect of language on price dispersion. It is based on Melitz (2008) and Melitz and Toubals’ (2010) analysis and methodology to capture language and cultural effects. The additional mechanisms are Latin, which is a dummy indicating if countries share the same culture (Latin) but not the same language, and Trust, an index of trust taken from Guiso et al. (2009). These mechanisms are included to test if language effects are mainly driven by cultural similarity and trust. Lastly, the controls are Internet Use, which is the product of the proportion of frequent internet users in each country in $j$, $\ln(Distance)$, the natural logarithm of distance in km between the capitals of the countries in pair $j$ and , the absolute of the natural logarithm of the real GDP per capita ratio between the two countries in pair $j$. Notably, $\ln(Distance)$ allows contrasting this analysis with comparable studies about the effect of distance on price differences of homogenous virtual goods (Rauch, 1996; Blum and Goldfarb, 2006).

The main model applies flight, time and country fixed effects. Time and flight fixed effects absorb, and thus control for the market and flight characteristics that influence price discrimination. The impact of these characteristics has been identified in the preliminary model but also in the literature, e.g. the impact of concentration (Stavins, 2001). The time fixed effects additionally replace the estimation of a specific trend function. The data as well as the literature (Salanti et al., 2012) suggests the existence of distinct long run and short run time trends. Given that the analysis only covers a period of 17 days, the model could not estimate the long run trend and thus would be subject to biased results. Interpreting coefficients within time periods alleviates this bias. Lastly, the analysis employs (non-directional) country fixed effects to capture country fixed unobserved heterogeneity (e.g. country specific demand curves and pricing strategies). The inclusion of these fixed effects will be of particular interest as they capture if lower price discrimination is mainly driven by the fact that countries with the same language might also have similar market and demand characteristics (e.g. Germany and Austria). In general, the country fixed effects capture any country specific heterogeneity which is constant for this country, such as differences in demand, regulation and taxation. In order to account for the fact that the errors of the regression might be correlated within each flight, the model clusters the standard errors at the flight level.

4. Results
4.1 Preliminary Analysis: Market & Flight Characteristics
The preliminary model outlined above analyses the impact of flight and market characteristics on price dispersion. The detailed results can be found in table 3 (Appendix B). Interestingly, the model predicts a negative quadratic relationship between concentration and price dispersion on Expedia. This is partly in line with the literature which finds that more competition is associated with more price discrimination (Stavins, 2001). The spatial model might justify this result if consumers attach preferences to certain (e.g. local) airlines. A higher number of locally differentiated airlines allow for more differentiated prices. Notably, this relationship is inverted, albeit relatively flat, for Skyscanner. This might be explained by the fact that Expedia and Skyscanner have different business models with the former mostly buying on the wholesale market and selling to clients directly, i.e. setting prices, and the latter being a meta search engine, i.e. receiving prices.

Similar to the relationship with concentration, the effect of flights being closer to departure is opposite for both websites. Whilst price discrimination increase when flights are closer to departure on Expedia, they fall on Skyscanner. Again, the relationship is more pronounced on Expedia and rather flat on Skyscanner. These results provide the econometric justification for the inclusion of time and flight fixed effects in the main analysis.

4.2 Main Analysis: Language & Cultural Effects
The main analysis is the estimation of equation (2) using time, flight and country fixed effects. The results can be found in table 1 below. The coefficient on Same Language is highly significant at the 1% level on both websites with values of 1% and 0.6% respectively. Thus, on Expedia sharing the same website language is associated with a decrease of price dispersion of 1% for the same flight on a given day, ceteris paribus. Put differently, language effects explain about 27% of the variation in price dispersion on Expedia and 32% on Skyscanner, compared to the respective average levels of price discrimination. Following an interpretation used by Engel and Rogers (1996), the variation induced by sharing the same website language is equivalent to reducing the distance between the two countries by 1694km. It is worth noting that the coefficient on Same
Language has not fallen significantly after including measures of cultural proximity and trust (Table 4, Appendix C). This provides further evidence that language explains a separate effect from cultural proximity and trust. Lastly, the coefficient for language effects is negative on both websites, despite many other effects having opposite signs across the websites. This provides some evidence of the robustness of the language effect, for example to a change in business model.

The coefficient on Latin is negative but only significant on Expedia but not on Skyscanner. This provides some evidence of the importance of cultural proximity on the ability to price discriminate. In order to make a more substantiated interpretation, the analysis could be conducted using further websites with comparable business models to either website. Similarly, the coefficient on Trust, an index of trust between countries developed by Guiso et al. (2009), is negative but insignificant. The results for culture and trust effects stand in contrast to the literature, which found trust and cultural proximity to be a main determinant for pricing differences (Melitz and Toubal, 2014; Guiso et al., 2009). This might be explained by the fact that this analysis focuses only on developed European countries. Since the culture and level of development is very similar in the sample, the effects found in the literature might not materialise in this analysis.

### Table 1: Main Analysis: Language and Cultural Effects

<table>
<thead>
<tr>
<th></th>
<th>Expedia</th>
<th></th>
<th>Skyscanner</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) Abs[ln(RP)]</td>
<td>(2) Abs[ln(RP)]</td>
<td>(3) Abs[ln(RP)]</td>
<td>(4) Abs[ln(RP)]</td>
</tr>
<tr>
<td>Same Language</td>
<td>-0.0105***</td>
<td>-0.00999***</td>
<td>-0.00548***</td>
<td>-0.00556***</td>
</tr>
<tr>
<td></td>
<td>(0.000385)</td>
<td>(0.000420)</td>
<td>(0.000372)</td>
<td>(0.000424)</td>
</tr>
<tr>
<td>ln(Distance)</td>
<td>0.00104***</td>
<td>0.000701*</td>
<td>-0.000769***</td>
<td>-0.000416</td>
</tr>
<tr>
<td></td>
<td>(0.000282)</td>
<td>(0.000371)</td>
<td>(0.000207)</td>
<td>(0.000256)</td>
</tr>
<tr>
<td>Abs[ln(GDPPr)]</td>
<td>0.0121***</td>
<td>0.00454*</td>
<td>0.00434***</td>
<td>0.00487***</td>
</tr>
<tr>
<td></td>
<td>(0.00139)</td>
<td>(0.00257)</td>
<td>(0.000472)</td>
<td>(0.000910)</td>
</tr>
<tr>
<td>Latin</td>
<td>-0.00361***</td>
<td></td>
<td>-0.000295</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000841)</td>
<td></td>
<td>(0.000386)</td>
<td></td>
</tr>
<tr>
<td>Trust Index</td>
<td>-0.000177</td>
<td></td>
<td>0.000422**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000259)</td>
<td></td>
<td>(0.000170)</td>
<td></td>
</tr>
<tr>
<td>Internet use</td>
<td>-0.0331***</td>
<td></td>
<td>0.00595</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00938)</td>
<td></td>
<td>(0.00688)</td>
<td></td>
</tr>
</tbody>
</table>

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

The coefficients on the three control variables are in line with the literature but either statistically insignificant or economically unimportant. For example, the coefficient on distance on Expedia is small and a one standard deviation increase in distance would only result in a 0.07% fall in relative prices. This is in line with the literature analysing the effect of distance on virtual homogenous goods (Rauch, 1996; Blum and Goldfarb, 2006; Hortasu et al., 2009). These results hold whilst controlling for country constant unobserved heterogeneity. Estimating a simplified form of equation (2), with and without country fixed effects, the coefficient on Same Language remains significant at the 1% level, albeit falling 30% (Table 4, Appendix C). This demonstrates that the results hold when controlling for possible similarities.
in the pricing strategy for countries with the same website language, e.g. for operational reasons. To provide an example of the interpretation of the country fixed effects, Belgium is associated with a 1.28% higher price dispersion, compared to Germany.

4.3 Robustness Analysis
In order to further verify the validity of the results above, a series of robustness checks have been carried out. The results for the two main robustness checks are reported for each website in table 2. Firstly, in regression (1) and (3) only the first 10 search results and only the first page of each search query have been used. This aims to make sure that price dispersion is not only a phenomenon of flights which are expensive and unlikely to be bought (since flights are ordered by price). In particular, websites might mark up more on these flights as they are more likely to be bought by price insensitive consumers. However, table 2 depicts that the language effect is robust to focusing on the cheaper and most commonly bought flights. Secondly, in regression (2) and (4) the prices of flights have only been compared when their position on the different regional websites has been the same over time. The relevance of this robustness test emanates from the fact that the websites might use different orderings to “nudge” consumers towards buying certain flights, e.g. by changing prices so that different flights are positioned on the top for different regional websites. Therefore, in order to conclude intended price discrimination, this robustness test controls for the intent to sell using position as a proxy. The effect remains significant on Expedia but not Skyscanner. This might be explained by the large drop in observations on Skyscanner for this robustness test.

Further robustness tests have been carried out and are reported in Appendix C. In particular, the results remain robust to using direct flights and to applying a restriction on flight duration. Moreover, as common in the literature, standard errors have been clustered at the pair level as opposed to the flight level. In short, the results are weakened, due to the relative small number of pairs, however they remain significant. Lastly, a methodological robustness test has been carried out, closely replicating the model of Melitz and Toubal (2014) using the collected data. Overall, the alternative methodology, albeit not applying perfectly, confirms the results of this paper. Further details are included in Appendix C.
5. Limitations & Conclusion

5.1 Limitations
Primarily, this research focuses on price discrimination in the market for flight tickets and in the Eurozone. Therefore, any conclusions might be limited to this specific setting, and therefore not generalisable. Secondly, the dataset does not include data on the demand side of the market, such as tickets sold and competitive pressure between the websites. This might result in endogeneity of the language variable, if these effects are not constant within flight, time or country. Future research could control for this by including a proxy variable for demand such as data from Google Trends. Thirdly, this analysis is not able to fully account for the interaction of short-run and long-run price trends inherent in the market of flight tickets. Extending the dataset to include prices for a longer range of dates would facilitate this analysis. It would also allow contrasting with the literature on holiday effects in flight prices (Salanti et al., 2012). Furthermore, an extension of the dataset could include prices taken directly from the airlines to compare the effects across the production chain. Lastly, whilst sharing the same language reduces the cost for consumers to exploit price differences, it does not necessarily mean that they are aware of the opportunity. Particularly, Expedia and Skyscanner “nudge” consumers to the respective country website by redirecting them automatically. This might cause an omitted relevant variable bias in the model resulting in an upward bias in the coefficient on the language variable. Therefore, whilst this analysis can establish that language effects are associated with lower price discrimination, it cannot provide direct evidence for a causal mechanism.

5.2 Conclusion
This paper aims to empirically investigate the ability of firms to price discriminate online in the EU in the absence of geoblocking. It does so by investigating the impact language effects have on price differentials in the flight market in the

<table>
<thead>
<tr>
<th>Table 2: Robustness Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Same Language</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Latin</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Trust Index</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Internet use</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>ln(Distance)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Abs[ln(GDP)]</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Observations 155275  11293  171603  8528
R² 0.255  0.271  0.426  0.444
Cluster Flight  Flight  Flight  Flight
Flight FE Yes    Yes    Yes    Yes
Time FE Yes    Yes    Yes    Yes
Country FE Yes    Yes    Yes    Yes

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

In the literature, clear-cut conclusions are often distorted by factors such as trade costs, nominal exchange rate pass-through, and product heterogeneity. Therefore, this analysis focuses on price discrimination of a homogenous virtual good within the euro zone, in which members share a currency and regulatory framework. More precisely, this paper exploits a novel set of price data on over 10,000 identical flights across 8 different countries in the Eurozone, which was collected repeatedly over a period of two weeks in January 2017.

In line with the research hypothesis, this paper establishes a significant effect of sharing the same language on price discrimination. In particular, sharing the same website language results in a 1% reduction in price discrimination on expedia and 0.6% on skyscanner. Therefore, language effects account for 27% and 32%, respectively. Following the interpretation of Engel and Rogers (1996), this effect has a distance equivalent of 1694km. Notably, these results hold when controlling for country specific unobserved heterogeneity. These results are also robust to accounting for strategic manipulation of the pricing orders across websites and also to limiting the analysis to flights that are most likely to be bought. Contrary to the literature (Melitz and Toubal, 2014), the analysis only finds mixed evidence for the importance of cultural proximity and trust. This might due to the fact that this analysis is limited to developed countries with similar cultures. In line with the literature, it is found that distance is not economically important for virtual homogenous goods (Blum and Goldfarb, 2006).

These results imply that policymakers should consider language barriers when forming policies to further integrate trade between the Eurozone member states. In particular, policies could aim to streamline the use of foreign websites, for example by requiring the existence of an English version of a website.
References
Melitz, J. and Touval, F. (2014). Native language, spoken language, translation and trade. Universit Paris1 Panthon-Sborne (Post-Print and Working Papers) hal-01306641, HAL.
Exploring conflict in Nigeria: Does the rapacity effect dominate across local government areas in Nigeria?

DAVID HENNING

University of Warwick

ABSTRACT

How do sudden windfalls in revenues to local governments affect conflict? One channel suggested by the literature is the rapacity effect. It posits that a rise in contestable income increases the incentives to engage in conflict, by expanding the potential gains from appropriation. We explore whether this effect is dominant across local government areas in Nigeria. To avoid endogeneity issues, we use international oil price spikes as an exogenous estimate for variations in transfers to local governments. This is possible because of the institutional setting of Nigeria, and the federal budget's high dependence on oil revenues. Our preliminary results suggested that the rapacity effect dominates. However, none of the results are consistently robust. We thus conclude that the rapacity effect does not dominate; in fact, we find little evidence of a relationship between changes in income and conflict. Possible explanations are the difficulty of extracting money from local governments and the limited visibility of our income shocks. We also provide suggestive evidence that corruption is the main impeding factor for increasing transfers to local governments.

1 I thank Stephan Kyburz for trusting me with access to the data on budget allocations, which he collected and operationalised. I thank Thiemo Fetzer and Rabah Arezki for giving me access to the CIA data on the distribution of ministers in Nigeria. I also thank Thiemo Fetzer for valuable ideas and suggestions, which greatly helped me develop the estimation strategy for this paper. Furthermore, a big thanks to Stefania Paredes Fuentes for her patience, consistent support and advice throughout the entire process. Finally, I thank James Fenske in particular, and numerous other professors and fellow students more broadly, for comments and suggestions.
Pharmaceutical Royalty Licensing Transaction Impacts on Shareholder Wealth:
An Event Study

ALEXANDRA MALINDI COLYER
Georgetown University

ABSTRACT
The cost of developing a single pharmaceutical drug from discovery to final commercialization has skyrocketed over the past few decades without declines in risk or significant increases in revenues. The industry is thus faced with increased needs for capital funding; however, traditional methods of debt and equity financing remain volatile due to market characteristics. Thus, new financial vehicles have become commonplace in the industry, one of which is royalty stream licensing. As a relatively new financial vehicle, royalties and their associated impact on shareholder wealth has little been studied. This study finds that the announcement of royalty license transactions generates significant positive abnormal returns for the royalty receiver (licensor) but none for the royalty payer (licensee). Such findings indicate that these transactions are positive information signals for the licensor, while not for the licensee presenting further study possibilities regarding information asymmetry and optimism bias.

1 I would like to thank my thesis advisor Robert Cumby for his helpful input on this paper.
1. Introduction
The biopharmaceutical industry has historically been a leader in financial performance generating just over one trillion U.S. dollars in revenue worldwide in 2015; however, the industry is facing significant funding and growth challenges. (Statista) Though generating large revenues, revenue growth has declined from 9.6% in 2002 to just 1% in 2015 with peak sales per asset falling 11.4% year-on-year since 2010. (Deloitte, 2016) Compounding this topline growth slowdown, the industry’s R&D capacity has been stymied. Average costs per approved molecule remain extremely high at estimates of $2.87 billion. (Tufts Center for the Study of Drug Development, 2016) Moreover, R&D returns have continued to decline over the past five years from 10.1% in 2010 to 3.7% in 2016. (Deloitte, 2016) To make matters worse, in a time of late-stage pipeline shortage and patent cliff, capital flow to the industry has also slowed, leading to a dearth commonly referred to as the “Valley of Death.” At present, the biopharmaceutical industry is on the verge of revolutionary breakthroughs in treating diseases and prolonging life, yet rising risks and costs, compounded with insufficient funding, hamper drug development and overall innovation.

These challenges facing the biopharmaceutical industry over the past few decades have led to a need for new financing vehicles for the drug development process. Since many biotech firms are finding it difficult to secure financing in equity and debt markets, they have begun to supplement their funding through royalty-based license. As a new financial vehicle, biopharmaceutical royalty licenses have been little studied and their value generation or destruction would prove to be an interesting topic of inquiry. This paper thus conducts an event study around the transaction of royalty licenses between pharmaceutical companies in order to analyze its effects on shareholder wealth, finding significant positive abnormal returns for the royalty receiver, but none for the royalty payer.

2. Methodology
An event study evaluates whether the cross-sectional distribution of returns at the time of the event deviate from predicted. The focus is on the mean distribution of the abnormal returns. This study aims to evaluate whether royalty transactions, the event, on average, is associated with a change in shareholder wealth for investors on both sides of the transaction.

2.1 Event Definition, Event Window, and Estimation Window
The event, designated as \( t = 0 \), is defined as the publicly available date of a royalty-licensing transaction between a royalty payer (licensee) and receiver (licensor) where both firms are publicly listed.

The event window is the period before and after the event, in which stocks are analyzed for abnormal returns. Following precedent in related biopharmaceutical event studies, the event window is set at \( \pm 1 \) days (\( t = \pm 1 \)) around the event. (Hwang, 2013) A short event window is used because they have been shown to capture the significant effect of the event (Ryngaert and Netter, 1990). Moreover, long event windows reduce the power of the test statistic, leading to false inferences about the significance of the event (Brown and Warner, 1980, 1984), and make it difficult to control for confounding effects. (McWilliams and Segel, 1997) We will consider other event windows of \( \pm 2 \) and \( \pm 5 \) days to check for robustness.

The estimation window is the period over which the parameters are estimated to determine normal returns. The event window data is then compared to these estimates to determine abnormal returns. Since the estimation window constructs normal expected returns absent of the event, it is important that no royalty transaction occurs during that time frame. In line with the Sharma and Lacey (2004) paper, an estimation period of 250 days is used referring to \( t = -251 \) to -2.

The post event window refers to the time period after the event window that is used to evaluate market efficiency and that returns went back to normal. A post event window of 120 days will be used, indicating time \( t = +2 \) to +121.

2.2 Abnormal Returns
To test for the royalty transaction’s effect on stock prices, the abnormal returns must be calculated over the event period. The abnormal return is defined as the actual \textit{ex post} return of the security over the event window minus the expected return of the firm if the event did not take place over the event window.” (MacKinlay, 1997) For each firm \( i \) and event date \( \tau \)

\[
\varepsilon_{it} = R_{it} - E[R_{it} | X_{it}]
\]
where $\varepsilon_{it}$, $R_{it}$, and $E(R_{it})$ are the abnormal, actual, and normal returns, respectively, for time $t$. Additionally $X_t$ is the conditioning information for normal performance model. As to be explained subsequently with model choice, $X_t$ is the market return.

In event studies, there are numerous models to estimate actual returns. In this study the market model will be used as it the most prevalent and has been shown to offer more powerful tests in detecting abnormal returns (Dyckman, 1984). For any security $i$,

$$\begin{align*}
R_{it} &= \alpha_i + \beta_i R_{mt} + \varepsilon_{it} \\
E[\varepsilon_{it}] &= 0 \quad \text{Var}[\varepsilon_{it}] = \sigma^2 \varepsilon_i
\end{align*}$$

where $R_{it}$ is the actual return, $R_{mt}$ is the market portfolio, and $\varepsilon_{it}$ the zero-mean disturbance. The parameters $\alpha_i$ and $\beta_i$ are estimated using OLS regression. Under the normal market model, the abnormal returns become

$$\varepsilon^*_i = R_{it} - (\alpha_i + \beta_i R_{mt})$$

Summing the abnormal returns around the event window from $T_1$ to $T_2$ to yield the cumulative abnormal returns for security $i$.

$$\text{CAR}_i(T_1, T_2) = \sum_{t=T_1}^{T_2} \varepsilon_{it}$$

Since some of the firms differ in size so that abnormal returns are not directly comparable, market capitalization $K$ on the first day of the event window are considered. Consequently, to see the impact on shareholder wealth as a result of the event the cumulative abnormal returns are multiplied with the company’s market capitalization on $T_3$ around the transaction date.

$$\text{Wi}(T_1, T_2) = \text{CAR}_i(T_1, T_2) \times K_{T_3}$$

2.3 Hypotheses

The null hypothesis is that there are no abnormal mean returns. If the null hypothesis proves to be true, the following diagram results (Figure 6).

Based off the findings of Lev and Gu (2004) that royalty income has a significant positive effect on stock prices, the following hypothesis is to be tested in relation to the null. This study hypothesizes that **royalty transactions in the biopharmaceutical sector generate significant, positive abnormal mean returns**. Two additional hypotheses will be tested. 1) There is a significant difference between the abnormal returns generated from royalty transactions on drugs in later development stages than earlier ones, likely favoring later stages as they represent a more guaranteed source of cash flow in the near future. And 2) There is a significant difference between the abnormal returns generated from royalty transactions on drugs in “hot” therapeutic areas like oncological and neurological as these areas have higher demands in the market place.
If the hypothesis that the event contains valuable information content is perfectly efficient, the cumulative abnormal return change will look like Figure 7. Since we are presuming a semi-strong efficient market hypothesis, Figure 7 likely will not occur, rather it will result in price return effects related to Figure 8 with potential over or under-reactions as provided by the University of Houston (2014). Furthermore, a graph of abnormal returns should witness abnormal returns averaging around zero and then a sharp spike on the event date that equilibrates back down to zero immediately afterwards. Thus, if our hypothesis proves true we expect to find significant abnormal returns and cumulative abnormal returns different than zero on the event date t=0.

3 Statistical Tests
3.1 Parametric Tests
The Central Limit Theorem guarantees that the abnormal return mean’s distribution converges to normality as long as the cross sectional abnormal returns of firms are iid. The first test will be a traditional t-test.

\[ t_{\text{CAR}} = \frac{\text{CAR}_{it} \times \sqrt{n}}{\sigma \times \text{CAR}_{it}} \]

This will be followed by a cross-sectional dependence test for robustness.

\[ t_{\text{cross}} = \frac{\text{CAR}_{it}(T_1,T_2)}{\sigma_{\text{CAR}(T_1,T_2)}} \]

3.2 Non-Parametric Tests
Non-parametric tests serve two purposes. One, they solve the problems that might occur if the data sample size is not large enough for the Central Limit Theorem to apply. Two, they provide a robustness check on the parametric tests' conclusions as they provide robustness against non-normally distributed data. The two non-parametric tests to be evaluated are the Wilcoxon Signed-Rank Test and the generalized sign test.

The Wilcoxon signed-rank test assesses both the sign and magnitude (through ranking based on the absolute value of returns) of abnormal returns\(^1\), assuming that none of the absolute values are equal and that each is a nonzero value. The test statistic (\(H_0: AAR=0\)) is given by the following equation where \(i\) is the positive rank of the absolute value of abnormal returns at the time point \(t\) for firm \(i\).

\[ W = \sum_{i=1}^{n} R_{i} \]

\[ R_{i} = \text{sign}(AAR_i) \times R_{i} \]

\[ W_i = \frac{1}{2} \times \frac{n(n+1)}{2} - \sum_{i=1}^{n} \frac{R_{i}^2}{2} \]

\[ Z = \frac{W - \frac{1}{4}n(n+1)}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \]

\[ p-value = 2 \times \Phi(-|Z|) \]

\(\text{AAR}_{it}\) is the abnormal return for firm \(i\) at time \(t\). \(\text{CAR}_{it}\) is the cumulative abnormal return for firm \(i\) at time \(t\). \(\text{AR}_{it}\) is the average return for firm \(i\) at time \(t\). \(\sigma\) is the standard deviation of the abnormal return for firm \(i\) at time \(t\).
\[ W_i = \sum_{t=1}^{N} \text{rank}(A_{it})^+ \]

\[ z_{\text{Wilcoxon}} = W - N(N - 1)/4 \div \sqrt{(N(N+1)(2N+1)/12)} \]

Under the generalized sign test proposed by Cowan (1992) the ratio of positive cumulative abnormal returns over the event window is a binomial random variable giving the following test statistic

\[ t_{\text{sign}} = \frac{p_0 - p}{\sqrt{p(1 - p)/N}} \]

where \( p_0 \) is the observed fraction of positive returns computed over all stock in the event window and \( p \) is the observed fraction of positive returns computed over all stocks in the estimation period. It verifies the direction of abnormal returns, either positive or negative.

### 3.3 Cross sectional Regression

The cross-sectional regression permits analysis of certain factors as explanations of wealth effects from the event. Thus, a regression of the cumulative abnormal returns for the portfolio of stocks on an indicator variable will be conducted for the two optional hypotheses run with heteroskedastic standard errors. In the first, comparing drug development stage, the indicator variable will be instantiated as 1 for its stage of development (discovery, preclinical, clinical, commercial) and 0 otherwise with commercial represented in the constant. For the second hypothesis on innovative drug classification, the indicator variable will be instantiated as 1 for its therapeutic area and 0 otherwise with the constant representing areas with less than 9% data representation.

### 4. Data Description

The event data consists of 73 royalty transactions between publicly listed pharmaceutical companies\(^2\) from 2000 to 2016. In the data set there are 52 unique royalty receivers, or licensors, and 41 unique royalty payers, or licensees. Announcement data comes from Evaluate Group press releases, LexisNexis investment announcements, SEC 8-Ks, and company financial news reports. Equity returns and stock data are exported from Bloomberg Terminals (Bloomberg L.C.). Important to designate, market returns are taken from the S&P Pharmaceuticals Select Industry Index. This ETF was chosen to represent market returns as it is equally weighted across holdings and, most importantly, is balanced across market cap, which mirrors the royalty transactions data's composition.

### 5. Results

#### 5.1. Royalty Receiver

Focusing on the royalty-receiving firms’ returns around the event window, there is a statistically significant positive abnormal return on the event date across firms and across varying event window lengths. Table 2 summarizes the cumulative abnormal returns across all events on the event date. Overall the cumulative abnormal return is approximately 0.07 with significance at the 99% confidence level. A five and eleven-day window offer similar results to the three-day window, verifying that the results are accurate.\(^3\)

<table>
<thead>
<tr>
<th>Event Window</th>
<th>3 Day (±1)</th>
<th>5 Day (±2)</th>
<th>11 Day (±5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Royalty Receiver</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAR</td>
<td>0.0706</td>
<td>0.0727</td>
<td>0.0674</td>
</tr>
<tr>
<td>t-stat</td>
<td>(4.68)***</td>
<td>(4.15)***</td>
<td>(2.66)***</td>
</tr>
<tr>
<td>AR</td>
<td>.0605</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>(4.71)***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Each column indicates the length of the event window to evaluate cumulative abnormal returns and abnormal returns. *, **, *** signify significance at the 90, 95, and 99% respectively. The best event window is thus chosen based off data from this table. The most ideal event window is the shortest without compromising high levels of statistical significance,

\(^2\) As a note PDL Biopharma is traditionally a pharmaceutical company but has since shifted to be more like an investment firm managing patents, royalties, etc.

\(^3\) The drop in significance for the 11-day window is expected as the majority of the days experience normal returns.
which for the case of the royalty receivers will be a three-day event window.

Now, Table 3 illustrates the average abnormal returns and cumulative abnormal returns relative to the S&P Pharmaceutical Select around the event date.

<table>
<thead>
<tr>
<th>t</th>
<th>AAR</th>
<th>t-stat</th>
<th>CAAR</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>.0022379</td>
<td>-18.2</td>
<td>-.1238855</td>
<td>26.3</td>
</tr>
<tr>
<td>-1</td>
<td>-.000559</td>
<td>-10.1</td>
<td>-.1239414</td>
<td>26.4</td>
</tr>
<tr>
<td>0</td>
<td>.059406</td>
<td>-220</td>
<td>-.0645317</td>
<td>4.56</td>
</tr>
<tr>
<td>1</td>
<td>.0093646</td>
<td>-43.4</td>
<td>-.0551671</td>
<td>1.12</td>
</tr>
<tr>
<td>2</td>
<td>-.0009728</td>
<td>-6.85</td>
<td>-.05614</td>
<td>1.48</td>
</tr>
</tbody>
</table>

Table 3: Each row represents the time distance from the event date. The t statistic is calculated based off the 253 AAR or CAAR from the estimation window and event window. Any absolute t value greater than 2.6 is significant at the 99% level. As a note to avoid confusion, the CAAR's total AAR for all t recorded up to that date. For example, the CAAR recorded at t=-5 is the sum of AAR from t=-251 to t=-5.

Important to note, on the event date the average abnormal returns is .059406, significant at the 99% confidence level. The AAR on day 0 is the largest positive abnormal return relative to those of the surrounding 40 days. Moreover, it is the most highly significant average abnormal return within that time period analyzed. Analysis of the 40 days surrounding the event date reveal that the AAR at t=0 is statistically different and larger than the other daily abnormal returns at the 99% significance level. On top of that the AAR represents a 289% growth from the average AAR (-.0002063) within those surrounding days. Thus, a royalty transaction for the royalty receiver is associated with positive abnormal returns. Now looking at the cumulative average abnormal returns, there is significant 47% increase between t=-1 and the event day t=0. Figures 11 and 12 display these trends for both AAR and CAAR and are the strongest evidence for positive abnormal returns on specifically the event date t=0 due to the royalty transactions for the licensor.

Over time, including the estimation period, event window, and post-event period, it is expected that the average abnormal returns should be approximately 0. However, if the event holds information content, during the event window the average
abnormal returns should be different than zero. The working hypothesis states that the royalty transactions should be a wealth generator and thus it is expected to have positive abnormal returns around the event date. This has been substantiated by Table 2 and Figure 12. Table 4 illustrates a sign test over three periods – the entire time, event window, and the event date – to further analyze these hypotheses.

Table 4: Each column represents the results of a sign test for various time frames. The entire time represents the estimation period, event window, and post event period with an expected positive and negative average abnormal return instance at 186.5 each. The event window consists of solely the event window (±1) for all 73 observations, with an expected positive and negative abnormal return instance at 109.5 each. The event date represents only t=0 for all 73 observations with an expected positive and negative abnormal return instance of 36.5. Since the prior analyses have indicated positive abnormal returns on the event date, alternative hypothesis of the median abnormal return <0 has been left out. The table reveals that the transaction in fact generates statistically significant positive abnormal returns for the royalty receiver.

Table 5: Each column represents the results of a sign test for the different days within the event window in order to capture potential information leakages or delays. The expected positive or negative abnormal return instance is predicted to be 36.5. Since the prior sign test revealed significant positive abnormal returns the alternative hypothesis of median AR<0 has been left out. Overall there does not appear to be significant information leakage or delay around the event.

From Table 5 there is no evidence of information leakage as abnormal returns are not different than zero in the event window prior to the event date. This is corroborated by Figures 11 and 12 as the jumps in AAR and CAAR occur at t=0. There may be slight information delay for royalty receivers as abnormal returns on t=1 are greater than 0 at the 80% significance level.

The Wilcoxon sign-rank test serves as a robustness check for the parametric tests and even for the nonparametric sign test, adjusting for potential non-normality, while offering a test for abnormal returns magnitudes. Overall as shown in Table 6 it finds similar results to the tests as abnormal returns have significant positive abnormal returns at the 99% significance level on the event date.
Table 6: Each column represents the results of a sign test for the different time periods tested. Overall it reveals similar results to the other parametric and non-parametric tests.

Having established that royalty receivers experience relatively efficient positive abnormal returns from royalty license transactions, such shareholder-wealth growth is broken down according to development stage and therapeutic area via cross-sectional regressions in Tables 7 and 8. Analysis of the different stages of drug development serves as a proxy for differentiating between actual wealth generation and future wealth signaling. Theoretically, earlier stages are bets on the future and thus their royalty transactions should serve as an indication or signal of future wealth. Later stages, particularly commercial, represent a definitive cash flow for a current or near time frame and thus should estimate actual wealth generation. Since the earliest stage, discovery, is the most highly significant indicator, the positive abnormal returns represented in this study are best described as a signal for future wealth generation rather than current shareholder wealth generation. Additionally, the relative contribution to CAR decreases the later the stage in pharmaceutical development. Originally, it was expected that the commercial level would produce the largest abnormal returns as it was a guaranteed cash inflow for the receiver in the near term; however, this is evidently not the case. These results likely reflect the higher risk-reward demands for earlier stages; since the early development stages have a high risk of failure, shareholders may demand higher returns to compensate for the increased risk.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>AR=0 (in event window)</th>
<th>AR = 0 (on event date)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Instances</td>
<td>129</td>
<td>51</td>
</tr>
<tr>
<td>Negative Instances</td>
<td>90</td>
<td>22</td>
</tr>
<tr>
<td>Sum Ranks Positive</td>
<td>15164</td>
<td>2135</td>
</tr>
<tr>
<td>Sum Ranks Negative</td>
<td>8926</td>
<td>566</td>
</tr>
<tr>
<td>P Value</td>
<td>0.001</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 7: The table represents the cross-sectional regression of drug development stage against cumulative abnormal returns and abnormal returns on the event date. Each variable represents an indicator for its respective development stage. The constant represents the commercial stage. T-statistics are represented in parentheses below the coefficient values. *, **, *** signify significance at the 90, 95, and 99% respectively.

<table>
<thead>
<tr>
<th>t = 0</th>
<th>CAR</th>
<th>AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant (Commercial)</td>
<td>.00940</td>
<td>.0161</td>
</tr>
<tr>
<td></td>
<td>(0.48)</td>
<td>(0.76)</td>
</tr>
<tr>
<td>Discovery</td>
<td>.0908</td>
<td>.0637</td>
</tr>
<tr>
<td></td>
<td>(2.64)**</td>
<td>(1.91)*</td>
</tr>
<tr>
<td>Preclinical</td>
<td>.0780</td>
<td>.0493</td>
</tr>
<tr>
<td></td>
<td>(1.72)*</td>
<td>(1.35)</td>
</tr>
<tr>
<td>Clinical</td>
<td>.0671</td>
<td>.0523</td>
</tr>
<tr>
<td></td>
<td>(2.14)**</td>
<td>(1.65)</td>
</tr>
<tr>
<td>R-Squared</td>
<td>.0636</td>
<td>.0433</td>
</tr>
<tr>
<td>No. Observations</td>
<td>73</td>
<td>73</td>
</tr>
</tbody>
</table>
Table 8: The table represents the cross-sectional regression of therapeutic area against CAR and ARs on the event date. Each variable represents an indicator for its respective therapeutic area. The constant represents the therapeutic category of other, which includes areas with less than 9% of data representation. T-statistics are represented in parentheses below the coefficient values. *, **, *** signify significance at the 90, 95, and 99% respectively.

With regards to drug therapeutic area, the three most numerous areas - oncology (21), technology (10), neurology (6) - were chosen as separate dummy regressors while the remaining areas were designated for other to be represented in the constant. The cross-sectional results are shown in Table 8. Interestingly, the constant appears to significantly positively impact AR and CAR. Oncological drugs do increase abnormal returns only; however, this is only significant at the 90% level. Importantly, though, oncological drugs are associated with the largest abnormal returns at approximately .08.

5.2 Royalty Payer
The results for royalty payers follow the same order and format as that of the royalty buyers. The previous results commentary also similarly applies.

Table 9: Each column indicates the length of the event window to evaluate cumulative abnormal returns and abnormal returns. *, **, *** signify significance at the 90, 95, and 99% respectively. The best event window is thus chosen based off data from this table. The most ideal event window is the shortest without compromising high levels of statistical significance, which for the case of the royalty receivers will be a three-day event window.

Unlike the royalty receivers, the royalty payer is not associated with abnormal returns around the event date as seen by the lack of significance above the 90% level on any event window and abnormal returns extremely close to 0.4

---

4 The results for the royalty payer are only significant at the 80% level.
Figure 16: Cumulative average abnormal returns (CAAR) are expected to be stable around a certain value prior to the event window, then rapidly increase or decrease during the event window, and finally stay around the maximum/minimum for the remainder of the time. This does not occur for the royalty payers, indicating potential information leakages/delays or overreactions. Overall CAAR experience a 15% increase from t=-1 to t=0.

Table 10: Each row represents the time distance from the event date. The t statistic is calculated based off the 253 AAR or CAAR from the estimation window and event window. Any absolute t value greater than 2.6 is significant at the 99% level. As a note to avoid confusion, the CAAR's total AAR for all t recorded up to that date. For example, the CAAR recorded at t=-5 is the sum of AAR from t=-251 to t=-5.

From Table 10 important to note, on the event date the average abnormal returns is .00733. The AAR on day 0 is the largest positive abnormal return relative to those of the surrounding 40 days. Moreover, it is the most highly significant average abnormal return within that time period analyzed. Analysis of the 40 days surrounding the event date reveal that the AAR represents a 44% growth from the average AAR (-.0001712) within those surrounding days. Now looking at the cumulative average abnormal returns, there is significant 15% increase between t=-1 and the event day t=0. Figures 15 and 16 display these trends for both AAR and CAAR.
Interestingly, Figures 15 and 16 are more volatile than expected, especially in comparison to royalty receivers, indicating that the royalty license transaction is not as important in causing increased returns. Under normal circumstances such volatility would complicate results interpretation even given the lack of significance. It is important to note that the peak abnormal returns registered on the event date are only one-tenth that experienced by the licensors. Abnormal returns of .00766 under the royalty receiver represent fluctuations under normal market conditions outside the event window. Thus, due to the increased volatility, abnormal returns one-tenth that of the royalty receivers, and insignificance at the 85% level, royalty license transactions do not likely affect royalty payers in a meaningful way for shareholders.5

<table>
<thead>
<tr>
<th>Time</th>
<th>Entire Time (AAR)</th>
<th>Event Window (AR)</th>
<th>Event Date (AR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Instances</td>
<td>180</td>
<td>112</td>
<td>40</td>
</tr>
<tr>
<td>Negative Instances</td>
<td>193</td>
<td>107</td>
<td>33</td>
</tr>
<tr>
<td>P value: Ha median &gt;0</td>
<td>0.7645</td>
<td>0.3935</td>
<td>0.2414</td>
</tr>
<tr>
<td>P value: Ha median ≠0</td>
<td>0.5366</td>
<td>0.7870</td>
<td>0.4828</td>
</tr>
</tbody>
</table>

Table 11: The entire time has an expected positive and negative average abnormal return instance at 186.5 each. The event window has and expected positive and negative abnormal return instance at 109.5 each while the event date's is 36.5. Since the prior analyses have indicated potential, but likely insignificant positive abnormal returns on the event date, alternative hypothesis of the median abnormal return <0 has been left out. The table reveals that the transaction in fact does not generate statistically significant abnormal returns. It is, though, more likely to find positive abnormal returns than negative abnormal returns on the event date. (significant at the 75% level)

Now Table 11 verifies that over the entire time period AAR equilibrate to 0; however, sign tests on the event window and event date itself reveal that abnormal returns around the event are most likely 0. In fact, only at the 75% significance level can we claim that abnormal returns on t=0 are positive. This analysis further corroborates the prior results that royalty license transactions do not generate significant positive returns for shareholders.

<table>
<thead>
<tr>
<th>Time</th>
<th>t = -1</th>
<th>t = 0</th>
<th>t = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Instances</td>
<td>41</td>
<td>40</td>
<td>31</td>
</tr>
<tr>
<td>Negative Instances</td>
<td>32</td>
<td>33</td>
<td>42</td>
</tr>
<tr>
<td>P value: Ha median AR &gt;0</td>
<td>.1746</td>
<td>.2414</td>
<td>.9201</td>
</tr>
<tr>
<td>P value: Ha median AR ≠0</td>
<td>.3492</td>
<td>.4828</td>
<td>.2416</td>
</tr>
</tbody>
</table>

Table 12 The expected positive or negative abnormal return instance is predicted to be 36.5. Overall there does not appear to be statistically significant ARs in any direction. When taken in comparison with prior results, information leakage is not likely the case as fluctuations in abnormal returns outside the event date are similar in magnitude to that of the payers’ AR.

Given that Table 11 shows that it is more likely to find a positive abnormal return within the event window than a negative instance, Table 12 illustrates a sign test focusing on such returns to identify potential information leakages and/or delays. Seeing that at the 80% significance level abnormal returns on t=-1 are greater than 0, there may be slight information leakage prior to the actual event within royalty payers. There, however, does not appear to be significant information delays.

5 Solely looking at Figures 15 and 16, there is a decrease in CAAR after the event window, which would indicate shareholder overreaction to the event; however, the CAAR returns to its maximum around t=12 and then quickly plummets again. A similar occurrence happens with regards to AAR as there are secondary maximums around t=±12. These, however, must be taken with a grain of salt as abnormal returns are only significant at the 80% level and register at magnitudes equivalent to that of normal return volatility for the royalty receiver.
The Wilcoxon-rank test for the licensees finds that on average the licensee experiences slightly positive abnormal returns around the transaction; however, these are not statistically significant and thus we cannot reject the hypothesis that abnormal returns for the licensee as a result of the royalty license transaction announcement are different than zero as shown in Table 13.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>AR=0 (in event window)</th>
<th>AR=0 (on event date)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Instances</td>
<td>112</td>
<td>40</td>
</tr>
<tr>
<td>Negative Instances</td>
<td>107</td>
<td>33</td>
</tr>
<tr>
<td>Sum Ranks Positive</td>
<td>12583</td>
<td>1488</td>
</tr>
<tr>
<td>Sum Ranks Negative</td>
<td>11507</td>
<td>1213</td>
</tr>
<tr>
<td>P Value</td>
<td>0.567</td>
<td>0.450</td>
</tr>
</tbody>
</table>

Table 13: The Wilcoxon sign-rank test for the licensee reveals similar results to the overall results. Even though the data seems to indicate that there is a slightly higher chance of finding positive abnormal returns, these are statistically insignificant and thus cannot reject the null hypothesis that the abnormal returns are different than zero on the event date.

Now turning our attention to the drug development stage, there do not appear to be any remarkable trends. The preclinical stage appears to be the most significant determinant of abnormal returns; however, it negatively impacts them. Outside of this, the commercial stage is the most important positive contributor to abnormal returns as it has the highest positive coefficients and largest t-values relative to the other positive contribution stages. On the event date, commercial drug royalty licenses generate .0177 in abnormal returns at the 90% significance level. This indicates that for royalty payers, the event likely acts more as a wealth generator than a signal, contrary to findings with the royalty receivers. This opposing trend may seem to invalidate our findings for the receivers; however, given that the abnormal returns for the payers are insignificant and one-tenth the size of that for the licensors, this trend of commercial's influence is not that strong of an argument. These findings do not debunk that for the receivers since we would expect similar finding, but rather encourage further study in the case of drug development impact on payers’ returns.

<table>
<thead>
<tr>
<th>t = 0</th>
<th>CAR</th>
<th>AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant (Commercial)</td>
<td>0.01359  (0.73)</td>
<td>0.0177  (1.91)*</td>
</tr>
<tr>
<td>Discovery</td>
<td>-.00212   (-0.10)</td>
<td>0.00779 (0.34)</td>
</tr>
<tr>
<td>Preclinical</td>
<td>-.02271   (-0.93)</td>
<td>-.0237  (-2.17)**</td>
</tr>
<tr>
<td>Clinical</td>
<td>-.01022   (-0.52)</td>
<td>-.0186  (-1.88)*</td>
</tr>
<tr>
<td>R-Squared</td>
<td>.0253</td>
<td>.0790</td>
</tr>
<tr>
<td>No. Observations</td>
<td>73</td>
<td>73</td>
</tr>
</tbody>
</table>

Table 14: The table represents the cross-sectional regression of drug development stage against cumulative abnormal returns and abnormal returns on the event date. The constant represents the commercial stage.
Table 15: The table represents the cross-sectional regression of therapeutic area against cumulative abnormal returns and abnormal returns on the event date.

Now abnormal returns are not significantly impacted by any therapeutic areas for royalty payers as demonstrated in Table 15. This is likely because abnormal returns for the payers are very close to zero and insignificant at the 90% level in the first place along with presence of outliers.

6. Discussion
As this study has shown, shareholders view royalty licensing transactions as a positive indicator for success. Only one party, the licensors, truly receives significant positive returns. Such returns, interestingly, serve as a signal for the royalty receiver’s future success and wealth generation. Rather than actual wealth generation as products in the discovery phase incur the largest abnormal returns for the royalty receiver. This is particularly interesting because the drug’s initial creator, generally the royalty payer, on average does not receive abnormal returns different than zero. Thus, there may be an interesting case of information asymmetry or optimism bias that needs to be explored in the future as perhaps the royalty receivers’ shareholders are too optimistic on the chances for success of the drugs or do not have full information on the licensee firms and their products. For future development, it would be enlightening to conduct similar analyses on smaller, more discovery-stage-oriented firms like colleges and private small firms to see if the signaling effect still outweighs actual wealth generation, if returns to investment are worthwhile, if they encourage innovation, and most importantly if they help bridge the “Valley of Death.” In conclusion, this study presents interesting implications for pharmaceutical development financing.

---

6 There may be information asymmetry because the shareholders for the receivers may not be invested in nor fully follow the payers. As a result, they may not have all the information about the payers’ products and pipelines.
References


Financial Fragility and Contagion in Emerging Economies: A Panel Analysis

STEFANI KOSTADINOVA
American University in Bulgaria

ABSTRACT
The Great Recession of 2008-2009 is still a topic of heated debate among academics and has brought renewed scholastic interest to the study of crises. Considering how the crisis began in an advanced economy and quickly spread to the rest of the world, spillovers have been a popular topic of research alongside intrinsic qualities that make an economy more susceptible to a crisis. However, existing literature tends to focus too narrowly on a few indicators, while disregarding others. This paper aims to fill this gap by exploring crises and their determinants through several panel analysis models. It focuses on 25 emerging economies during the period from 2000 to 2015 inclusive. Through Arellano-Bond GMM, fixed and random effect, logit, and probit models, I come to the conclusion that crisis indices are aggravated by weak fundamentals, financial fragility, and close trading relationships with other countries.
1. Introduction

Crises are a subject of intense study by economists, largely because they affect all layers of society. There is ripe literature on the topic of why different types of crises occur and what policies we could undertake to limit them. Due to the interconnectedness of markets and economies in today’s global world, significant contagion occurs between countries during times of crisis, especially from advanced to emerging economies. For example, if we consider the 2008-2009 financial crisis, we can now look back in retrospect and observe that even though it started in the United States’ housing market, it soon spread not only to the entire financial system and eventually to the real economy, but also to many countries worldwide. In particular, emerging economies were especially hard hit by the recession, which lead not only to decreased growth, but also to shaken banking systems and overall financial instability. For that reason, in this paper, I look at the problem of crisis transmission and the indicators that make it more likely that a certain economy will experience a crisis. From a policy perspective, it is important to note what factors are responsible for strong contagion from advanced to emerging economies because if we can pinpoint the problem, we can then design policies to make emerging economies more resilient to outside shocks.

As suggested by academic research on the topic, crises are more likely to hit some emerging countries than others – usually, those which are subject to financial fragility and stress in their domestic economies as shown by Corsetti et al. (1999), Feridun (2009), and Filipozzi and Harkmann (2010). Additionally, trading relationships can also increase contagion as indicated by Glick and Rose (1999), as can financial linkages, as shown by Park and Mercado (2013). However, most of existing literature is focusing too narrowly on one or two channels of crisis transmission and painting an incomplete picture. I would like to add as a novelty how fragility and contagion occurred in emerging economies during the global financial crisis of 2008-2009 by looking at a multitude of financial and macroeconomic fundamental factors and their relationships. Specifically, with the help of a number of indices on fragility, I aim to see what determines which emerging countries will be most severely affected by crises, what factors influence this, and how the transmission happens. The analysis covers twenty-five countries over the period 2000-2015. The countries were selected by data availability.

The findings indicate that the major factors contributing to the likelihood of crisis transmission to emerging economies are current account deficits and high levels of non-performing loans, especially when accompanied by weak reserves and macro fundamentals, along with banking sector fragility, overvaluation of the real exchange rate, stock market weakness and contagion, and close trading relationships with other countries. First, I provide a literature review on the topic in section 2, followed by a description of methodology in section 3, a discussion of the empirical results and analysis in section 4, and a conclusion in section 5. The appendix contains the list of countries, data sources, and references.

2. Literature Review

Corsetti, Pesenti, and Roubini (1999) explore the reasons behind the Asian financial crisis of 1997-98 and whether it was caused by fundamental weaknesses and harmful policies or financial panic and behavioral factors. They conduct a cross-sectional analysis of 24 emerging economies in order to determine which factors are significant for crisis occurrence. For that purpose, they construct a crisis index as well as indices of financial fragility, reserves, external imbalances, and fundamentals. Corsetti et al. show that the crisis index, which measures currency pressure, is influenced by current account deficits (connected with currency appreciation) and the amount of non-performing loans (connected with lending booms). More specifically, these factors worsen the crisis index especially strongly in the presence of low foreign reserves. Reduced availability of the latter makes the economy more vulnerable to speculative attacks. Non-performing loans as a share of GDP, however, negatively affect the economy if high, regardless of fundamentals and reserves availability.

Similarly, Feridun (2009) explores the relationship between exchange market pressure and underlying financial fragility. By using an autoregressive distributed lag (ARDL) bounds testing procedure and Granger causality tests with a vector error-correction model, he finds a long-run relationship between exchange market pressure and banking sector fragility, reserves inadequacy, and real exchange rate overvaluation. In addition, Filipozzi and Harkmann (2010) also look at the exchange rate channel, linking it to financial fundamentals, as well as exploring the money market channel of crisis incidence. They find that government and banking sector dynamics are linked with crisis indicators. Nevertheless, it is worth noting that they find the exchange rate channel to be the much more significant one.
In addition to the role that financial fragility and fundamentals’ performance play in crisis incidence, interconnectedness is also important. For example, as Glick and Rose (1999) discuss, crises are often regional and a country’s vulnerability to a crisis depends not only on domestic macroeconomic factors, but also on trading relationships. In addition to trade, financial linkages can play a role in contagion as well. By using a financial stress index and variance decomposition to measure the impact of external shocks on domestic markets, Park and Mercado (2013) find that not only do financial shocks spread from advanced to emerging economies, but also between emerging economies themselves. They also show that whether the shocks come from the same region as the countries they affect does not matter, pointing to the interconnectedness of economies beyond geographical proximity. This is less true for emerging Asia, where regional shocks play a bigger significance.

Similarly, Kilic and Ulusoy (2015) discuss contagion in the context of financial system linkages as well as FDI and trade. They use foreign exchange and stock markets and the transmission mechanism of endogenous excessive volatility to measure the latter’s significance for contagion. They find that such volatility is a cause for contagion across economies; this effect can be somewhat reduced by diversification of trade exposure. However, in some cases diversification does not help much either, as large economies’ influence and their price-setter role can be too strong and greatly reduce the effect of diversification.

Even though as mentioned above emerging economies can also spread contagion between one another, the effects of transmission from advanced to emerging economies are very significant and more widely discussed. For example, Balakrishnan et al. (2009) also discuss contagion from advanced to emerging economies. By constructing financial stress indices for both types of economies, they show that the depth of financial linkages between economies is a major determinant of contagion in times of crisis. A major factor that influences this transmission are higher foreign liabilities, as bank lending linkages especially from Western European countries have been a major channel of transmission. Current account and fiscal balance do not seem to help much with financial crisis contagion, but the authors note that they provide an important buffer from damage to the real economy.

3. Methodology

The chosen methodology for this paper is panel analysis. I will use an Arellano-Bond GMM model to regress the crisis index, which is a measure of currency crisis pressures, on indices of financial fragility, current account imbalances, foreign reserves adequacy, and fundamentals’ performance. The crisis index is a weighted average of changes in the exchange rate and in foreign reserves; it indicates a speculative attack against a currency, as the latter can be signaled by a depreciation of the exchange rate or a contraction in foreign reserves. Regressing this index on the ones mentioned above will help measure the effects on the crisis index of not only financial fragility and imbalances, but also their fiscal implications, as well as both financial and real economy weaknesses. Furthermore, I will test whether there is a link between speculative pressure and exchange rate overvaluation, banking-sector fragility, and the foreign reserves level. As Feridun (2009) notes when exploring exchange market pressure in Turkey, it has been suggested in crisis literature that currency crises are more likely when the exchange rate is appreciated relative to its long-run average, the banking system is fragile, and the level of foreign reserves does not fully cover liabilities. In order to test this theory, I use an Arellano-Bond GMM model.

I also construct a second exchange market pressure index that is differently specified than the first one as well as a money market pressure index. During crises, the exchange market should experience an increased demand for foreign reserves and the money market – for liquidity. By using a group of variables on the banking sector, “real” economy, fiscal situation, and external balance as listed below and regressing the two indices on these variables, we can explore which of these factors determine crises. Following this approach, I use two models. First, I run both fixed and random effects panel regressions to study the relationship between the two indices and the explanatory variables. Second, I set the EMP/IMP relationship as a binary variable and explore the relationship between this binary and the other variables through a panel logit model. The goal is to see how the explanatory variables influence the probability of a crisis.

Since it has been suggested in the literature on crises that financial interconnectedness is not the sole channel of contagion and that trade relations can contribute as well, I will additionally use a binary probit model to estimate crisis incidence
based on volume of trade and macroeconomic controls. It is reasonable to assume that the more a country trades with others, the stronger the transmission in case of a crisis. By using the probit model, I will test whether or not stronger trade and therefore connectedness of economies lead to larger spillover effects and an increased likelihood of crisis incidence.

4. Empirical Results and Analysis
4.1 Current Account Imbalances and Non-performing Loans
Financial fragility has often been associated with crises. Lending and asset booms can act as sources of stress, as they can increase banks’ exposure to risky sectors. As a general rule in finance, the higher the leverage, the higher the potential for profit or upside risk, but also the higher the downside risk. If we have too many players in the financial system that are highly leveraged, all it takes is one sector to fall to set off a domino effect in all others. We saw this in 2008, when the subprime mortgage market went underwater, which lead to contagion in the whole U.S. financial system before the crisis spread around the world. This is hardly surprising, considering the interconnectedness of financial markets. Additionally, the financial liberalization that took place before the subprime crisis was both ubiquitous and extensive; it made banks and other financial institutions less relaxed about capital cushions and detailed credit check-ups. As a result, the amount of non-performing loans (NPLs), for example, increased. The first regressions will deal with that exact issue to determine whether NPLs affect the crisis index.

Following Corsetti et al.’s (1999) paper on the Asian final crisis, I create similar indices for the 25 emerging economies explored in the paper for the period 2000-2015. I am constructing a crisis index to measure speculative pressure on a country’s currency. The crisis index, \( \text{ind} \), is a weighted average of a country’s change in the domestic exchange rate relative to the U.S. dollar and its change in foreign reserves.

\[
\text{ind}_{i,t} = 0.75 \times \Delta \text{exch}_{i,t} + 0.25 \times \Delta \text{forres}_{i,t}
\]

The assigned weights are the same ones that Corsetti et al. (1999) used in their analysis of the Asian financial crisis of 1997. If the exchange rate depreciates sharply, that can signal a speculative currency attack; similarly, if foreign reserves contract, this could be another signal of an attack, as reserves can be decreased to prevent a devaluation. Therefore, a large negative value of the crisis index signifies a more severe currency crisis. While an increase in domestic interest rates can also point to a speculative attack, it is highly correlated with decreased foreign reserves. As a result, adding this variable becomes unnecessary and potentially harmful for empirical work. For that reason, the crisis index is composed of changes (\( \Delta \)) in \( \text{exch} \), national currency per U.S. dollar, and \( \text{forres} \), foreign reserves.

Following Corsetti et al.’s (1999) example, I have constructed indices of financial fragility, current account imbalances, and foreign reserves adequacy and fundamentals’ performance. A major indicator of weaknesses in the banking system are non-performing loans. In the data, they are expressed as a percentage of total gross loans. They are interacted with the lending boom variable, defined as the growth rate of credit to the private sector as a percentage of GDP. I create a new variable, \( \text{nplb} \), which takes the value of NPL, if the lending boom is positive, and 0 otherwise. This is done because usually NPLs do not significantly increase a country’s vulnerability to a crisis unless accompanied by excessive bank lending. Since a country’s vulnerability to both currency and financial crises increases with bailout costs, I also create another indicator, \( \text{nply} \), which takes this into account. If we assume that NPLs are a good proxy for potential bailouts, the associated fiscal costs can be expressed as NPLs as a percentage of GDP.

Another major determinant of crisis pressure is the current account of a country. However, this is only meaningful if expressed as a share of GDP and even in that case, it is hard to know whether changes in the CA balance are due to investments in profitable projects or are eroding profitability. Nevertheless, if a CA deficit is accompanied by exchange rate appreciation, this is a red light according to the consensus in academic literature. For that reason, the current account index, \( \text{cai} \), equals the current account balance as a share of GDP if we have an exchange rate appreciation; otherwise, it is 0.1

1 Note: Corsetti et al. try regressions with both 0% and 10% appreciation thresholds and obtain similar results. In the case of the 2008-9 crisis, however, we rarely see such significant appreciations of over 10% and the cai variable be-
In addition to these indices, fundamentals could also adversely affect economies. In particular, it is interesting to note how the insufficiency of foreign reserves affects external imbalances and financial fragility. We can measure the adequacy of foreign reserves when we consider them with respect to a measure of domestic liquid assets. If a country is strong with regards to foreign reserves, then the ratio of M1 to foreign reserves is in the lowest sample quartile and I assign the dummy variable d1res a value of 0. If it is weak, that is if M1/forres is above the lowest sample quartile, then I assign d1res a value of 1. The same process is repeated with M2 to form the dummy d2res.

Additionally, we can create a fundamentals dummy by focusing on current account imbalances and financial fragility. I created a weak fundamentals dummy, dfundb, which takes the value of 0 if fundamentals are strong, that is if either cai is in the highest sample quartile or nplb is in the lowest; otherwise, dfundb takes the value of 1, signifying weak fundamentals. Similarly, the dummy dfund indicates the same fundamentals qualities with regards to cai and nply.

It would also be useful to include a measure of “real” weaknesses in addition to the measures of financial fragility and current account imbalances. I test whether capital productivity is significant for the crisis index; it is important to note that excessive borrowing from foreign countries is less dangerous when it finances new investment rather than increase spending. Therefore, if a current account deficit is accompanied by rising investment, it could have potential benefits in the future and is not as problematic as one accompanied by lower savings rates. In order to measure investment efficiency, I take the incremental capital output ratio (icor), the ratio of the investment growth rate and the output growth rate, which is the measure that is widely used in empirical literature. Additionally, I also create the variable icorlb, which takes the value of icor when the lending boom is positive and 0 otherwise. That way, we can observe whether low capital productivity is more problematic in the presence of excessive lending. The expectation is that it should be, since if investment projects are not profitable and banks increase their lending, then they are likely not taking into account adverse selection and moral hazard issues, which can significantly increase default rates on loans and expose the financial system to the risk of a crisis. This is certainly reminiscent of the excessive lending to subprime borrowers prior to the Great Recession, and for that reason, checking the significance of icorlb is important to my analysis.

I begin with testing whether or not unit roots are present in the independent variables by running a series of Dickey-Fuller tests; it appears that no unit roots are present and the results are significant at a 1% significance level. Whereas Corsetti et al. use basic cross-sectional regressions, my data spans the period 2000-2015 and requires panel analysis. However, a fixed effects estimation will not do in this case for several reasons. First, there could be endogeneity between the dependent variables and the independent ones, especially between NPLs (and its derivative variables) and the crisis index. For example, we can have bidirectional causality between the two arising from the fact that a currency crisis can cause people to default on their loans, but an increase in NPLs could cause a crisis if significant enough. In addition, considering the data's time dimension of 16 years and the tendency of crisis indicators to depend on their past values, it is likely that the model is dynamic. After running both fixed effects and random effects regressions and observing insignificant results, this appears to be the case. Since misspecification of the model can produce weak regressions and I believe that the model could use more dynamics and an address of the endogeneity issue, I chose to run an Arellano-Bond generalized method of moments (GMM) estimation. Such an estimation allows for unobserved effects in the data, while working with lagged variables and addressing endogeneity issues. What is more, the panel data has a shorter time dimension and a larger country dimension, making it especially suitable for a GMM model. I use a twostep robust specification to make sure the model is free of heteroscedasticity, serial correlation, and covariance among regressors. I add the robust option to correct for the downward bias in standard errors arising from a twostep procedure. Post-estimation of autocorrelation was performed after every regression, ensuring that we have no AR(2) as required when using a GMM model.

I begin by running several basic regressions (Table 1, appendix) with interactive terms that include the dummy variable comes redundant if I use 10% as the threshold. I, therefore, use a threshold of 0%, that is any exchange rate appreciation. Corsetti et al. also consider these in regard to short-term debt and interest payments on foreign debt. However, since I have no access to the data source they used, Datastream, and could not find such data anywhere else, I have used only ratios with liquid assets to measure the adequacy of foreign reserves. Through the use of “estat abond” in Stata.
for low reserves. As can be observed from the initial regression of ind on cai and nplb, the signs are as expected, that is both a large current account deficit accompanied by exchange rate appreciation and a high share of non-performing loans during a lending boom are indicative of a more severe currency crisis. However, these results should be interpreted with caution, as they are not significant at the 10% level and should not be relied upon. Nevertheless, it is interesting to note that the lagged value of the dependent variable is significant at the 10% level, thereby confirming that there are dynamics in the model and GMM is an appropriate method of estimation.

In columns (2) and (3), the signs of cai and nplb are now “flipped,” but they are once again insignificant variables. Therefore, these inconsistent results should not come as a surprise. Once again, past values of the dependent variable are highly significant, suggesting a dynamic model structure. Considering the persistence of crises and the fact that economic recovery takes time, the high influence of lagged variables is expected. When crises conditions persist, it takes time for exchange rates and foreign reserves to rebound. Similarly, when the crisis index is high and the economy is not threatened, exchange rates and reserves tend to maintain healthy levels as a result of overall public confidence in the economy. Therefore, influenced by perceptions and trends, past values of the crisis index are likely to influence present ones.

Furthermore, when cai and nplb are interacted with the low reserves dummies, all interactive variables but one are significant. In the case of the regression with M1/forres, the interactive term between nplb and d1res is significant at the 5% level, while cai x d1res is not. In the regression with M2/forres, both interactive terms are significant. This could be due to the more extensive nature of M2 as compared to M1; it is broader and a better measure of liquidity. If we consider the M2/forres regression, it is evident that even though cai and nplb might not be significant on their own, when there are low reserves present in the economy, they can influence the crisis index as expected – that is, current account deficits associated with an exchange rate appreciation and high NPLs associated with lending booms can both worsen the severity of currency crises if accompanied by low reserves. This implies that structural imbalances play a role in the incidence of a crisis if accompanied by low reserves, whereas in times when reserves are sufficient, there is not enough statistical evidence to suggest that these imbalances can cause a crisis. When reserves are low, however, the country becomes more vulnerable to speculative attacks and a currency crisis.

Next, I take the regression with M1/forres and add two more interactive terms, cai x d1res x dfundb and nplb x d1res x dfundb. I do the same with M2/forres. In this case, all of the variables except nplb x d1res x dfundb and nplb x d2res x dfundb turn insignificant and inconsistent (Table 2, appendix). This captures the effect of NPLs accompanied by a lending boom, low reserves, and weak fundamentals. Both the M1 and M2 versions are highly significant, showing that NPLs can worsen a crisis, if foreign reserves are insufficient and the country is weak fundamentally. Considering the previous regression, interactive variables with non-performing loans are especially important for the onset of a crisis, even more so than current account imbalances.

4.2 Current Account Imbalances and Fiscal Costs
Now I turn to fiscal implications rather than financial fragility per se. Instead of nplb, the regressions now include nply, the proxy for the costs of fiscal bailouts. It is important to note that unlike the regressions above, these regressions displayed evidence of AR(2) autocorrelation. Therefore, the estimation results below are obtained by running the regressions with the independent variables’ differenced versions, which eliminated the correlation present.

Even in the first basic regression, nply is significant, unlike nplb in the corresponding regression from above. It loses significance in regressions (2) and (3) of Table 3, possibly because of the proximity of another variable, namely the interactive variable between nply and a low reserves dummy in both regressions. This indicates the crucial role of reserve adequacy for financial fragility and its fiscal implications. Low NPLs pose a threat not only to finance, but also to the government’s budget. For that reason, they could be connected to different types of crises, as they appear to influence the economy through more than one channel. Just as before, the variable cai is only significant if interacted with d2res, which could be taken as another indicator that the M2/forres ratio gives a better idea of reserves adequacy than the M1/forres ratio.
If to these basic regressions we also add $cai \times d1res \times dfundy$ and $nplb \times d1res \times dfundy$ and $cai \times d2res \times dfundy$ and $nplb \times d2res \times dfundy$ respectively (output in Table 4), we again end up with a bigger and also mostly insignificant regression, just like we did in the $nplb$ case above. In both cases, the interactive terms between $nply$ (or $nplb$) and low reserves on the one side and $nply$ (or $nplb$), low reserves, and weak fundamentals on the other are very strongly correlated across all possible pairs (over 0.90). This could be the reason behind the highly insignificant results and the models reducing to only two significant variables each. Nevertheless, in tables 2 and 4, we see the importance of $nply/nplb$ when accompanied by low reserves.

I also ran several different regressions, in which I regressed the crisis index on $icor$ or $icorlb$ and $cai$, $nplb$, or $nply$. In all cases, $icor$ or $icorlb$ are insignificant. Therefore, it is likely that crises onsets are much more dependent on financial fragility, fiscal costs, and external imbalances, which make a country susceptible to contagion, rather than productivity weaknesses such as investment inefficiency. While the latter can be an important factor in economic growth, it is not as influential when it comes to crisis incidence as the former factors are.

4.3 Determinants of Exchange Market Pressure

When exploring the determinants of exchange market pressure in Turkey, Mete Feridun (2009) suggests a similar set of factors that influence crises. He links speculative pressure to real exchange rate overvaluation, fragility in the banking sector, and foreign reserve levels.

I am creating similar indices for the 25 emerging economies that are explored in this paper. The dependent variable, the emerging market pressure index, is constructed in the following way:

$$emp_{it} = \alpha \Delta exch_{it} + \beta \Delta spread_{it} - \gamma \Delta resm1_{it}$$

The weights $\alpha$, $\beta$, and $\gamma$ are the inverses of the standard deviation of the given variable, whereas $\Delta$ represents percent change. The variable $exch$ is the exchange rate per USD, $spread$ is the spread between the 3-month deposit rate and the 3-month Treasury bill rate, and $resm1$ is the ratio of reserves to M1. The higher this index, the higher the pressure in the exchange market, as low reserves, devaluation, and large spreads are indicative of a strained economy.

The $emp$ index is regressed on three independent variables: $reer$, the percentage change in the real effective exchange rate obtained by IMF calculations; $m2res$, the ratio of M2 to international reserves; and the $bsfi$ (banking sector fragility index), shown below.

$$bsfi_{it} = \left( \frac{\Delta credit_{Lt} - \mu_{credit}}{\sigma_{credit}} + \frac{\Delta deps_{Lt} - \mu_{deps}}{\sigma_{deps}} + \frac{\Delta forliab_{Lt} - \mu_{forliab}}{\sigma_{forliab}} \right) / 3$$

$Credit$ refers to bank credits to the domestic private sector; $deps$ to bank deposits; and $forliab$ to banks’ foreign liabilities. The initial idea that Feridun (2009) presents is that the higher this index, the higher the risky behavior of banks and therefore, the more aggravated the exchange market.

Before beginning empirical work, I conducted unit root tests, which suggest that only $emp$ and $reer$ are free of unit roots. After first-differencing $bsfi$ and $m2res$, however, the unit root problem was solved, indicating that these two variables were of integrated order one. Having addressed these issues, the variables can now be freely used in regressions.

In his paper, Feridun first uses autoregressive distributed lag (ARDL) bounds-testing procedure in order to establish whether a long-run relationship exists between the speculative pressure and the three explanatory variables. Once he has established that, he runs Granger causality tests based on a vector error correction model and finds a long-run causality from the explanatory variables to exchange market pressure, as well as a feedback mechanism between banking-sector fragility and exchange market pressure. Considering that Feridun (2009) is only exploring the Turkish economy and this paper is analyzing 25 economies within a panel context, I decided to select several representative economies of each region,
Europe, Latin America, Middle East and Africa, and Asia. However, upon running ARDL and Granger causality tests for each country individually, the results were highly insignificant and inconsistent. I assume that this is because of the short time dimension – all the data behind this paper is annual, spanning the period from 2000 to 2015. Considering the data availability for all variables, I could not expand the time dimension further. In a time series model like VECM, a T of 16 is insufficient for proper analysis. Therefore, I borrowed Feridun’s indices and their construction methods and incorporate them into an Arellano-Bond GMM model that is more appropriate for dynamic panel data with larger N rather than T, as previously discussed. However, as I ran this model, I noticed that the lagged value of $emp$ is insignificant and furthermore, there was no AR(1) autocorrelation in the model’s post-estimation as is usually the case with GMM. If indeed some sort of adjustment mechanism is present between such variables as Feridun also found out with his VECM work, then it makes sense that a lagged value can turn insignificant if a short-run adjustment variable is instead more significant and driving the model towards long-run equilibrium. Regardless of whether that is the case or not, the insignificance of $emp$ led me to disregard dynamics and run a fixed effects model followed by a random effects model. The Hausman test on the estimates, operating under the null hypothesis that the preferred model is the random effects one, indicates that the null should be rejected, considering the p-value of 0.0000, which is less than any significance level. Therefore, fixed effects are to be preferred, and the results from the estimation are presented in Table 5.

Considering that the $emp$ index can be high when we see currency devaluation, an expansion of the interest rate spread, a decrease in international reserves, or a combination of the above, a high $emp$ index is indicative of pressure in the exchange market. As expected, overvaluations of the real exchange rate increases exchange market pressures, as does the lowering of foreign reserves. However, it must be noted that $d.m2res$ has a much lower coefficient, suggesting a smaller role that reserves play in terms of magnitude, even though they are highly significant for the $emp$ index. What is interesting to note here is the unexpected sign of the $bsfi$ index. Feridun (2009) discovers that a high $bsfi$ index aggravates crisis conditions, as it is an indication of excessive risk-taking behavior by banks. In my sample, however, the opposite seems to be true. In my opinion, the $bsfi$ index could be positively related to the $emp$ index, if banks are being negligent about credit risk or lending in an excessive way that eats up their capital cushions. However, considering that the $bsfi$ index includes deposits, credit to the private sector, and foreign liabilities without any indication of credit-worthiness, capital controls, or NPL data included in it, for example, one does not necessarily have to think that an increase in this index signifies risky behavior. It could also mean a pick-up in economic activity, an increased confidence about economic conditions, or recovery from a credit crunch, for example. In fact, it seems that interpreting this index can indeed go both ways – in region-specific regressions, it appears that the $bsfi$ index has a negative value for the regions of Asia, Europe, and Middle East and Africa, while having a positive value for the region of Latin America. This could be due to the high amount of non-performing loans in Latin America during the period 2000-2015 and the region’s wobbly economies throughout much of that period and not only during the subprime crisis. The $bsfi$ index could very well mean increased risk-taking behavior in this case. In Asia, on the other hand, the $bsfi$ index has the highest negative value, perhaps signifying times of economic rebound. Since Asian cultures also tend to be less risk-loving than Latin American ones, for example, we could also be seeing an indicator in which cultural differences play a role. Overall, I would consider indices like the ones incorporating NPLs, for example, to be a better measure of banking fragility, since we see real financial weaknesses in their computations, rather than variables that could be interpreted in multiple ways.

4.4 Determinants of Exchange Market and Money Market Pressure
Filipozzi and Harkmann (2010) also construct an exchange market pressure index in addition to constructing a money market pressure one. However, their specification is a bit different:

$$emp^2_{i,t} = \alpha \Delta \text{lnexch}_{i,t} + \beta \Delta \text{lnmm}_{i,t} - \gamma \Delta \text{lnforres}_{i,t}$$

The weights $\alpha$, $\beta$, and $\gamma$ represent the inverses of the standard deviation of the corresponding variables, namely change in the natural logarithm of exchange rates, money market rates, and foreign reserves, respectively. I have named this index $emp^2$ so as to distinguish it from the previous $emp$ index. The $imp$ index, measuring money market interest rates pressure, is derived in a similar way by using the inverses of the standard deviation of the change in the natural logarithm of the money market rate and the ratio of central bank funds to bank deposits. Each of these two indices is individually regressed
on a series of variables:

- `govt_borr` – change in government borrowing
- `L_m2res` – lagged value of the M2 to foreign reserves ratio; the expectation from literature is that this is a leading indicator
- `dom_cred` – changes in domestic credit
- `banks_for` – changes in the ratio of banks’ foreign assets to liabilities
- `cpi` – CPI-based inflation
- `l.return` – lagged value of local stock market index return; stock market prices can be a useful signal for an economy’s near future and are considered by many economists to be a leading indicator
- `l.reer` – lagged value of the real effective exchange rate, as measured by the IMF; the expectation from literature is that this is a leading indicator
- `export` – percentage change in exports
- `mm_spread` – spread between a country’s local money market rate and the money market rate of the U.S.4

After regressing the `emp2` index on the variables above both in a fixed-effect and random-effect fashion, I run a Hausman test, which strongly suggests using fixed effects. As can be observed from the estimates in Table 6, not all variables in the model proved to be significant – in fact, many of the insignificant ones have high p-values, so their coefficients should not be taken at face value. However, the model also provides us with several highly significant results. It has been suggested in literature that inflation, through increasing interest rates, can create pressure in the exchange markets. The positive value of its coefficient in the model supports this claim. The `m2res` variable’s positive relation to `emp2` is also in line with literature; a high ratio signifies a lax fiscal policy and increases the probability of a crisis, hence its positive sign in the model.

Stock returns are also highly significant and negatively related to the dependent variable; that is when returns are negative, we can expect an increase of the index – a greater probability of a crisis. Considering that I used the lagged value of stock returns, these results are consistent with the leading indicator theory of stock returns. Similarly, `reer` is also a significant leading variable. Increases in `reer` signify real exchange rate overvaluations, which can lead to devaluation pressure and an increase in the exchange market pressure index. An increase in the foreign assets to liabilities ratio, `banks_for`, can do the same as indicated by its positive sign. During periods of crisis, banks find it harder to finance themselves on the international markets, thereby decreasing their foreign liabilities, which increases their foreign assets to liabilities ratio. This difficulty to borrow can put pressure on exchange markets and increase the probability of a crisis. At the same time, if we have a rapid growth in domestic credit, this can also increase crisis pressures according to the regression results, probably working through the channel of offsetting financial markets.

Similarly to Filipozzi and Harkmann (2010), I also discover that `imp` is not as good of an indicator of crisis pressure as `emp2` is (Table 7). They attribute it to the fact that the `imp` is built on two variables that move in ways that are difficult to explain by our explanatory variables because they are sticky – overnight rates are such because they are closely related to central bank rates, which change rarely, and the ratio of central bank funds to bank deposits is also sticky. However, `l.reer` and `dom_cred` are both significant; what is more, they have similar coefficients to the ones they had in the `emp2` regressions, suggesting that both these variables are important determinants of the probability of a crisis, regardless of the crisis pressure measure used.

Additionally, we can use a binary logit model to explore these relationships further (Table 8). The variables `emp2` and `imp` take the value of 1 if they are higher than their respective thresholds and 0 otherwise. As suggested by Eichengreen et al. (1996), one can measure these threshold values by setting them at 1.5 standard deviations above the mean. Should an index exceed this level, it takes the value of 1, thereby creating a binary crisis variable.

---

4 Filipozzi and Harkmann use the Euribor, but they are only analyzing 3 Eastern European countries. In this paper, the sample contains countries from all over the world, so comparison to U.S. rates is much more applicable considering the status of the U.S. as a global economic power.
Only two variables are significant in these models. In the first regression, inflation behaves as predicted by theory and previous models; in the second model, high stock returns seem to add to crisis pressures. This is counterintuitive, unless we consider the possibility that perhaps this regression is capturing the effects of irrational exuberance. In any case, the results appear to be quite inconsistent, and perhaps the fixed effects regression based on the exchange market pressure index is a more trustworthy indicator of crisis pressures.

4.5 Trade and contagion
As it has been suggested in literature that crises are contagious and trade acts as a transmission channel, I am attempting to test this hypothesis. I regress the dummy \( d_{\text{crisis\_global}} \), which takes the value of 1 if the year is 2008 or 2009, the most severe crisis period, on the following macroeconomic controls:

- \( \text{trade} \) – total amount of exports + imports
- \( \text{lb} \) – lending boom variable
- \( \text{exp} \) – government expenditure as a percentage of GDP
- \( \text{ca} \) – current account as a percentage of GDP
- \( \text{gdpgr} \) – growth of GDP
- \( \text{m2res} \) – M2 to foreign reserves
- \( \text{cpi} \) – CPI-based inflation
- \( \text{exch} \) – national currency per USD

The results of the first regression are shown in Table 9. Several macroeconomic controls affect the probability of a crisis as expected. For example, GDP growth lowers it, as does exchange rate appreciation. However, trade seems to be insignificant in this model – whether or not a country trades with others does not affect its likelihood of being hit by a crisis. Because of this unexpected result, I take a look at regressions at the regional level (Table 10).

It is interesting to note that out of all the regions, Europe is the only one in which trade increases the likelihood of a crisis. In a region in which economic cointegration is stressed, trade acts as a crisis transmission channel because the trade relationships between European countries run very deep, thus making contagion much more likely. All but one country (Ukraine) of the European sample are members of the European Union, which means that they are much more connected economically than countries in Asia or Latin America, for example. Therefore, the significance of the trade variable solely in Europe is not that surprising. It is dangerous, however. As Joseph Stiglitz notes in his book *The Euro and its Threat to the Future of Europe*, European trade agreements appear to have generated negligent gains, while their role as a transmission channel of contagion has been much bigger, thereby making large parts of Europe's population worse off. Considering trade's significance in Europe in the model above and its high positive value, the regression does indicate a strong positive effect of trade on crisis incidence.

5. Conclusion
This paper analyzed the vulnerability of 25 emerging economies to crises and the factors that determine this vulnerability. Current account imbalances associated with exchange rate appreciation as well as non-performing loans associated with a lending boom indicate a more severe crisis. Crises can also spread through excessive risk-taking bank behavior, an overvaluation in the real effective exchange rate, or interest rate channels. The regression results indicate that crisis-prone fundamentals and susceptibility to transmission of adverse global shocks, especially when accompanied by insufficient foreign reserves and weak macro fundamentals, can strongly increase the likelihood of a crisis. The channels are multiple, as contagion can spread through stock markets or trade; weaknesses in both the real economy and the financial sector can increase fragility and vulnerability to a crisis; and excessive government spending and borrowing can affect the government's capability of fiscal intervention. Considering how many possible indicators of fragility or contagion channels can be included in models, any model run on the topic is bound to suffer from omitted variables and unexplored explanations. A good topic for further research would be to explore countries in a time-series setting rather than a panel one, as in the former case, models that are unavailable for panel analysis can provide a much clearer idea of transmission mechanisms and volatility. To get a clear view of crisis incidence, it is necessary to look both from a panel and a case study viewpoint.
References
## Output Tables

**Table 1. Basic regressions; Arellano-Bond GMM method**

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Regression with M1/ forres</th>
<th>Regression with M2/ forres</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$l.ind$</td>
<td>0.212*</td>
<td>0.202***</td>
</tr>
<tr>
<td></td>
<td>(0.108)</td>
<td>(0.0651)</td>
</tr>
<tr>
<td>$cai$</td>
<td>0.00228</td>
<td>-0.00140</td>
</tr>
<tr>
<td></td>
<td>(0.00183)</td>
<td>(0.00313)</td>
</tr>
<tr>
<td>$nplb$</td>
<td>-0.00234</td>
<td>0.00113</td>
</tr>
<tr>
<td></td>
<td>(0.00235)</td>
<td>(0.00183)</td>
</tr>
<tr>
<td>$cai \times d1res$</td>
<td>0.00372</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00802)</td>
<td></td>
</tr>
<tr>
<td>$nplb \times d1res$</td>
<td>-0.00555**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00227)</td>
<td></td>
</tr>
<tr>
<td>$cai \times d2res$</td>
<td></td>
<td>0.00391*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00220)</td>
</tr>
<tr>
<td>$nplb \times d2res$</td>
<td></td>
<td>-0.00705***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00239)</td>
</tr>
<tr>
<td>constant</td>
<td>0.0137</td>
<td>0.0183*</td>
</tr>
<tr>
<td></td>
<td>(0.0134)</td>
<td>(0.0107)</td>
</tr>
<tr>
<td>N</td>
<td>320</td>
<td>316</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01
**Table 2.** The role of fundamentals and reserves; Arellano-Bond GMM method; dependent variable – ind

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Regression with M1/forres (1)</th>
<th>Regression with M2/forres (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>l.ind</td>
<td>0.178** (0.0772)</td>
<td>0.198** (0.0869)</td>
</tr>
<tr>
<td>nplb<em>d1res</em>dfundb</td>
<td>-0.00588** (0.00263)</td>
<td></td>
</tr>
<tr>
<td>nplb<em>d2res</em>dfundb</td>
<td>-0.00718*** (0.00137)</td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>0.0199** (0.0101)</td>
<td>0.0233** (0.00933)</td>
</tr>
</tbody>
</table>

N = 316 310

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

**Table 3.** Fiscal implications basic regressions

Arellano-Bond GMM method; dependent variable – ind

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Regression with M1/forres (1)</th>
<th>Regression with M2/forres (2)</th>
<th>Regression with M2/forres (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>l.ind</td>
<td>0.214** (0.0954)</td>
<td>0.232*** (0.0686)</td>
<td>0.127 (0.0865)</td>
</tr>
<tr>
<td>d.cai</td>
<td>0.00140 (0.00289)</td>
<td>-0.00200 (0.00152)</td>
<td>-0.00368 (0.00268)</td>
</tr>
<tr>
<td>d.nply</td>
<td>-0.0000742** (0.0000333)</td>
<td>-0.0000362 (0.0000267)</td>
<td>-0.0000450 (0.0000439)</td>
</tr>
<tr>
<td>d.(cai x d1res)</td>
<td></td>
<td></td>
<td>0.00608 (0.00482)</td>
</tr>
<tr>
<td>d.(nply*d1res)</td>
<td>-0.0000748** (0.0000337)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
\[d.(cai \times d2res)\]
\[
\begin{array}{c}
0.00684^* \\
(0.00377)
\end{array}
\]

\[d.(nply\times d2res)\]
\[
\begin{array}{c}
-0.000131^{**} \\
(0.0000524)
\end{array}
\]

constant  
\[
\begin{array}{c}
0.0106 \\
(0.00767)
\end{array}
\]
\[
\begin{array}{c}
0.00839 \\
(0.00698)
\end{array}
\]
\[
\begin{array}{c}
0.0524^{***} \\
(0.0177)
\end{array}
\]

N 316 309 307

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

Table 4. The role of bail-out costs, fundamentals, and reserves

Arellano-Bond GMM method; dependent variable – ind

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Regression with M1/forres (1)</th>
<th>Regression with M2/forres (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>l.ind</td>
<td>0.174* (0.0893)</td>
<td>0.201** (0.0827)</td>
</tr>
<tr>
<td>d.(nply*d1res)</td>
<td>-0.0000783** (0.000386)</td>
<td></td>
</tr>
<tr>
<td>d.(nply*d2res)</td>
<td>-0.000109*** (0.000270)</td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>0.0102 (0.00878)</td>
<td>0.00878 (0.00794)</td>
</tr>
</tbody>
</table>

N 309 302

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01
### Table 5. Explaining the *emp* index

Fixed effects estimation; dependent variable – *emp*

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>rer</em></td>
<td>1202.6*</td>
</tr>
<tr>
<td></td>
<td>(503.9)</td>
</tr>
<tr>
<td><em>D_bsi</em></td>
<td>-249.6*</td>
</tr>
<tr>
<td></td>
<td>(122.7)</td>
</tr>
<tr>
<td><em>D_m2res</em></td>
<td>0.353***</td>
</tr>
<tr>
<td></td>
<td>(0.0438)</td>
</tr>
<tr>
<td>constant</td>
<td>129.6***</td>
</tr>
<tr>
<td></td>
<td>(42.39)</td>
</tr>
</tbody>
</table>

N 268

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

### Table 6. Estimating *emp2*

Fixed effects estimation; dependent variable – *emp2*

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>cpi</em></td>
<td>0.0557**</td>
</tr>
<tr>
<td></td>
<td>(0.0216)</td>
</tr>
<tr>
<td><em>l_m2res</em></td>
<td>0.000287***</td>
</tr>
<tr>
<td></td>
<td>(0.0000997)</td>
</tr>
<tr>
<td><em>l_return</em></td>
<td>-0.618***</td>
</tr>
<tr>
<td></td>
<td>(0.220)</td>
</tr>
<tr>
<td><em>l_reer</em></td>
<td>2.184*</td>
</tr>
<tr>
<td></td>
<td>(1.271)</td>
</tr>
<tr>
<td><em>export</em></td>
<td>0.604</td>
</tr>
<tr>
<td></td>
<td>(0.673)</td>
</tr>
<tr>
<td>Independent variable</td>
<td>(1)</td>
</tr>
<tr>
<td>----------------------</td>
<td>------------</td>
</tr>
<tr>
<td>cpi</td>
<td>0.0113</td>
</tr>
<tr>
<td></td>
<td>(0.0184)</td>
</tr>
<tr>
<td>l.m2res</td>
<td>0.00000648</td>
</tr>
<tr>
<td></td>
<td>(0.0000139)</td>
</tr>
<tr>
<td>l.return</td>
<td>-0.0990</td>
</tr>
<tr>
<td></td>
<td>(0.201)</td>
</tr>
<tr>
<td>l.reer</td>
<td>2.112*</td>
</tr>
<tr>
<td></td>
<td>(1.144)</td>
</tr>
<tr>
<td>export</td>
<td>-0.0683</td>
</tr>
<tr>
<td></td>
<td>(0.625)</td>
</tr>
<tr>
<td>mm_spread</td>
<td>0.0144</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01
Table 8. Logit regressions

Dependent variables – binary variants of emp2 and imp respectively

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cpi</td>
<td>0.0714*</td>
<td>0.0317</td>
</tr>
<tr>
<td></td>
<td>(0.0404)</td>
<td>(0.0559)</td>
</tr>
<tr>
<td>l.m2res</td>
<td>-0.000118</td>
<td>-0.0000868</td>
</tr>
<tr>
<td></td>
<td>(0.000186)</td>
<td>(0.000127)</td>
</tr>
<tr>
<td>l.return</td>
<td>-0.324</td>
<td>1.486**</td>
</tr>
<tr>
<td></td>
<td>(0.639)</td>
<td>(0.579)</td>
</tr>
<tr>
<td>l.reer</td>
<td>3.838</td>
<td>-0.527</td>
</tr>
<tr>
<td></td>
<td>(3.327)</td>
<td>(3.892)</td>
</tr>
<tr>
<td>export</td>
<td>1.038</td>
<td>-0.158</td>
</tr>
<tr>
<td></td>
<td>(1.939)</td>
<td>(2.079)</td>
</tr>
<tr>
<td>mm_spread</td>
<td>0.00328</td>
<td>-0.0464</td>
</tr>
<tr>
<td></td>
<td>(0.0438)</td>
<td>(0.0588)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01
### Table 9. Trade and contagion

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>trade</td>
<td>-5.54e-18</td>
</tr>
<tr>
<td></td>
<td>(6.65e-18)</td>
</tr>
<tr>
<td>lb</td>
<td>0.940</td>
</tr>
<tr>
<td></td>
<td>(0.630)</td>
</tr>
<tr>
<td>exp</td>
<td>-0.00972</td>
</tr>
<tr>
<td></td>
<td>(0.0232)</td>
</tr>
<tr>
<td>ca</td>
<td>0.000606</td>
</tr>
<tr>
<td></td>
<td>(0.00936)</td>
</tr>
<tr>
<td>gdpgr</td>
<td>-0.116***</td>
</tr>
<tr>
<td></td>
<td>(0.0234)</td>
</tr>
<tr>
<td>m2res</td>
<td>0.000276*</td>
</tr>
<tr>
<td></td>
<td>(0.000150)</td>
</tr>
<tr>
<td>cpi</td>
<td>0.0146</td>
</tr>
<tr>
<td></td>
<td>(0.0125)</td>
</tr>
</tbody>
</table>

Probit model; dependent variable $- d_{crisis\_global}$

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>exch</td>
<td>-0.00131*</td>
<td>(0.000739)</td>
</tr>
<tr>
<td>constant</td>
<td>-0.722</td>
<td>(0.449)</td>
</tr>
</tbody>
</table>

N = 355

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01
### Table 10. Regional significance of trade

Probit model; dependent variable – *d\_crisis\_global*

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Latin America</th>
<th>Europe</th>
<th>Middle East and Africa</th>
<th>Asia</th>
</tr>
</thead>
<tbody>
<tr>
<td>trade</td>
<td>-1.32e-16</td>
<td>2.22e-15*</td>
<td>5.02e-16</td>
<td>-6.59e-19</td>
</tr>
<tr>
<td></td>
<td>(8.77e-17)</td>
<td>(1.32e-15)</td>
<td>(7.36e-16)</td>
<td>(7.30e-18)</td>
</tr>
<tr>
<td>lb</td>
<td>-0.772</td>
<td>1.458</td>
<td>0.588</td>
<td>-2.023</td>
</tr>
<tr>
<td></td>
<td>(3.040)</td>
<td>(1.248)</td>
<td>(0.980)</td>
<td>(3.367)</td>
</tr>
<tr>
<td>exp</td>
<td>0.0301</td>
<td>0.0618</td>
<td>-0.00373</td>
<td>-0.0937</td>
</tr>
<tr>
<td></td>
<td>(0.0741)</td>
<td>(0.156)</td>
<td>(0.0408)</td>
<td>(0.112)</td>
</tr>
<tr>
<td>ca</td>
<td>-0.00304</td>
<td>-0.132***</td>
<td>-0.00231</td>
<td>0.0825</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.0439)</td>
<td>(0.0116)</td>
<td>(0.0564)</td>
</tr>
<tr>
<td>gdpgr</td>
<td>-0.143**</td>
<td>-0.239***</td>
<td>-0.0591</td>
<td>-0.253***</td>
</tr>
<tr>
<td></td>
<td>(0.0680)</td>
<td>(0.0626)</td>
<td>(0.0426)</td>
<td>(0.0925)</td>
</tr>
<tr>
<td>m2res</td>
<td>0.00109*</td>
<td>-0.00387</td>
<td>-0.00334</td>
<td>0.000220</td>
</tr>
<tr>
<td></td>
<td>(0.000619)</td>
<td>(0.00408)</td>
<td>(0.00351)</td>
<td>(0.000164)</td>
</tr>
<tr>
<td>cpi</td>
<td>-0.0527</td>
<td>-0.000452</td>
<td>0.0228</td>
<td>0.0956</td>
</tr>
<tr>
<td></td>
<td>(0.0575)</td>
<td>(0.0251)</td>
<td>(0.0209)</td>
<td>(0.0621)</td>
</tr>
<tr>
<td>exch</td>
<td>-0.00417</td>
<td>-0.0368</td>
<td>0.00237</td>
<td>-0.00109</td>
</tr>
<tr>
<td></td>
<td>(0.00268)</td>
<td>(0.0270)</td>
<td>(0.00916)</td>
<td>(0.000804)</td>
</tr>
<tr>
<td>constant</td>
<td>-0.966</td>
<td>-2.472</td>
<td>-0.887</td>
<td>-0.278</td>
</tr>
<tr>
<td></td>
<td>(1.054)</td>
<td>(2.908)</td>
<td>(0.921)</td>
<td>(1.561)</td>
</tr>
</tbody>
</table>

N 74 99 125 57

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01
List of countries
Argentina
Brazil
Bulgaria
Chile
Colombia
Czech Republic
Egypt
Estonia
Hungary
Indonesia
Israel
Jordan
Kuwait
Malaysia
Mexico
Nigeria
Oman
Pakistan
Poland
Romania
Russian Federation
Saudi Arabia
South Africa
Turkey
Ukraine
### Data Sources

<table>
<thead>
<tr>
<th>Series</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>M2</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>Foreign reserves</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>CPI inflation</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>Exchange rate (national currency per USD)</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>Real effective exchange rate</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>GDP</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>Banks' foreign liabilities</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>Money market rates</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>Deposit interest rates</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>Banks' foreign assets</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>Central bank funds</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>Government borrowing</td>
<td>International Financial Statistics</td>
</tr>
<tr>
<td>Current Account (% of GDP)</td>
<td>World Bank</td>
</tr>
<tr>
<td>Domestic credit to the private sector by banks (% of GDP)</td>
<td>World Bank</td>
</tr>
<tr>
<td>Government expense (% of GDP)</td>
<td>World Bank</td>
</tr>
<tr>
<td>Exports (% of GDP)</td>
<td>World Bank</td>
</tr>
<tr>
<td>NPLs (% of gross loans)</td>
<td>World Bank</td>
</tr>
<tr>
<td>Trade (% of GDP)</td>
<td>World Bank</td>
</tr>
<tr>
<td>Gross fixed capital formation growth</td>
<td>World Bank</td>
</tr>
<tr>
<td>Treasury rates</td>
<td>Federal Reserve Economic Data</td>
</tr>
<tr>
<td>Banks' deposits</td>
<td>Federal Reserve Economic Data</td>
</tr>
<tr>
<td>Stock market price level</td>
<td>Bloomberg Terminal</td>
</tr>
</tbody>
</table>
Working women: The Impact of Financial Crises on Female Labor Force Participation Rates in Developing Countries

ALYSON MATTHEWS
Georgetown University

ABSTRACT
In the wake of recent financial crises, female labor force participation rates in high income countries have not followed trends explained by gender norms and have instead remained steady and even increased. In contrast, little is known about the impact of financial crises on female labor force participation rates in middle income countries. I suggest a new framework for analyzing labor force participation rates in these countries by looking at country-level supply and demand determinants of female labor force participation. Specifically, I hypothesize that levels of financial stability are key in understanding changes in female labor force participation rates. Using regression analysis and testing existing models for explaining female labor force participation rates in 96 countries from 1990 to 2011, I find some support for my hypothesis, that domestic financial stability impacts female labor force participation. Accounting for instability also yields contradictory results for existing models of female labor force participation, suggesting that the financial health of a country has wide ranging implications and should be accounted for in future work.
1. Introduction
If serving as guardian of her family is one of a woman's central identities, whether as a wife, mother, or daughter, how should she respond when her family experiences economic hardship? Women increasingly face this challenge as the pace of financial crises hitting developed and developing countries increases. During the most recent financial crisis in 2008, OECD nations saw relatively steady labor force participation rates, although employment dropped by an average of three percentage points\(^1\). Strikingly, labor force participation rates and employment levels among women in these countries nations did not decline as traditionally expected by gender norms. Women seemed to protect their families by either joining or remaining in the workforce. Among developed countries, women have remained in the workforce even during economic downturns due to changes in family structures (Alpekina and Luo, 2015), government policies to encourage greater gender balance in the workforce (Thevenon, 2013), and changing attitudes about gender roles (Fortin and Lemieux, 1998).

In contrast to high income countries, emerging markets and lower income countries faced the financial crisis with limited ability to manage shocks to their domestic economic systems. The coupling of uncertainty and rising costs has the potential to disrupt trade and investment in developing countries, further weakening employment levels and labor force participation rates in the formal sector while driving many women into informal labor, raising the possibility of exploitation and unsafe working conditions (Mehra and Gammage, 1999). Because developing countries lack the capability to absorb economic shocks, they are often subject to market forces in determining labor allocation within their countries.

The developed country framework examining two main channels for determining participation rates, labor demand and labor supply, is a useful dichotomy for analyzing developing countries. Many firms demand female workers as cheaper, more flexible labor during cyclical downturns since they are more willing to work part time and often leave employment for family reasons (Mehra and Gammage, 1999). In both developing and developed country contexts, service sector firms often prefer women over men to handle customer interactions for their supposed higher levels of interpersonal skills. As service industries rise in value across the world, women's employment opportunities may rise, regardless of a country's level of development and industrialization. Yet as a counter trend in developing countries, textile and simple manufacturing industries are increasingly technologized, leading firms to shed female workers as machines replace human labor (Juhn and Potter, 2006). On the supply side, reasons for greater levels of female labor force participation are also diverse. Women in developing countries most often choose to enter the workforce to maintain family income (Mehra and Gammage, 1999), especially if their husbands lose their jobs in economic downturns. Yet technological improvements in the home may also reduce domestic workloads, providing additional time for a woman to work outside the home (Juhn and Potter, 2006) regardless of her family's economic status.

The current field of literature examining female labor force participation rates during financial crises in developing countries is small and varied. I seek to add to the literature by utilizing a cross-country analysis of 96 middle income countries from 1990-2011 to determine large-scale drivers of female labor force participation. Coupled with demand and supply side determinants of women's labor force participation, I hypothesize levels of stability in a country's domestic financial market will have varying effects on female labor force participation rates. During economic booms, higher levels of stability will lead to declines in the labor force participation rates of women, in line with the literature that easier access to credit and stronger inflows of Foreign Direct Investment (FDI) will increase opportunities for men and therefore deter firms from hiring women and women from entering the labor market. In contrast, I hypothesize that countries with higher levels of domestic financial stability will experience increases in female labor force participation during crises. Beyond levels of development, firms are more likely to have access to domestic and international lines of credit in countries with higher levels of domestic financial stability. Access to loans and business opportunities enables firms to maintain levels of production by cutting costs through increased hiring of women, even when worldwide demand for goods shrinks. This new direction brings together both supply and demand factors of female labor force participation, while adding to the relatively small body of literature studying the impact of financial crises on labor force participation rates in developing countries.

---

I begin with a review of the literature on female labor force participation in developing countries, the effects of financial crises on developing countries, and the small but diverse literature on the impact of financial crises on female labor force participation rates in developing countries. Section Three presents my theory, Section Four introduces my data and key variables, and Section Five outlines my methodological approach. I present my results in Section Six and conclude in Section Seven.

2. Literature Review

2.1 Explaining Female Labor Force Participation in Developing Countries

Early work studying female labor force participation in developing countries centers on Goldin's (1994) "feminization U-Hypothesis", a precursor to the gendered Kuznets Curve (Eastin and Prakash, 2013) for economic and political rights. Economists have applied the curvilinear framework to model issues in environmental economics and more recently, women's development. The feminization U-Hypothesis suggests that only lowest income countries and highest income countries experience strong levels of female labor force participation since women are employed in very low skill or very high skill jobs, respectively. In contrast, middle income countries experience declines in female labor force participation since sectoral shifts in production favor men over women and women often view staying home as a symbol of their husband's financial stability. Gaddis and Klasen (2011) test the U-Hypothesis among 28 developing countries and find little evidence for the curve when controlling for endogeneity, citing few differences in sectoral changes across country income levels. However, their study validated more recent sectoral explanations for growth and decline in female labor force participation based on the strength of different sectors within an economy.

This competing explanation starts with the premise that developing countries are increasingly hands off in their intervention and regulation of labor markets, preferring to allow firms to hire the cheapest sources of labor to attract investment, rather than putting workplace protections and anti-discrimination laws in place (Mehra and Gammage, 1999). The laissez-faire approach to the labor market and job allocation furthers the impact of gendered employment by sector, as agricultural and informal sectors are increasingly female dominated, since these sectors allow women to maintain domestic responsibilities while also holding additional employment outside the home.

Coupled with changes in the sectoral allocation of jobs in developing countries, discrimination in the status of women workers, as secondary or supplemental income rather than a core part of household income, further contributes to reductions in female labor force participation. Rubery and Rafferty (2013) find that gender segregation within sectors of the economy may have an impact on women's employment and labor force participation, since women may be classified as flexible labor easily shed during times of crisis. In contrast to the view that traditional gender roles account for declines in female labor force participation, a competing body of literature finds that sectoral segregation may also explain why women retain their jobs during economic downturns. Many women are often employed in "buffer jobs", or labor intensive jobs that enable firms to adjust employment numbers and wages with variations in demand from consumers to keep costs low (Bettio, 1998).

2.2 Effects of Financial Crises on Developing Country Labor Markets

As the global economy becomes increasingly integrated, linkages between developed economies, emerging markets, and developing countries not only result in production efficiencies due to knowledge spillovers, but also contagion effects during periods of economic instability. Increasingly, financial crises originating in one country or region quickly have world-wide implications for investors, companies, and workers. Lin (2008) points to a global feedback loop from north to south that is especially vulnerable during crises, since decreased demand in high income countries leads to declines in terms of trade for industrializing countries, risk aversion leads to capital flight from emerging markets, and lower levels of global employment reduces remittances from migrants working in developed countries to families in their home countries. Griffith-Jones and Ocampo (2009) and Demirguc-Kunt and Detragiache (1998) suggest more severe downturns from financial crises occur among emerging markets rather than low income countries because the latter are far less integrated or dependent on global private capital markets and investment inflows to drive the domestic economy. Emerging markets, they argue, experience significant declines in terms of trade in key export sectors, while financial vulnerabilities exacerbated by crises lead capital holders to flee to more trustworthy investments in developed countries.
Though financial crises can have serious implications for capital flows and trade between countries, previous research suggests that FDI remains relatively stable during crises. Yet even in countries with constant inflows of FDI the potential for benefits of this type of investment largely depends on the health and stability of domestic financial markets. Agosin and Machado (2005) find that when FDI enters into sectors with high levels of domestic investment, FDI typically crowds out domestic operations in favor of lower-cost, higher-efficiency opportunities. In contrast, when FDI enters in sectors that complement existing domestic industries, host countries see gains in both domestically operated sectors and foreign dominated sectors.

Because developing countries experience differing impacts of financial instability on FDI and capital flows during financial crises, their ability to manage changes in domestic labor demand and labor force participation varies. Previous literature suggests that changes in demand and trade particularly affect women, since most women work in labor intensive industries which typically rely on local business-owners who contract their work to multinational firms (Joekes, 1995). Maloney (2004) finds that declines in the labor market as a result of financial crises in Latin American countries shifted workers, especially women, towards informal sector work because of the flexibility it provided laborers. This effect was most noticeable in employment in domestic firms, because they cut costs in production by relying on informal workers paid at lower rates to compete with more efficient and larger MNCs. Examining Indonesia's economic crisis, Manning (2000) suggests that because labor markets remained highly flexible during the crisis despite changing sectoral demands for labor, Indonesia lessened the burden of the crisis on its labor market compared to more rigid labor markets in Latin American countries. It is clear that trade liberalization and financial openness are a double-edged sword for emerging economies. Greater access to credit and trade growth are paths to development for many emerging markets, yet these countries were also most impacted by the financial crises, which often have severe consequences for domestic labor markets.

2.3 Bringing the Literature Together: Financial Crisis and Female Labor Force Participation

Stemming from concerns about the changing trends in the labor force, a newer body of literature examines the impact of financial crises and instability on the labor force participation rates of women and men. The vulnerability of developing countries to shocks in export demands and capital investment suggests a possible link between economic stability and changes to the employment landscape for men and women. Cho and Newhouse (2013) support the findings that sectoral differences are important for women since many firms relying on cheap labor hire women since they are less likely to be the breadwinners of their families, to require severance pay, and to demand improved benefits and rights. Mehra and Gammage (1999) also show that during crises lower income women enter the labor force to supplement family income while wealthier women choose to leave the labor market to ensure the “core” male labor force remains employed in industries and positions requiring higher skill levels. Looking at the East Asian financial crisis and credit crunch in the Philippines, Lim (2000) shows that female intensive service sectors were not impacted by a decline in employment, since inflation and capital flight led to the collapse of the male dominated manufacturing sector and as a result, women served as the key coping mechanisms for families to weather the crisis.

3. Theory

Previous literature has found varying implications of financial crises for female labor force participation, although these studies have mainly focused on single-country analyses. Additionally, these studies focus on micro-level supply and demand effects, using survey data rather than panel data over an extended period of time. I intend to add to the field of literature by studying the role domestic financial stability plays in determining female labor force participation rates during financial crises for a cross-country sample over 21 years. Specifically, I hypothesize that countries with more pre-crisis stability and higher levels of economic globalization will experience declines in female labor force participation (in favor of male labor force participation) during normal times. This hypothesized trend supports views that more stable countries attract higher investment flows from both FDI and institutional investors, allowing domestic and foreign firms to hire more expensive laborers, most often men (Kose et al., 2003). Reduced demand for low-skilled and cheap workers coupled with high and steady incomes limit the imperative for wives, daughters, and sisters to work outside the home. Rising wages for females is also associated with higher levels of development, which pushes women out of the workforce as foreign-owned firms move to cheaper production centers and domestic firms cut women to limit costs (Seguino and Grown, 2006). Since firms owned and operated by foreign companies rely on FDI rather than domestic sources of lending
(Borensztein et al., 1994) and domestic firms in financially stable countries have an easier time accessing domestic and international credit markets, greater levels of stability in the domestic financial sector of a country acts as a marker of prosperity and reduces the likelihood that women will enter the labor market.

However, I hypothesize that during financial crises, more financially stable countries will see increases in female labor force participation rates as women enter the labor market to supplement their families’ incomes. Women are overwhelmingly employed in firms owned and operated by domestic employers (Maloney, 2004), which suggests that labor demand is tied to the ability of domestic firms to access credit and continued investment flows. During financial crises, FDI remains relatively steady while foreign institutional lending can quickly dry up in financially unstable countries as investors look to minimize risk in their portfolios (Aghion et al., 2003). In turn, this negatively impacts the ability of domestic banks and lending institutions to provide lines of credit to domestic firms. More financially stable countries may attract continued investment from foreign sources, thereby improving opportunities for domestic lending (Hermes and Lensink, 2010 and Lensink, 2001). This allows domestic firms to access credit domestically and internationally and hire female workers more likely to maintain production levels at cheaper labor rates. Beyond levels of development, more financially stable countries are better able to provide their domestic firms with paths for economic growth and women with opportunities to supplement family income during economic downturns.

4. Data
To test my hypotheses, I use data from the World Bank's World Development Indicators\(^2\) and Global Financial Development\(^3\) datasets. Surveying 96 lower-middle and middle income countries from 1990-2011, I use the Female Labor Force Participation Rate (FLFPR) and the Ratio of Female to Male Labor Force Participation Rates (RLFPR) as my dependent variables. The first measure provides a direct measure for female labor force participation rates, in line with existing literature in developed countries. Utilizing the ratio of the two labor force participation rates allows for analysis on possible differing impacts for men and women, in line with literature studying the effects of financial crises on the labor force participation rates in developing countries.

My independent variables of interest are measures of dependence on trade, foreign direct investment, level of development proxied by GDP, and measures of financial instability. I use trade and FDI as a percentage of GDP to capture the dependence of a country on international trade and financial flows. As a measure of financial stability, I use bank z-score, an indicator of the stability of a country’s domestic financial sector. Bank z-score is a World Bank constructed ratio of “return on assets plus the capital–asset ratio to the standard deviation of return on assets. If profits are assumed to follow a normal distribution, the \(z\) -score is the inverse of the probability of insolvency. Specifically, \(z\) indicates the number of standard deviations below the expected value of a bank’s return on assets at which equity is depleted and the bank is insolvent” (Beck, et al., 2010). This suggests that a higher \(z\) -score serves as an indicator of higher financial stability, while trends over the past 10 years have shown that it accurately indicated weaknesses in the banking sector prior to the 2008 financial crisis. I also verify my results using the log transformation of central bank assets to GDP, which serves as a proxy for the stability and strength of domestic financial institutions.

To capture the effect of sectoral changes in the economic landscape of a country, I use variables for the value added, measured as percent annual growth, of the manufacturing, agricultural, services, textiles, and industrial sectors. Increases in the value added of any single sector to a country’s economy could have a lasting impact on a woman’s decision to enter the labor force, since certain sectors tend to employ workers along gender lines.

To address omitted variable bias, I control for supply-side determinants of labor force participation. Education, fertility rates, and cultural factors have been shown to be key indicators of increasing female labor force participation rates in developing countries (Psacharopoulos and Tzannatos, 1989). To address the impacts of education, I use the average years of primary schooling and the average years of secondary schooling within a country from the Barro and Lee education

dataset. I use the fertility rate, measured as the total births per woman, and GDP per capita to account for fertility rates and level of development, respectively. Other measures typically included in analyses of labor force participation rates in developed countries such as tax incentives, availability of childcare, and maternity leave are shown to have little impact on developing countries, since many of these programs simply do not exist (Behrman and Wolfe, 1984). I also use net migration and personal remittances received as a percentage of GDP as controls to account for the inflow and outflow of family members and income earned outside the labor force, both of which may impact a woman’s decision to join the labor force during periods of financial hardship. I also account for religious factors by including an index of the percentage of the population identifying as Muslim within a country, since Muslim women tend not to work outside of the home and participate in the formal economy. I also control for political impacts on female labor force participation rates by including a dichotomous measure for democracy, coded one if the regime is democratic and zero otherwise.

5. Methodology
I begin my analysis with a baseline specification to verify my data matches previous analyses of labor force participation rates in developing countries. I use OLS to regress FLFPR and RLFPR on traditional predictor variables. So that participation rates are linearly related to the explanatory variables, I transform FLFPR and RLFPR into:

$$\log(FLFPR / (100-FLFPR)) \text{ and } \log(RLFPR / (100-RLFPR))$$

The baseline specification is as follows:

$$\log(RLFPR / (1-RLFPR)) = \beta_1(\log \text{gdp}) + \beta_2[(\log \text{gdp})]^2 + \beta_k X + \epsilon$$

According to previous literature on the supply-side determinants of FLFPR, I expect to see a negative coefficient for the logarithm of GDP and a positive coefficient for the squared GDP term. I also expect positive coefficients for democracy and education, since cultures with strong democratic regimes and higher levels of education are likely to support women working outside of the home. In contrast, higher fertility rates, a stronger prevalence of Islam, greater immigration to OECD countries, and higher flows of remittances are all likely to decrease FLFPR, so I expect these coefficients to be negative.

While previous literature on labor force participation rates in developing countries focuses on supply side factors and levels of overall development, these analyses miss demand side factors, a key part of labor market decisions. To address this gap, I use a second specification estimating the effect of instability and growth in the value-added of individual sectors of the economy on labor force participation rates:

$$\log(RLFPR / (1-RLFPR)) = \beta_1 \text{Bank Zscore} + \beta_2 \text{Capital Volatility} + \beta_3 \text{Trade} + \beta_4 \text{Textiles} + \beta_5 \text{Industry} + \beta_6 \text{Agriculture} + \beta_7 \text{Services} + \beta_k X + \epsilon$$

Higher levels of stability, measured through a higher Bank Z-score, are likely to decrease female labor force participation rates, as women leave the labor market for socioeconomic status, since their husbands can provide for the entire family without the need of a second income. In contrast, countries with higher levels of financial instability will experience increases in FLFPR, as women are hired by firms as cheaper sources of labor to cut costs and reduce investment in riskier countries. Increases in the value added of certain sectors are also likely to increase or decrease female labor force participation. Gains in the textiles and services sectors are likely to increase female labor force participation rates, as women traditionally enter these sectors to balance domestic and work responsibilities. A decline in the share of agriculture in the economy should indicate falling female labor force participation, since women are traditionally employed in the agricultural sector while men work in manufacturing and industry. Consequently, an increase in the value added of the manufacturing and industrial sectors could have ambiguous impacts. These sectors may move from male to female

5 Dataset courtesy of Professor Vreeland

165
employment as employers look to cheaper sources of labor, which would increase female labor force participation, or they may continue to employ males, supporting the breadwinner and cultural attitudes hypotheses, which could reduce the level of female labor force participation. I expect that the ratio of female to male labor force participation will be crucial in understanding this impact.

To test my main hypothesis, I introduce interaction terms into my model, interacting key economic and instability measures with a crisis dummy to test whether the trends in FLFPR during times of prosperity shift during financial crises, which are periods of decreased investment and high levels of uncertainty in businesses and families alike. I follow the specification:

$$\log(\frac{RLFPR}{1 - RLFPR})_t = \beta_1 [\text{Bank Zscore}]_{t-1} + \beta_2 [\text{Bank Zscore} \times \text{crisis}]_t + \beta_k X_t + \epsilon$$

Where Bank Z-score is the lagged term for instability to measure prior levels of instability, the interaction term measures the effect of instability when there is a crisis, and X is a vector of controls from the baseline specification. I predict that higher levels of stability during non-crisis periods will decrease female labor force participation, in line with previous hypotheses about FLFPR in more financially stable countries. However, I expect that the instability term will reverse signs during financial crises, as women enter the labor market to supplement family income as men lose their jobs. Additionally, firms may demand more women during financial crises, as they look to cut production costs to cope with reduced demand in developed country markets. These dual effects should lead to greater impacts on the labor force participation rates of females, so I do not expect to see significant differences using the ratio of female to male labor force participation rates, since men are likely to continue to remain in the labor market, either working in their job or looking for work to help their families weather the economic downturn.

To address endogeneity concerns about the reverse-causal effect of fertility levels on female labor force participation, I introduce instrumental variables into my regressions. I use total health expenditure per capita and infant mortality rates as instruments for fertility. Health spending is less likely to impact a woman's decision to work outside the home, but is highly correlated with fertility rates in the developing countries in my analysis. Likewise, infant mortality rates are also highly correlated with fertility rates, but less strongly correlated with FLFPR. I use two instruments to account for possible weaknesses in either one, and employ robustness checks to ensure validity in my instruments.

6. Results

Beginning with the baseline specifications, I find my data consistent with previous literature on supply-side determinants of FLFPR. Table 1 shows that the coefficients for log GDP, the squared term of log GDP, democracy, Islam, immigration, remittances, and education dummy variables display expected signs and significance at the 99% level. Across specifications, including instrumenting fertility, the coefficients for the GDP terms and fertility maintain 99% significance. I find that a 1% increase in GDP reduces RLFPR by approximately 7.6% while the positive term on the squared value of the log of GDP indicates a nonlinear effect of GDP on RLFPR. The u-shape to the effect of GDP growth on RLFPR gives weight to the Feminization U-hypothesis: that as a country develops, women reduce their labor force participation rate until a threshold level at which point they enter again. The relatively small coefficient for the squared term of GDP suggests that not as many women enter the labor force in more prosperous countries, which supports previous literature finding an uneven impact in the socioeconomic status of working women in developing countries. While these results give some indication of the drivers of FLFPR, they clearly do not provide the whole picture.

To test the strength of the U-hypothesis given sectoral changes and levels of instability, I include the squared term of the log of GDP to analyze the impact on RLFPR in Table 2. Across both measures of labor force participation rates, higher levels of stability leads to reductions in female labor force participation rates, even when controlling for nonlinearities in GDP, lending support for the hypothesis that women do not enter the labor force in more stable countries since high levels of stability and development offer many opportunities for men to work and reduces the need for women to work outside the home. Though many of the sectors remain statistically insignificant with small impacts on RLFPR, the coefficient for an increase in the value added of the textile industry loses significance which suggests high levels of variation across countries when controlling for nonlinearities in development. Perhaps most interestingly, when controlling for changes in sectoral
emphasis in countries and the role levels of domestic financial stability play in determining RLFPR, the orientation of the U-shape for the log of GDP changes. In the baseline regressions, GDP followed a normal U-shape pattern with a negative coefficient on the linear term and a positive coefficient on the squared term. In contrast, Table 2 shows an inverted U-shape, with a positive coefficient on log of GDP and a negative coefficient on the squared term, both statistically significant at the 99% level. These findings indicate that levels of development play a significant role in determining labor force participation rates and that the debate in the literature is far from over.

Even in accounting for the effect stability plays in supply and demand drivers of female labor force participation, the previous models in my analysis have not accounted for possible changes in FLFPR as a result of financial crises. In a final series of regressions, I propose answers to my main hypothesis, that domestic financial stability plays a role in changing the direction of FLFPR during financial crises. Using instrumental variables regressions with country fixed effects, Table 3 shows that financial crises do shift the effect of stability from negative to positive, with results statistically significant at the 99% level. A history of strong levels of stability, noted by a higher lagged bank z-score results in declines in the female labor force participation rates when the economy is strong. But during financial crises, increases in the financial stability of a country result in increases in the FLFPR. This shift holds across all specifications and for the ratio of female to male labor force participation, suggesting that financial stability especially impacts women over men in developing countries. Since women work in domestically owned industries which require access to domestic and foreign lines of credit to maintain operations during crises, higher levels of stability signal safer investments and easier access to credit, which improves the likelihood that foreign investments will enter the country and that domestic lending institutions will have the ability to extend lines of credit even through economic downturns. Families may need women to enter the market to bolster family income during financial crises and in more stable countries, firms seeking to maintain production levels at reduced costs will demand female labor. While non-crisis trends suggest that more stable countries follow the Feminization U-hypothesis that women leave the workforce as a status symbol of economic prosperity, during financial crises women face greater economic needs from within the family and incentives from labor demand channels to work outside the home. Foreign Direct Investment inflows are insignificant throughout stable economic conditions and financial crises, as noted by the interaction of FDI on the crisis dummy variable. This result is entirely in line with the literature suggesting that Foreign Direct Investment has little impact on the labor force decisions of women, from either the supply or demand side. Fertility is still highly significant, with a one child per woman increase in the fertility rate reduces female labor force participation by an estimated 35%, significant at the 99% level. Accounting for the possibility of the U-hypothesis holding, I find that both the linear term and the squared term for log of GDP follow the originally hypothesized signs and are highly significant. These results suggest that the overall wealth of a country has significant implications for the labor decisions of women, although higher levels of stability during financial crises do play a supporting role.

7. Conclusion
Analyzing female labor force participation rates in 96 countries in the period 1990-2011, I find some support for my hypothesis, that higher levels of stability reduce labor force participation rates during times of economic stability but increase rates in financial crises. My results give weight to existing literature on the demand side determinants of female labor force participation; that foreign and domestic lending institutions are more likely to continue working with domestic firms in more stable countries, which allows these domestic firms to find cheaper sources of labor to cut production costs during economic downturns. These results also confirm existing evidence that women often enter the workforce to supplement family income when male members of the family lose their jobs or take cuts in pay. I find mixed evidence for both the Feminization U-Hypothesis and the Sectoral Changes hypothesis when accounting for levels of instability in a country’s domestic financial sector. The discrepancies in my results suggest that more than simple increases in wealth or the changing nature of work explain rises and falls in female labor force participation in middle income countries. As emerging markets and rapidly industrializing countries enter the global market, it seems that domestic financial stability plays a key role not only in the economic health of a country, but also in the composition of its labor force.
### Appendix

**Table 1.** Baseline specification of the demand side determinants for the ratio of female to male labor force participation rates

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{log}<em>\text{std}</em>{\text{RLFPR}}$</td>
<td>-7.612***</td>
<td>-7.273***</td>
<td>-2.384***</td>
<td>-3.528***</td>
</tr>
<tr>
<td></td>
<td>(0.547)</td>
<td>(0.582)</td>
<td>(0.371)</td>
<td>(0.431)</td>
</tr>
<tr>
<td>$\text{log}<em>\text{std}</em>{\text{RLFP}}$</td>
<td>0.401***</td>
<td>0.383***</td>
<td>0.138***</td>
<td>0.199***</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.033)</td>
<td>(0.021)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>$\text{log}_{\text{gdp}}$</td>
<td>-7.612***</td>
<td>-7.273***</td>
<td>-2.384***</td>
<td>-3.528***</td>
</tr>
<tr>
<td></td>
<td>(0.547)</td>
<td>(0.582)</td>
<td>(0.371)</td>
<td>(0.431)</td>
</tr>
<tr>
<td>$\text{log}_{\text{gdp}}^{\text{sq}}$</td>
<td>0.401***</td>
<td>0.383***</td>
<td>0.138***</td>
<td>0.199***</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.033)</td>
<td>(0.021)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>Fertility</td>
<td>-0.215***</td>
<td>-0.197***</td>
<td>-0.197***</td>
<td>-0.349***</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.026)</td>
<td>(0.018)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Remittances</td>
<td>-0.0161***</td>
<td>-0.0147***</td>
<td>0.00667***</td>
<td>0.00537***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Muslim</td>
<td>-1.484***</td>
<td>-1.428***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.084)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immigrant</td>
<td>-1.37e-06***</td>
<td>-1.31e-06***</td>
<td>-8.59e-08</td>
<td>-2.69e-07</td>
</tr>
<tr>
<td></td>
<td>(3.25e-07)</td>
<td>(3.10e-07)</td>
<td>(1.67e-07)</td>
<td>(6.28e-07)</td>
</tr>
<tr>
<td>Democracy</td>
<td>-0.439***</td>
<td>-0.414***</td>
<td>0.0283</td>
<td>0.0226</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.059)</td>
<td>(0.025)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>School_4_to_7</td>
<td>0.0979</td>
<td>0.0384</td>
<td>-0.123***</td>
<td>-0.0343</td>
</tr>
<tr>
<td></td>
<td>(0.102)</td>
<td>(0.110)</td>
<td>(0.037)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>School_above_7</td>
<td>0.465***</td>
<td>0.395***</td>
<td>0.0202</td>
<td>-0.00428</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
<td>(0.110)</td>
<td>(0.048)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>Constant</td>
<td>37.83***</td>
<td>36.22***</td>
<td>11.46***</td>
<td>17.16***</td>
</tr>
<tr>
<td></td>
<td>(2.366)</td>
<td>(2.547)</td>
<td>(1.661)</td>
<td>(2.037)</td>
</tr>
<tr>
<td>IV Regression</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1

**Table 2.** Sectoral and Instability demand side drivers of RLFPR controlling for possible nonlinearities in GDP. Baseline control variables were included in the regression but left out of the table for clarity

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{log}<em>\text{std}</em>{\text{RLFPR}}$</td>
<td>-0.0288</td>
<td>-0.0352</td>
<td>-0.232***</td>
</tr>
<tr>
<td></td>
<td>(0.094)</td>
<td>(0.094)</td>
<td>(0.075)</td>
</tr>
<tr>
<td>$\text{log}<em>\text{std}</em>{\text{RLFPR}_\text{lead1}}$</td>
<td>-0.0414***</td>
<td>-0.0416***</td>
<td>-0.000758</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Bank_zscore</td>
<td>0.446***</td>
<td>0.435***</td>
<td>-0.115</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
<td>(0.102)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>$\text{log}_\text{trade}$</td>
<td>0.0177</td>
<td>0.0160</td>
<td>-0.0185</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.062)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>$\text{log}_\text{industry}$</td>
<td>0.394*</td>
<td>0.359</td>
<td>0.264***</td>
</tr>
<tr>
<td></td>
<td>(0.232)</td>
<td>(0.230)</td>
<td>(0.098)</td>
</tr>
<tr>
<td>VARIABLES</td>
<td>(1) log_std_FLFPR</td>
<td>(2) log_std_FLFPR</td>
<td>(3) log_std_RLFPR</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Fertility</td>
<td>-0.350***</td>
<td>-0.349***</td>
<td>-0.517***</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.059)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>Bank_zscore_lag1</td>
<td>-0.00324***</td>
<td>-0.00325***</td>
<td>-0.00568***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Bank_zscore_crisis</td>
<td>0.00038*</td>
<td>0.00035*</td>
<td>0.00372*</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td></td>
<td>(0.415)</td>
<td>(0.416)</td>
<td>(0.500)</td>
</tr>
<tr>
<td>Log_gdp_sq</td>
<td>0.185***</td>
<td>0.185***</td>
<td>0.212***</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.023)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>FDI</td>
<td>-0.000230</td>
<td>-0.000179</td>
<td>-0.000899</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>FDI_crisis</td>
<td>-</td>
<td>-0.000267</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>16.04***</td>
<td>16.05***</td>
<td>17.90***</td>
</tr>
<tr>
<td></td>
<td>(1.928)</td>
<td>(1.932)</td>
<td>(2.326)</td>
</tr>
<tr>
<td>IV Regression</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>
</tr>
</tbody>
</table>

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

**Table 3.** The impact of financial crises and instability on both FLFPR and RLFPR. Baseline controls were included in the regression but are not reported for the sake of clarity.
References


Thévenon, O. 2013. "Drivers of female labour force participation in the OECD."
Voting for Credibility:
The Political Economy of Sovereign Credit Ratings

TY GREENBERG
Georgetown University

ABSTRACT
This paper focuses on the determinants of sovereign credit ratings, which are assessments of a government’s willingness and ability to repay its debt. I argue that countries that correspond to American foreign policy interests receive reputational capital in global debt markets. This in turn increases investors’ perception of creditworthiness and leads to higher sovereign credit ratings, all else equal. My empirical results demonstrate that, conditional upon democratic status, countries that vote with the United States in the United Nations General Assembly are assigned more favorable credit ratings. This finding is robust to a number of alternative empirical specifications, control variables, and the inclusion of time and entity fixed effects.
“Comrade, Have You Seen the Latest Polls?”
A Quantitative Analysis of the Impact of National Election Polls on Vote Defection in the European Parliament

MARCEL SCHLEPPER
University of Warwick

ABSTRACT
Previous research disregards that election polls are a major element of media coverage of politics and that they are continuously available to policymakers. By providing a theoretical framework and an empirical analysis of the distortive effects of election polls on parliamentary voting in the European Parliament, this paper extends the current knowledge in a weakly developed field of political economy. This paper concludes that a deterioration in election polls increases the probability of defection from the European party. Further, it is shown that anti-European parties react stronger to polls when national elections are approaching and that pro-European parties react less to polls when European elections are closer.

---

1 This paper has greatly benefited from the excellent research supervision by Professor Vera E. Tröger, who has never stopped asking the difficult questions, which have been necessary in order to develop the research idea further. Thank you for your time and your effort. The author wants to thank everyone who has contributed to this research project, namely: for the provision of data (Professor Simon Hix, Professor Philip Manow, Dr. Holger Döring, Institut für Demoskopie Allensbach, Ipsos, Kantar TNS, Institute of Sociology of the Czech Academy of Sciences), for guiding the implementation in Stata (Professor Jeremy Smith and Gytautas Karklius), and for overall suggestions (Dr. Gianna Boero and Marc Walter).
1. Introduction
This paper investigates the impact of election polls on parliamentary voting in the European Parliament (EP). Because of the dual party membership of legislators, the EP provides a unique possibility to assess the effect of election polls on policymakers. This paper argues that adverse performance in polls increases the probability of vote defection from the European party.

Although election polls reach back to U.S. elections of 1824 (Smith, 1990), research has exclusively analysed the accuracy of these estimates and their impact on voters. This narrow focus disregards that election polls are a major element of media coverage of politics and are continuously available to policymakers (Patterson, 2005). By providing a theoretical framework and an empirical analysis of the impact of election polls on policymakers, this paper extends the current knowledge in a weakly developed field of political economy.

In recent months, pollsters have been criticised for false predictions of major political events. At the same time, the existence of an effect of polls on voters is widely acknowledged, causing the publication of polls to be partially banned in many countries before election days (Walsh et al., 2009). This paper goes further and assesses the negative externalities of election polls on policymakers’ behaviour. It raises the question whether the dominance of polls in the media promotes the implementation of national rather than European policies in the EP.

First, this paper lays the theoretical foundation by guiding the reader through the historical development of the analysis of parliamentary voting. Thereafter, the theoretical models of voting behaviour within the EP are extended and hypotheses are derived. The next section introduces the data and outlines the empirical model employed to investigate the importance of polls. Finally, the results are presented and fields for future research are outlined.

2. Literature review
2.1 Do parties discipline their representatives?
Initially, two theoretically derived hypotheses about the source of party cohesion (i.e. similar voting within parliamentary groups) competed. While Cox and McCubbins (1993) claim that the party leadership influences the behaviour of representatives, Krehbiel (1993) suggests that party cohesion is only a reflection of similar preferences within a party. The analysis of roll-call votes (i.e. votes in which each representative's behaviour is reported for) has provided the possibility to critically assess these ideas.

Hager and Talbert (2000) have argued that representatives who switch party affiliation (e.g. former Democrats joining the Republicans) also change their voting behaviour due to leadership pressure. In addition, in key votes (i.e. procedural, organisational, and label-defining), party pressure causes 20 to 40 representatives to differ from their expected position (Cox and Poole, 2002). Hence, representatives in Congress frequently follow the party line in spite of their preferences.

2.2 Which party enforces cohesion in the European Parliament?
Members of the European Parliament (MEP) are organised in European parties. Even though European party affiliation is chosen freely with the aim to minimise policy incongruence (McElroy and Benoit, 2010), dual party membership requires a new theoretical model explaining MEP behaviour. The literature is dominated by the idea that a MEP is an agent to two principals. Both principles desire high cohesion and sanction deviating voting behaviour (Gabel and Hix, 2002).

In the *two-principal model* the possibility to sanction arises from the fact that the European parties determine policy implementation and office allocation within the EP, while the national parties control re-election (Hix, 2004).

2.3 What drives vote defection in the European Parliament?
Even though high levels of party cohesion are observed, occasionally individual MEPs deviate from group positions. In order to study these phenomena, the previous theoretical model is complemented by a second one. It is built on the idea that the European party leadership does not possess strong enough means to punish deviating behaviour as offices are distributed by national proportionality (Ringe, 2009).
In this light, the perceived-preference-cohesion model describes European party cohesion as a bottom-up process. Ringe (2009) proposes that policy positions are not exogenous but emerge from negotiating legislation in committees. As MEPs outside the respective committee are uncertain which legislative choices satisfy their own ideological preferences, they trust experts (i.e. fellow party members who have led the negotiations). The model is extended by Roger and Winzen (2014) to provide an intuition for the development of conflicts (i.e. situations in which conflicting vote instructions are received from the national and European party). When facing domestic pressure or anticipating intra-party conflict in the plenary meeting, national parties decide upon their position in advance of committee negotiations which may lead to incompatibility of standpoints and consequently cause vote defection from either party.

3. Theory and hypotheses
The hypotheses, which will be presented hereafter, are based on a binary choice model which regresses vote defection from the European party for individual ‘i’ and vote ‘j’ on polls, defined as the difference between current polls and previous election outcomes.

\[
\text{Prob}(\text{EP defection})_{ij} = G(\alpha + \beta_{\text{polls}})
\]

3.1 Effect of polls on probability of defection from European party
The perceived-preference-cohesion model describes the emergence of conflicts endogenously. Vote cohesion is high when experts give similar voting advice and when non-experts follow these suggestions. Once election polls are integrated into the model, their impact on the development of conflicts can be observed. If the national party experiences diminishing polls, its MEPs have greater incentives to express policy positions that reflect domestic political considerations in order to satisfy the domestic electorate. Ultimately, this increases the number of conflicts as MEPs may either prevent the development of consensus positions in the committees or defect from these positions in parliamentary voting.

The two-principal model focuses on the behaviour of an MEP once he finds himself in a conflict caused exogenously. The decision to deviate from leadership’s instructions will be based on the ability of the principals (i.e. the party’s leadership) to induce costs on the MEP. Once election polls are integrated into the model, their impact on defection costs can be observed. If the polls deteriorate for the national party, its representatives anticipate that fewer government offices and parliamentary seats will be available in future. This makes vote defection from the national party less likely as it can induce greater costs.

The theoretical framework is based on the idea that vote defections in the EP are – to some extent – employed to address the domestic electorate. This is in line with previous research by Slapin and Proksch (2010) who have empirically shown that defecting MEPs demand more speaking time in order to publically signal their support for the national party in the respective vote.

Table 1. Effect of election polls on defection probability

<table>
<thead>
<tr>
<th></th>
<th>European party</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Perceived-preference-cohesion</td>
<td>(\beta_{\text{ppc}} &lt; 0)</td>
</tr>
<tr>
<td>2. Two-principal</td>
<td>(\beta_{\text{tp}} &lt; 0)</td>
</tr>
<tr>
<td>Total</td>
<td>(\beta = \beta_{\text{ppc}} + \beta_{\text{tp}} &lt; 0)</td>
</tr>
</tbody>
</table>

Table 1 shows the hypothesis that a deterioration of national poll performance leads to a rise in the probability of vote defection from the European party. The total effect consists of the impact within both theoretic models. In the perceived-preference-cohesion model, a deterioration in poll performance increases the number of conflicts (\(\beta_{\text{ppc}} < 0\)) and in the two-principal model the European party is less able to discipline its members (\(\beta_{\text{tp}} < 0\)).

**Hypothesis 1:** A deterioration in national election polls increases the probability of vote defection from the European party.
3.2 Effect of polls for pro- and anti-European parties

The initial hypothesis treats all parties identically. In reality, ideological and institutional differences exist along the lines of pro-European parties (ALDE, EPP, Greens-EFA and S&D) on one side and anti-European parties (ECR, EFD, and GUE-NGL) on the other side. These dissimilarities manifest themselves in contrary political objectives. The relationship between left-right preferences and the sympathy towards Europe is frequently described as an inverted U-shape, where parties at the boundaries (i.e. ECR, EFD, and GUE-NGL) strongly oppose the European Union (McElroy and Benoit, 2010). Less visible, but not less important, is the different ability of the two groups to implement policies on the European level. Various European institutions (namely the Commission, Council, and Parliament) interact throughout the legislative procedure until a law is adopted. However, all three are dominated by (or consist entirely of) individuals who belong to parties which are classified as pro-European (Corbett et al., 2014).

Figure 1. Share of votes on winning side (Source: Hix)

The institutional dominance of pro-European parties goes along with higher importance in the process of policy implementation because ALDE, EPP, and S&D frequently negotiate compromises before voting on the floor (Bressanelli et al. 2015; Bowler and McElroy, 2015). Figure 1 shows that pro-European parties support a majority of the adopted policy proposals. For example, in only 10% of the cases, a proposal is implemented which has not been supported by the EPP. As anti-European parties have a lower ability to implement policies, the establishment of high party cohesion is less important in comparison to the policy-driving pro-European parties. Further, the latter may experience pressure from national governments to vote cohesively in order to support proposals which have achieved agreement in the European Council (Faas, 2003). In fact, defection rates in the sample are, on average, 20% for anti-European parties and hence are four times higher than for their pro-European colleagues. Thus, an analysis of the voting behaviour in the EP should account for the different roles of these two groups.

Hypothesis 2: Pro- and anti-European parties react differently to a deterioration in national election polls, which is due to their ideological and institutional distinctness.

3.3 Variation of the poll effect in dependency on time until election

The mechanisms, which lead to the responsiveness of political behaviour to polls, are based on individual career concerns and the collective desire of a political party to increase its vote share and political influence. While these aims exist throughout the legislative cycle, the need to react to poor performance increases when elections approach. This does not only reduce the likelihood that polls will turn around in the remaining time, but also diminishes the discount placed on a potential realisation of the concerns (i.e. losing office and power). Thus, it is predicted that the response to polls is conditional on the time until the next election. However, the direction of this effect depends on the weight each party group attributes to the respective election and the preferences of its electoral base. This is represented in Table 2.
Table 2. Impact of time until election on poll-effect

<table>
<thead>
<tr>
<th>Reaction on elections</th>
<th>National</th>
<th>European</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Anti-European</td>
<td>$\beta_{\text{anti,nat}}$ increases</td>
<td>$\beta_{\text{anti,Euro}}$ decreases</td>
</tr>
<tr>
<td>2. Pro-European</td>
<td>$\beta_{\text{pro,nat}}$ increases</td>
<td>$\beta_{\text{pro,Euro}}$ decreases</td>
</tr>
</tbody>
</table>

Both pro- and anti-European parties have a clear focus on national elections. The EP offers a platform to strengthen the profile for these. Further, Meserve et al. (2009) outline that it is important to please the national party in the run-up to national elections for those MEPs who use the EP as a stepping-stone for domestic offices. Thus, the effect of polls increases when national elections approach ($\beta_{\text{anti,nat}}$ increases) and ($\beta_{\text{pro,nat}}$ increases)

Supporters of pro-European parties reward cooperative and constructive behaviour in European elections. Furthermore, MEPs are not only concerned about their re-election, but also about their ability to shape policies once offices are allocated after the European election. As the European parties have significant power over the allocation (Trumm, 2015) individual MEPs optimise their future prospects by responding less to polls. Thus, the effect of polls diminishes for pro-European parties when European elections come closer ($\beta_{\text{pro,Euro}}$ decreases). However, European elections are not a priority for anti-European parties, so no significant impact on the poll effect is expected ($\beta_{\text{anti,Euro}}$ decreases).

Hypothesis 3: The poll effect on vote defection increases when national elections approach (both party types) and decreases when European elections approach (pro-European parties).

4. Data and methodology
4.1 Data on individual voting behaviour

The data on voting in the 7th European Parliament (2009-2014), including additional information on MEPs, has been provided by Simon Hix. It contains voting records for 853 MEPs in 6,963 votes. While the MEPs belong to 206 different national parties, they have organised themselves within seven European parties. Both the individual voting records and the European party membership have been employed to generate the dependent variable. It measures whether an individual MEP defects from the majority position of his European party in a specific vote.

While the dataset includes 5.8 million individual voting decisions, these do not represent the entire population of votes in the EP as only 20% to 30% of the voting is conducted by roll-call (Thierse, 2016). The other techniques (i.e. show of hands and electronic voting) do not record each individual's decision. They can neither be included in the analysis nor is the public informed about the voting decision. Thus, incentives and voting behaviour for unobserved votes is expected to differ and conclusions cannot be extended to votes that are not recorded.

Even though the 5.8 million voting decisions are only a subset of the entire population, they comprise different types of votes within themselves. Table 4 presents the share of each vote type that is (1) final vs. non-final and (2) legislative vs. non-legislative. The procedures of the EP require that each final legislative vote is recorded. In the dataset, they have a share of 13.75%. For votes which are either non-final or non-legislative, roll-call votes (RCV) have to be requested by a group of at least 40 MEPs (Yordanova and Mühleböck, 2014).

<table>
<thead>
<tr>
<th>Vote type</th>
<th>Final</th>
<th>Non-final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legislative</td>
<td>13.75%</td>
<td>23.48%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>37.23%</td>
</tr>
<tr>
<td>Non-legislative</td>
<td>10.41%</td>
<td>52.35%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>62.76%</td>
</tr>
<tr>
<td></td>
<td>24.16%</td>
<td>75.83%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100%</td>
</tr>
</tbody>
</table>

This is problematic as Carrubba et al. (2008) show that the request of RCVs is non-random, e.g. they do not proportionally represent committees, suffer group bias and over-represent resolutions. Furthermore, RCVs are strategically requested in order to obtain disciplined voting behaviour (Carrubba et al., 2008), to reveal disunity in competing European parties (Thierse, 2016) and/or to signal a position to voters (Faas, 2003).
Figure 2 illustrates the strategic and non-random nature by comparing *average vote defection rates* in final legislative votes (which are unbiased as the entire population is recorded) with (1) legislative votes which are non-final and (2) final votes which are non-legislative. For pro-European parties, significantly higher defection rates are observed for those vote types which are not representative of their respective population. This may be due to the specific nature of the vote, but it is more likely that these differences are driven by the bias arising from the fact that RCVs have to be requested. Thus, the analysis will only include final legislative votes and draw conclusions for this group. This reduces the sample to 807,000 votes, but it allows the estimation of unbiased results.

Figure 2. Average defection rates by European parties (Source: own calculations)

4.2 Data on party performance
The data on national election polls has been provided by several polling institutes. It has been merged with data on election results collected by Manow and Döring (2016) in order to generate the variable of interest *polls*. It is defined as the difference between current national election polls and the previous election outcome. Hence, negative values signal that a party is performing poorly.

The data has been collected for six member states. As motivations among MEPs (e.g. re-election and policy-implementation) do not vary substantially between countries, the results of this paper are expected to also hold for an analysis of the entire EP. This is reinforced by investigating the sample more closely. It shows that even though the sample consists of 6 of the 27 member states, it replicates the distribution of parties in the EP well. The sample includes between two and eight national parties for each European party, which represent 17%-22% of the total national party number in each European party. As countries such as Germany, Poland, and the United Kingdom have a high share of MEPs, this translates into 29%-42% of the total MEP number in each European party being included in the dataset. Hence, the sample does not over-represent certain European parties.

4.3 Model
In order to test empirically whether national election polls have an impact on vote defection, a binary response model is employed. The dependent variable is the *probability of vote defection* from the European party for vote “i” and individual “j”. The variable of interest is *polls*.

\[ \text{Prob}(\text{EP defection})_i = G(z) = G[\alpha + \Omega_i \text{NEG}_i \text{POLLS} + \Phi_i \text{POS}_i \text{POLLS} + \Theta_i \text{CONTROLS}] \]

The vector of interest includes the variable *negative_polls* individually and as an interaction with the time until the next European and national elections. Furthermore, it differentiates between pro-European and anti-European parties. A similarly defined vector is included for positive polls.
\[ \Omega_i \text{NEG}_P_O_L_L_S = \sum_{\text{pro}=0}^{1} \left[ \beta_{1,\text{pro}} \text{negative_poll} \times \text{pro}_i + \beta_{2,\text{pro}} \text{negative_poll} \times \text{time}_\text{national_election} \times \text{pro}_i + \beta_{3,\text{pro}} \text{negative_poll} \times \text{time}_\text{European_election} \times \text{pro}_i \right] \]

Further, the vector of control variables includes: (1) time until next European election, (2) time until next national election, (3) European party membership, (4) nationality, and (5) topic.

The result section will present the marginal effects from a random effects logit regression. A probit model has equal levels of significance. The random effect model is selected because the control variables account for the impact of European party membership and nationality on parliamentary voting. Thus, the remaining errors are randomly distributed on MEPs.

5. Results
5.1 Impact of polls on vote defection from the European party
Figure 4 shows that the marginal effect of polls on vote defection from the European group is negative. For pro-European parties, a 1pp. increase in polls reduces the probability of defection by 3pp. For anti-European parties, the effect is slightly lower, but not significantly different. Thus, it is concluded that the first hypothesis cannot be empirically rejected. When a party, which is performing below its previous election result, experiences a further decrease in election polls, then its MEPs react by defecting more frequently from the European party in parliamentary voting.

|               | dy/dx     | Std. Error | z   | P > |z| |
|---------------|-----------|------------|-----|-----|---|
| Pro-European  | -0.03059  | 0.005865   | -5.22| 0.000|
| Anti-European | -0.02690  | 0.008877   | -3.03| 0.002|

5.2 Effect of polls in dependency on national elections
Figure 5 shows how the size of the poll effect varies depending on the time until the next national election. At the start of the legislative cycle anti-European parties do not respond to fluctuations in polls by altering their voting behaviour. However, during the last months before national elections, they react to a 1pp increase in polls by reducing the probability of defection by 6pp. In contrast, pro-European parties experience a constant effect of polls on vote defection over the legislative cycle.

Figure 5. Marginal poll-effect for national elections (Source: own calculations)
Thus, the hypothesis that MEPs react stronger on polls when national elections approach is not rejected for anti-European parties and is rejected for pro-European parties. The reason for the differences in behaviour is the following. For both parties, the EP offers a platform to pursue policies, which increase their popularity in the domestic arena. However, misusing voting behaviour for such purposes requires the willingness and ability to do so. For anti-European parties the first is made possible by the lack of a European identity and the second is achieved through the direct control of their national party leadership over MEPs. Their key figures are in various instances themselves members of the EP (e.g. Bernd Lucke (until 2015), Marine Le Pen, and Nigel Farage).

6. Conclusion
This paper has developed a theoretical framework to explain the impact of election polls on voting behaviour within the European Parliament. The perceived-preference-cohesion model suggests that polls affect the number of conflicts between the European and national party. In addition, the two-principal model proposes that polls alter the loyalty of MEPs towards the two parties. Both models predict that adverse performance in election polls increases the probability of vote defection from the European party.

This paper has empirically investigated the hypothesis by analysing the voting behaviour within 180,000 final legislative votes. This paper has been the first to establish a link between election polls and parliamentary voting. In addition, it has been shown that anti-European parties react stronger to polls when national elections approach and that pro-European parties react less to polls when European elections are closer. Finally, the paper has outlined that the effect of election polls is non-linear and that the mechanisms, which link polls and voting behaviour, are weaker when parties perform strongly.

This paper has demonstrated the distortive effects of polls on policymaking as they lead to the implementation of national rather than European policies in the EP. Thus, extensive media coverage of polls has negative externalities. In the European context, undesired effects arise such as the prioritisation of domestic interests, which puts the implementation of truly European policies at risk. The research should be extended on national parliaments as a higher importance is attached to these meaning the effect of polls is expected to be even larger.

It has been out of the scope of this paper to process polling data for all national parties in the 7th European Parliament. Moreover, future research should attempt to identify national political shocks, such as scandals, and employ these as instruments in order to deal with the endogeneity issue of election polls. Finally, the conclusions have been restricted to final legislative votes. An extension on all types of votes should be approached once the selection bias is more explored.
References
COUNTY-LEVEL CRIME RATES AND SANCTUARY JURISDICTIONAL STATUS

THOMAS M. SEGERSTROM

Gettysburg College

ABSTRACT
Some policy analysts suggest that sanctuary jurisdictions in the United States lower local crime rates because sanctuaries promote higher levels of trust between Hispanic populations and police. However, little research has been done on the relationship between crime rates and sanctuary status. In this paper, I employ a rational-crime model from Becker (1968) in order to econometrically estimate the effect of sanctuary county status on crime rates for 495 counties in the United States in 2015. Consistent with prior research county level research into crime rates, I find that when accounting for simultaneity of police expenditure, there is no statistically significant difference in crime rates between sanctuaries and their counterparts.

---

1 I wish to thank my Honors Thesis advisor John Cadigan, Ph.D. for his mentorship and extensive feedback through this project. I’d also like to thank Professors Iva Reyes, Brendan Cushing-Daniels, James O’Brian, and the Honors Research Seminar for their suggestions and assistance in the development of this paper. I also wish to thank my parents for their support. Any errors that exist in this research are my own.
1. Introduction

Since the seminal work of Gary Becker (1968), economists and policy analysts have investigated determinants of crime rates (see for example Ehrlich, 1972; Lee et al. 2001; Levitt 1995; Glaeser and Sacerdote 1999). After the events of September 11, 2001 policy analysts have given increasing attention to "sanctuary status" as issues of national security and immigration are more prevalent and contemporary (Jordan and Elinson, 2015; Collingwood et al. 2016). In the United States, a "sanctuary city" is loosely defined as a jurisdiction where law enforcement, as stated by the Los Angeles Police Department, "shall not initiate police action with the objective of discovering the alien status of a person" nor "arrest [or] book persons for violation of Title 8, Section 1325 of the United States Immigration Code (Illegal Entry)” (Gates, 1979). Sanctuaries may lower crime rates if they create increased cooperation with the police (Kittrie, 2006; Theodore, 2013; and others). Alternatively, sanctuary status may increase crime rates if potential criminals face reduced deterrence effects because they do not fear deportation. Available research indicates that areas in the U.S. with high immigrant populations have lower crime rates (Spenkuch, 2014; Reid et al. 2005), and that sanctuary policy may lend itself to lower crime rates (Theodore, 2013). Conversely, sanctuary jurisdictions are believed by some public policy researchers to be a “public safety problem” for releasing jailed aliens back to the public (Vaughn, 2017). This paper adds to empirical literature on crime by evaluating impact of sanctuary status on violent and property crime rates.

Results across the log-log models that treat police expenditure both exogenously and endogenously (using 2SLS), and the negative binomial model which accounts for the count-variable nature of the dataset, indicate that sanctuary status has no statistically significant effect on crime rates.

2. Literature Review

Most research that investigates the relationship between crime rates and variables of interest employ the framework developed by Becker (1968) in evaluating crime rates. The rational crime model primarily points to the importance of elasticities of crime with respect to different variables of interest. Ehrlich (1972) develops Becker's model empirically and uses logarithms for ease of use in estimating the marginal impact of different indicators of crime. Glaeser and Sacerdote (1999), Levitt (1995) are common applications of Becker's seminal work on crime. Glaeser and Sacerdote (1999), and Kelly (2000) incorporate logs of crime rates and logs of explanatory variables to explain variation in crime and estimate elasticities of crime with respect to different demographic and socioeconomic variables. Glaeser and Sacerdote (1999) employ cross-sectional data of several variables to determine the effect of population using logs. Kelly (2000) also incorporates logs in estimating the effects of inequality in counties across the United States, and uses relevant instruments in accounting for potential police expenditure endogeneity (Levitt, 1995).

It is also crucial to consider the existing theories on what may make non-sanctuaries more crime-ridden. In this analysis, Merton's strain theory (Merton, 1938) articulates a causal mechanism for why an area in the US might have higher or lower crime than another (Martinez et al. 2010; Kelly, 2000; Lee et al. 2001). This theory of crime is often cited as it provides a rather intuitive explanation for criminal behavior, and may explain why non-sanctuaries might have more crime in theory. However, it is not the only theory to explain why disadvantaged groups may be associated with higher crime in empirical results as immigrants and other groups may be motivated by economic constraints to live in areas that already are crime ridden (Lambert, 1970; Lee et al. 2001). Shaw and McKay's social disorganization theory (1942) suggests that communities without much of a social structure will lead to higher crime. Macro county-level policy research into the phenomenon often points to interactions between immigration and crime before evaluating anything about the effect of sanctuary policy. Empirical evidence among several studies call into question whether recent immigrants commit more crime than the natively-born in the US (Lee et al. 2001, Ousey & Kubrin 2009 among others). This empirical evidence provides a suitable counter-claim to assertions that immigrant communities are disorganized and thus more crime-ridden, as many are organized in ways where crime is actually discouraged and found to have a negative effect on crime rates (Ousey & Kubrin 2009; Spenkuch 2014). It is also possible that areas with sanctuary policies may have less crime because deportations are not as frequent.

Limited research evaluates sanctuary policy on an empirical basis, however, there are two notable exceptions. First, Hararkin (2015) compares sanctuaries and non-sanctuaries and their respective crime rates, and points out that sanctuaries
have lower crime rates. Collingwood et al. (2016) find that there is no statistically significant difference between sanctuary cities and matched non-sanctuaries from the mid-2000s, even after controlling for relevant covariates. Secondly, Collingwood et al. (2016) runs sensitivity analysis and matching estimations for 55 known sanctuary cities in the U.S. and observes that sanctuaries have no statistically significant impact on crime when compared to a matched city. A limitation of existing research is that there has yet to be a study that tests for an effect of sanctuary on crime using the standard econometric techniques and accounts for the count variable nature of crime statistics. This paper incorporates appropriate controls for different explanations of crime, and the effect of sanctuary cities and fills in a gap in existing literature. Using cross-sectional data and a log-log regression framework in analysis, I find that sanctuary cities have no statistically significant impact on crime rates. This finding is consistent across three different empirical specifications.

3. Theory
In this research, I model crime as a function of several covariates and the sanctuary status of a county. Using the standard theoretical framework from Becker (1968), the log of a county’s crime rate is expressed by several explanatory covariates and its sanctuary status:

\[
\text{crime rate}_i = \beta_0 + \beta_1\text{density}_i + \beta_2\text{inequality}_i + \beta_3\text{immigration}_i + \beta_4\text{race}_i + \beta_5\text{unemployment}_i + \beta_6\text{youth}_i + \beta_7\text{poverty}_i + \beta_8\text{police}_i + \beta_9\text{sanctuary}_i
\]

(1)

Under the rational crime framework, areas with greater population densities offer more anonymity to criminals and opportunities to “hide” among a large populace (Becker 1968). As for inequality, individuals living in highly unequal areas of the U.S. may have a higher perception of their material lack of success, and be more likely to pursue criminal activity. This framework is direct from strain theory and is referred to in some papers (Merton, 1938; Kelly, 2000; Lee et al. 2001). Increases in unemployment may lead to higher crime rates. Unemployment has been often identified as an important control (Cantor and Land, 1988, Freeman, 1991, and Grogger, 1997). Race is included as it directly relates to sociological theories of crime (Grogger, 1997; Ousey and Kubrin 2009). Race is measured as the percentage of the county that is non-white. Several authors contend that younger segments of the population are more likely to commit to criminal activity (Cohen and Land, 1987). Youth is represented as the percentage of the population that is under the age of 25. Poorer areas have less economic means, and thus areas with higher poverty are more likely to be associated with violent and property crime under existing literature on crime.

Factors that would intuitively lead to lower crime rates are prevalence of immigrants, and police deterrence. Several studies indicate that crime rates in areas with high levels of foreign-born persons are associated with less crime. Contrary to much stigma in the USA, immigrants are less likely to commit crime than native-born Americans (Graham and Kubrin, 2009; Miles and Cox, 2016). Police activity is a factor that would lead to lower levels of crime, in expectation, as police presence offer a deterrence to crime. In addition to the existing previously researched deterrence variables, I include sanctuary status, or whether a county has a policy to restrict cooperation with federal authorities regarding immigration law, as a variable. I expect that the coefficient on the sanctuary status variable will be negative, and I also expect that the model will be subject to omitted variable bias if sanctuary status is not included. In expectation, if Theodore’s (2013) research into Hispanic “trust” in sanctuary cities and local law enforcement is well reasoned, jurisdictions that effectively protect illegal immigrants from federal deportation laws will exhibit lower crime rates as the local population will feel empowered to report tips to local authorities. Additionally, illegal aliens may feel more comfortable in their own community and not have as high of a propensity to commit crime, assuming strain theory. On the other hand, if the model expects crime to be higher in sanctuary counties, this would refute such a trust theory, and lend credence to the idea that crime is actually worse in sanctuary cities.

Following existing literature, I first use OLS multiple regression to express the logs of crime rates as functions of the logs of explanatory variables consistent with Ehrlich’s econometric specification. Then, I use a Negative Binomial regression model to account for dependent variables being counts. Ordinary Least Squares (OLS) fails on simply the levels of crime, however when crime rates (i.e. crimes per 10,000 persons) are used as dependent variables and then logged, OLS can be useful in determining elasticities of crime with respect to explanatory variables. Given the nature of the data, and evidence
across multiple papers that endogeneity is present in any police presence variable, I also use instruments suggested by Kelly (2000) and Levitt (1995) to account for simultaneity issues.

Since many of the observations of violent and property crime are reported to be zero or a low count, I also take the opportunity to account for this effect and thus employ a Negative Binomial regression in order to account for the non-negative data and the effect that such data have on classical OLS assumptions.

4. Data
The data used in this project primarily come from Federal government agencies that track demographic and public safety trends around the United States. In particular, this paper makes use of data from the 2015 Uniform Crime Reporting from the FBI, which includes data on law enforcement agencies, their size, and crimes that occur in each jurisdiction. Variables obtained from the FBI include violent crime, and property crime levels. I obtain data on unemployment from the Bureau of Labor Statistics. Various other demographic statistics come from the American Community Survey 2015 estimates.

I obtain the most recently available data regarding police expenditure per capita from the 2012 Census of State and Local Governments from the 2015 County and City Databook. Police expenditure is counted as the dollars spent per person in the county allocated by local budgets for police spending. Use of older police expenditure data has been used empirically (see Kelly, 2000) who used data three years older than demographic data in regressions. To account for simultaneity between police expenditure and crime rates, I employ a Two-Staged Least Squares instrumental variables approach using instruments suggested by Kelly (2000). These instruments used for police expenditure per capita include the non-police spending in dollars as a percentage of aggregate county income, median earnings of the county, as well as the percentage of the county that supported the Democratic ticket in 2012.

Most importantly, data on a jurisdiction’s sanctuary status and deportation levels in 2014-2015 are obtained from the Immigration and Customs Enforcement (ICE) and provided by the Texas Tribune. Sanctuary status is denoted by a binary variable where “1” indicates that a jurisdiction “has a policy or legislation that limits or prohibits cooperation with ICE,” and “0” otherwise (Griffith and Vaughn 2017). While there is no legally agreed-upon definition of what constitutes a sanctuary, the agency reports whether a local jurisdiction declined a detainer in the year 2015. There is potential for endogeneity in the data between sanctuary status and crime rates.

| Table 1 - Summary Statistics for Dependent and Independent Variables |
|--------------------------|-----------------|-----------------|-----------------|
|                         | Mean            | SD              | Minimum         | Maximum         |
| Violent Crime†          | 12.60           | 16.90           | 0               | 124.85          |
| Property Crime†         | 78.13           | 85.79           | 0               | 630.60          |
| Population              | 368,615         | 706,864         | 1335            | 10.2mil         |
| Income Inequality Gini  | 0.45            | 0.033           | 0.37            | 0.57            |
| Density                 | 563.84          | 1172.5          | 0.82            | 10,897.5        |
| Foreign-born (%)        | 10.28           | 8.30            | 0.398           | 51.65           |
| Non-white (%)           | 22.28           | 14.36           | 1.92            | 79.60           |
| Unemployment Rate (%)   | 5.51            | 2.01            | 2.4             | 24              |
| Youth (%)               | 32.82           | 4.28            | 21.20           | 47.093          |
| Poverty (%)             | 14.00           | 5.30            | 3.9             | 36.6            |
| Law Enforcement Spending ($) per capita | 240.49 | 113.84 | 22.87 | 1089.74 |
| Sanctuary Jurisdiction  | 0.318           | 0.466           | 0               | 1               |

† indicates value is calculated per 10,000 people.
5. Regression Results & Interpretation

OLS results of the elasticities of crime and variables of interest, as preliminary tests of the data, are summarized in Table 2. The regressions in Table 2 support prior findings (Collingwood et al. 2016) that sanctuaries are not associated with higher crime rates. There is no statistically significant difference between sanctuary counties and non-sanctuary counties. One finding notable, also noted by Kelly (2000), is that the effect of the percentage of the population that is young, or less than 25 years of age, is negative. It is possible that areas with more families have fewer crimes. Coefficients on police spending per capita are negative yet statistically significant.

| Table 2 - Determinants of Crime Treating Police Spending as Exogenous |
|--------------------------|--------------------------|
|                          | (1)                      | (2)                      |
|                          | Property Crime           | Violent Crime            |
| Police Expenditure       | -0.0335                  | -0.0568                  |
|                         | (0.171)                  | (0.160)                  |
| Sanctuary                | 0.0444                   | 0.0482                   |
|                         | (0.156)                  | (0.144)                  |
| Density                  | -0.0727                  | -0.0754                  |
|                         | (0.0505)                 | (0.0475)                 |
| Non-white (%)            | 0.164                    | 0.186*                   |
|                         | (0.121)                  | (0.112)                  |
| Unemployment (%)         | 0.427                    | 0.496**                  |
|                         | (0.263)                  | (0.245)                  |
| Youth (%)                | -1.861***                | -2.482***                |
|                         | (0.557)                  | (0.523)                  |
| Poverty (%)              | 0.309                    | 0.501**                  |
|                         | (0.232)                  | (0.212)                  |
| Foreign-born (%)         | -0.0565                  | 0.0617                   |
|                         | (0.0997)                 | (0.0947)                 |
| Constant                 | 8.922***                 | 8.434***                 |
|                         | (2.199)                  | (2.095)                  |

N 475 464  
R-squared 0.068 0.111

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1  
Dependent and independent variables (except sanctuary) are in logs.

These regressions are early indicators do not consider potential simultaneity between police expenditure and crime. Perhaps more police spending is a result of higher crime. After testing for endogeneity as outlined in Chapter 15 of Wooldridge (2015), I conclude that police expenditure per capita is in fact endogenously related to crime rates.

Table 3 indicates an IV approach in OLS results with logarithms, however, police expenditure per capita is treated endogenously via a simultaneous system of equations. Kelly (2000) uses the percent of the aggregate county income per which is dedicated to non-police government spending, the percentage share who vote Democratic in the presidential election in 2012, and the median income of the county. Levitt (1995) and Kelly (2000) consider police expenditure to be functions of electoral politics. Furthermore, percentage of the aggregate county income dedicated to non-police government spending indicates how generous a county can be towards government spending, and median income is also used to account for the relative wealth of the local economy.
<table>
<thead>
<tr>
<th></th>
<th>Property Crime</th>
<th>Violent Crime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Police Expenditure</td>
<td>-0.654</td>
<td>-0.401</td>
</tr>
<tr>
<td></td>
<td>(0.483)</td>
<td>(0.460)</td>
</tr>
<tr>
<td>Sanctuary</td>
<td>0.127</td>
<td>0.0895</td>
</tr>
<tr>
<td></td>
<td>(0.168)</td>
<td>(0.153)</td>
</tr>
<tr>
<td>Density</td>
<td>-0.0658</td>
<td>-0.0750</td>
</tr>
<tr>
<td></td>
<td>(0.0510)</td>
<td>(0.0473)</td>
</tr>
<tr>
<td>Non-white (%)</td>
<td>0.213*</td>
<td>0.216*</td>
</tr>
<tr>
<td></td>
<td>(0.127)</td>
<td>(0.118)</td>
</tr>
<tr>
<td>Unemployment (%)</td>
<td>0.378</td>
<td>0.464*</td>
</tr>
<tr>
<td></td>
<td>(0.266)</td>
<td>(0.247)</td>
</tr>
<tr>
<td>Youth (%)</td>
<td>-2.313***</td>
<td>-2.766***</td>
</tr>
<tr>
<td></td>
<td>(0.648)</td>
<td>(0.631)</td>
</tr>
<tr>
<td>Poverty (%)</td>
<td>0.271</td>
<td>0.480**</td>
</tr>
<tr>
<td></td>
<td>(0.235)</td>
<td>(0.213)</td>
</tr>
<tr>
<td>Foreign-born (%)</td>
<td>0.0625</td>
<td>0.132</td>
</tr>
<tr>
<td></td>
<td>(0.132)</td>
<td>(0.129)</td>
</tr>
<tr>
<td>Constant</td>
<td>13.57***</td>
<td>11.15***</td>
</tr>
<tr>
<td></td>
<td>(4.037)</td>
<td>(3.991)</td>
</tr>
</tbody>
</table>

Wu-Hausman Test p-val. 0.1669 0.4273
N 475 464
R-squared 0.042 0.102

Dependent and independent variables (except sanctuary) are in logs.

6. Conclusion
In this paper, I examine the effect of sanctuary counties in the United States and whether or not they have a meaningful impact on crime rates on the local level. The result that sanctuary counties have no effect on crime is noticeable across three different empirical estimation methods, most notably in the examination of elasticity of crime when police expenditure is treated endogenously and exogenously in a semi-log model. Results of a negative binomial regression that treats police expenditure exogenously confirm that the effect of sanctuary counties on crime is not statistically significant from zero. The impact of this paper is intended to address the public policy issue of sanctuary counties in the United States. Knowing that sanctuaries have no meaningful impact on crime confirms the limited empirical research available (Collingwood et al. 2016), and several non-empirical socioeconomic research works on the subject. While some may view sanctuary counties as a public safety problem, the data does not suggest that sanctuaries have any meaningful impact or positive effect on crime. Further research into this topic would involve acquiring a longitudinal dataset on sanctuary counties, or cities and also finding a more appropriate instrumental variables approach, and accounting for the potential endogeneity issue with sanctuaries and crime. Longitudinal data would allow for study on whether sanctuaries were any more crime-ridden or not after adopting the legislation or policy. Methods that might be useful in future papers on this subject would account for sanctuary policy over time using longitudinal data, and include factors that would affect sanctuary policy selection such as the political leanings of a county, or lagged crime, or some potential instrument that can account for the sanctuary effect.
REFERENCES


Appendix

<table>
<thead>
<tr>
<th>TABLE A: Instruments for Police Expenditure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Democratic support (%)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Non-police expenditure as a percentage of aggregate county income</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Median income ($)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

\[ \text{N} = 495 \]

\[ \text{R-squared} = 0.254 \]

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1
### TABLE B: Negative Binominal Robustness Check

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(3)</th>
<th>(5)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Property Crime</td>
<td>Property Crime</td>
<td>Violent Crime</td>
<td>Violent Crime</td>
</tr>
<tr>
<td>Sanctuary</td>
<td>-0.212*</td>
<td>-0.175</td>
<td>-0.116</td>
<td>-0.103</td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
<td>(0.116)</td>
<td>(0.115)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>Police Expenditure</td>
<td>0.108</td>
<td></td>
<td>0.352***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.120)</td>
<td></td>
<td>(0.125)</td>
<td></td>
</tr>
<tr>
<td>Endog. Police Exp.</td>
<td></td>
<td>-0.235</td>
<td></td>
<td>0.206</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.287)</td>
<td></td>
<td>(0.290)</td>
</tr>
<tr>
<td>Density</td>
<td>0.0787**</td>
<td>0.0790**</td>
<td>0.0501</td>
<td>0.0311</td>
</tr>
<tr>
<td></td>
<td>(0.0343)</td>
<td>(0.0341)</td>
<td>(0.0348)</td>
<td>(0.0345)</td>
</tr>
<tr>
<td>Non-white (%)</td>
<td>0.250***</td>
<td>0.278***</td>
<td>0.282***</td>
<td>0.334***</td>
</tr>
<tr>
<td></td>
<td>(0.0909)</td>
<td>(0.0902)</td>
<td>(0.0915)</td>
<td>(0.0920)</td>
</tr>
<tr>
<td>Unemployment (%)</td>
<td>0.585***</td>
<td>0.573***</td>
<td>0.684***</td>
<td>0.647***</td>
</tr>
<tr>
<td></td>
<td>(0.204)</td>
<td>(0.203)</td>
<td>(0.210)</td>
<td>(0.212)</td>
</tr>
<tr>
<td>Youth (%)</td>
<td>-1.952***</td>
<td>-2.090***</td>
<td>-2.259***</td>
<td>-2.469***</td>
</tr>
<tr>
<td></td>
<td>(0.397)</td>
<td>(0.390)</td>
<td>(0.430)</td>
<td>(0.430)</td>
</tr>
<tr>
<td>Poverty (%)</td>
<td>0.222</td>
<td>0.183</td>
<td>0.445***</td>
<td>0.442***</td>
</tr>
<tr>
<td></td>
<td>(0.154)</td>
<td>(0.157)</td>
<td>(0.157)</td>
<td>(0.163)</td>
</tr>
<tr>
<td>Foreign-born (%)</td>
<td>-0.0447</td>
<td>-0.00127</td>
<td>-0.00524</td>
<td>0.0491</td>
</tr>
<tr>
<td></td>
<td>(0.0742)</td>
<td>(0.0744)</td>
<td>(0.0770)</td>
<td>(0.0770)</td>
</tr>
<tr>
<td>Constant</td>
<td>7.999***</td>
<td>10.28***</td>
<td>5.092***</td>
<td>6.512***</td>
</tr>
<tr>
<td></td>
<td>(1.601)</td>
<td>(2.186)</td>
<td>(1.687)</td>
<td>(2.258)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neg. B. appropriate?</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>495</td>
<td>495</td>
<td>495</td>
<td>495</td>
</tr>
</tbody>
</table>

Explanatory variables are in logarithms. Crime rates are not in logarithms.
Relative Deprivation, Rural-Urban Migration and Rural Inequality in China

MOHAMMAD TAIMUR ALI AHMAD
Georgetown University Qatar

ABSTRACT
Rural-urban inequality in China has been increasing over the past few decades, with the prospects of higher wages in urban centers leading to mass rural-urban migration. I examine how rural inequality affects the household decision to invest in domestic migration. I use relative deprivation to measure inequality and find that relative deprivation, measured cardinally or ordinally, pushes households to invest in migration. I also estimate the impact of rural-urban migration on rural inequality in China. I use conditional quantile regressions and define inequality as the gap between the 80th and 20th percentiles. Considering that remotely-earned income is a substitute for income that otherwise would have been locally earned, I simulate the gap using counterfactual income distributions where all households field migrants or none do. I find that inequality would have been higher without migration, and would have been lower if every household had fielded a migrant.

1 Acknowledgments: I would like to thank Professor Daniel Westbrook for his constant support, mentorship and guidance throughout this project, as well as Professor Zhaoyang Hou for his insights and feedback at every stage. I would also like to thank Professor Oded Stark for his seminar at Georgetown University Qatar on relative deprivation, and Professor Mongoljin Batsaikhan and Professor Sulagna Mookerjee for their help throughout the year.
1. Introduction

During the past three decades China has made tremendous economic progress, reducing poverty from around 80% in the early 1980s to 1.9% in 2013. However, while most of the coastal cities in China have prospered, the interior rural areas have not enjoyed the same pace of development. Real urban income has increased more than real rural income for most of the past three decades; in 2005, rural income per capita was only 39% of urban income per capita (Park, 2008). Given the sharp divide between rural and urban China in terms of opportunities, labor mobility can play a significant role in economic development (World Bank, 2001). The gap between rural and urban earning potential provides an incentive for rural workers to migrate to urban centers; the incentive is stronger for those at the low end of the rural income distribution. The stock of rural-urban migration has consistently increased, reaching 155 million migrants in 2010 (Chan, 2013).

China’s rapid increase in average income – and in the incomes of the poor – has been accompanied by a surge in inequality. The contribution of the rural-urban gap to overall income inequality has been extensively documented (Du, Park and Wang, 2005; Xing, 2014, but relatively little attention has been paid to the concomitant increase in rural inequality. This is somewhat surprising, given that the Gini coefficient for rural incomes increased from 0.28 to 0.67 during 1985 to 2009 (Zhijun, 2012).

Focusing on the role of relative deprivation is justified by recent innovations presented by Stark (2017). Stark shows that the cardinal relative deprivation described in Stark and Yitzhaki (1988) can be expressed as the product of two terms: the first term is ordinal relative deprivation and the second is the burden of relative deprivation. Ordinal relative deprivation depends only on the proportion of the target’s reference group that lies above the target in the distribution of interest. The burden of relative deprivation is the average gap between the target and the outcomes of those who lie above the target in the distribution of interest.

In this paper, I do not use measures of relative deprivation based on income because they are endogenous. Instead, I use measures based on the amounts of farmland and the sizes of houses owned by households in the reference group. Rural farmland ownership is not likely to be endogenous to migration, as farmland was administratively allocated long ago. I also use measures based on house size but the exogeneity of that variable is less secure and will be investigated in subsequent work. While amounts of farmland and sizes of houses belonging to neighbors are more easily visible to the household of interest, ease of observation may not be the sole determinant of the perspective households take, so the question remains: which perspective is relevant? Indeed, if the ordinal perspective is correct but a cardinal measure is used in empirical work, then the coefficient on the measure of relative deprivation would be attenuated due to measurement error bias.

I find that ordinal relative deprivation based on the amount of farmland owned by households has a statistically significant positive effect on the decision of the household to field a migrant. In the same regression, the coefficient on burden of relative deprivation is statistically insignificant.

The paper is organized as follows. The second section contains the description of the data. In Section 3 I describe the migration determination model, the estimation methodology, and the results. Section 4 contains the income determination model, the estimation methodology, the quantile regression results, and the counterfactuals. Section 5 concludes and provides a discussion of future work that will focus on addressing endogeneity.

2. Relative Deprivation and the Migration Determination Equation

Households invest in migration not just to increase household income but also to improve the position of the household within a certain reference group (Stark, 1984; Stark and Taylor, 1989). While many basic models of household behavior are based on utility functions that are functions of income alone, Stark and Yitzhaki (1988) extend the analysis to include relative deprivation as an argument in the utility function, where relative deprivation is a measure of the degree to which the household of interest falls below other households in the reference group. The main elements of Stark and Yitzhaki
Consider the continuous random variable $w \geq 0$ which represents household income or wealth. This random variable is distributed as shown by the probability density function $f(w)$ in Figure 1. The corresponding distribution function is $P(w < z) = F(z)$.

Sample households are arranged in an ascending order:

$$(w_n < w_{n-1} < \ldots < w_z < \ldots < w_2 < w_1)$$

and $i$ is the rank of household $z$ in the distribution.

Stark (2017) shows that the cardinal measure of relative deprivation developed in Stark and Yitzhaki (1988) is the product of two terms: ordinal relative deprivation and the burden of relative deprivation. The ordinal measure reflects only the proportion of households that lie above the household of interest. Thus, ordinal relative deprivation is:

$$RDi_o = (i - 1) / n$$

Cardinal relative deprivation is ordinal relative deprivation weighted by the burden of relative deprivation. The burden is the gap between the household of interest and the mean of the households above it. Thus, the cardinal measure is:

$$RDi_c = (i - 1) / n \times \sum_{j=1}^{i-1} (w_j - w_i) / (i - 1)$$

Table 1 shows sample averages of ordinal relative deprivation, the burden of relative deprivation, and cardinal relative deprivation for households in the CHIPS 2002 data set, where the reference group for each household consists of the other households in the same county. These measures were calculated for the entire sample, using farmland per working age adult and house size per working age adult. The sample was then split into migrant and non-migrant household subsamples, and the sample averages were calculated. Relative deprivation measures based on household income were not used because they would be endogenous. However, farmland was assigned to households much earlier, during the land reform period, and hence, is not endogenous. By extension, house size is likely to be exogenous as well.

The figures in Table 1 show that households with a migrant have a higher ordinal relative deprivation, consistent across farm size per working age and house size per working age. Since these measures were constructed for the entire sample, they provide preliminary evidence that households without a migrant are more likely to be lower in the income distribution.
Interpreting the burden measure in Table 1 is more complicated. For farm size per capita, the difference suggests that households without a migrant are lower down the distribution in terms of the gap, but this could be due to the fact that a small number of non-migrant households that are at the bottom of the income distribution do not have farmland, which accentuates the gap and skews the burden measure.

Table 1. Measures of Relative Deprivation

<table>
<thead>
<tr>
<th>Variables</th>
<th>RD using farm size per working age member</th>
<th>RD using house size per working age member</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample Average</td>
<td>Difference</td>
</tr>
<tr>
<td>Ordinal RD No Migrant</td>
<td>0.46</td>
<td>0.07***</td>
</tr>
<tr>
<td>Migrant</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Burden</td>
<td>1.46</td>
<td>-0.41***</td>
</tr>
<tr>
<td>Migrant</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>Cardinal RD No Migrant</td>
<td>0.72</td>
<td>-0.13***</td>
</tr>
<tr>
<td>Migrant</td>
<td>0.59</td>
<td></td>
</tr>
</tbody>
</table>

The empirical literature on relative deprivation uses the cardinal measure and finds that it has a significant effect on migration in only some cases. Stark (2017) argues that ordinal relative deprivation should be tested separately from the cardinal measure because most households will not know the income or wealth of the households above them and will only have a general sense of where they lie in the ranking. Thus, the ordinal measure might be the more reliable measure for empirical work. Stark (2017) cites a number of empirical papers in psychology and behavioral economics.

For this paper, I assume that households do not compare themselves to the whole society; rather, I assume that they place themselves within a geographically-defined reference group. For rural households, I take this reference group to be the county of residence. In some papers migration allows a household to change its reference group by moving, but that is not the case here. The households in the data set remained in their official hukou locations and fielded individual migrants.

Because household income is endogenous in a migration equation, I base my measures of relative deprivation on two household characteristics that are less likely to be endogenous: farmland per capita and house size per capita. Both variables are positively correlated with household income per capita, making them good candidates to proxy relative deprivation in income or wealth. The two sets of relative deprivation measures – in log form – are used in separate regressions to compare their performances.

The migration determination equations are:

\[
Pr[P_i = 1 | Z_i, lrd_c, lburden] = \alpha_0 + Z_i \alpha_i + \delta_1 lrd_c + \delta_2 lburden_i + \epsilon_i \quad (1)
\]

and,

\[
Pr[P_i = 1 | Z_i, lrd_o] = \alpha_0 + Z_i \alpha_i + \delta lrd_o + \epsilon_i \quad (2)
\]

where \( P_i \) is an observed binary variable that is equal to 1 if the household has a migrant and equal to 0 for households without migrants. \( Z_i \) is a vector of household-level and village-level control variables that affect the household decision to field a migrant. Logs of cardinal relative deprivation, ordinal relative deprivation and the burden of relative deprivation are variables \( lrd_c, lrd_o, \) and \( lburden, \) respectively.

Estimating equation (1) lets me test the hypothesis that . Failure to reject this hypothesis indicates that equation (2) is the appropriate specification and that cardinal relative deprivation is the preferred measure. Estimating equation (1) also lets me test the hypothesis that . Failure to reject this hypothesis indicates that the appropriate specification contains only the ordinal relative deprivation measure. If both hypotheses are rejected, then equation (1) is the appropriate specification.
The control variables in the migration determination equation measure either the household's incentive to field a migrant or its capability to do so. Access to tap water, distance from urban centers and smaller farm size would incentivize households to invest in migration. On the other hand, the number of migrants per county and value of assets would measure the ability of the household to field a migrant. Migrants per county is a proxy for migration network, and hence, shows the ease of migration in terms of information and access, while the value of fixed and financial assets is a measure of the resources available to the household to invest in migration. Other household characteristics include an indicator for having a non-farm business, average education level of the working age adults in the household and the proportion of cash crops planted by the household.

3. Results
I used the Probit model to estimate the migration determination equations (1) and (2). Table 2 shows the coefficients on measures of relative deprivation that are based on household farm land per working age household member; complete results are presented in Appendix I. Results for relative deprivation measures based on farmland per capita, house size per working age and house size per capita are in Appendix II.

The results for the specification that includes ordinal relative deprivation and the burden of relative deprivation indicate that the hypothesis cannot be rejected; on the other hand, the coefficient on ordinal relative deprivation is highly statistically significant and the hypothesis is emphatically rejected. While the coefficients are for ordinal relative deprivation and cardinal relative deprivation both are highly statistically significant, the coefficient on the cardinal measure is smaller than the coefficient on the ordinal measure. One potential explanation for this outcome is that cardinal relative deprivation is simply ordinal relative deprivation measured with error, hence its coefficient is attenuated due to measurement error bias.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Equation (1)</th>
<th>Equation (2)</th>
<th>Equation (1) with $S_2 = 0$ imposed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coeff</td>
<td>p-value</td>
<td>coeff</td>
</tr>
<tr>
<td>Log Ordinal Relative Deprivation</td>
<td>0.0675** (0.0314)</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>Log Burden of Relative Deprivation</td>
<td>0.0311 (0.0454)</td>
<td>0.494</td>
<td></td>
</tr>
<tr>
<td>Log Cardinal Relative Deprivation</td>
<td>0.0552*** (0.0214)</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>County-Level Clustered Standard Errors</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Province Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Table 2. Coefficients on Relative Deprivation Measures

Note: Robust standard errors in parentheses.
This table shows coefficients on RD measures; refer to Appendix II for full results.

*** $p<0.01$, ** $p<0.05$, * $p<0.10$.

by estimating the probabilities of a household having a migrant at different levels of ordinal relative deprivation.
predicted probabilities are averages of the household-level predictions calculated by STATA's margins command for the observed values of all regressors except ordinal relative deprivation, which is fixed at selected values. Thus, we estimate the average migration rates for the existing sample households if the ordinal relative deprivation were constant across households, at the 20th, 50th and 80th percentiles of ordinal relative deprivation. The *p*-values are for tests of the hypothesis that the average predicted value is no greater than the average predicted value at the next lowest percentile. The results are highly statistically significant and confirm that higher ordinal relative deprivation increases the probability of the household having a migrant.

5. Endogeneity and Conclusion

An implication of Stark and Yitzhaki (1988) and Stark (2017) is that an increase in relative deprivation driven by an increase in inequality would incentivize households at the low end of the income distribution to choose some household members to engage in migration. China provides an excellent case study.

I find that, while controlling for household income and other characteristics, relative deprivation does motivate households to field migrants. Moreover, I find that Stark’s (2017) decomposition of relative deprivation into *ordinal* and *burden* components is informative: only the ordinal component is statistically significant.
References
## Table A1. Complete Migration Determination Regression

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Ordinal Relative Deprivation</td>
<td>0.0675**</td>
<td>0.0732**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0314)</td>
<td>(0.0298)</td>
<td></td>
</tr>
<tr>
<td>Log Burden of Relative Deprivation</td>
<td>0.0311</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0454)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Cardinal Relative Deprivation</td>
<td>0.0552***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0214)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total HH Income per Working Age Member (000000)</td>
<td>42.14***</td>
<td>41.83***</td>
<td>42.19***</td>
</tr>
<tr>
<td></td>
<td>(10.91)</td>
<td>(10.85)</td>
<td>(10.93)</td>
</tr>
<tr>
<td>House Size m2</td>
<td>-0.000522</td>
<td>-0.000515</td>
<td>-0.000541</td>
</tr>
<tr>
<td></td>
<td>(0.000374)</td>
<td>(0.000372)</td>
<td>(0.000376)</td>
</tr>
<tr>
<td>Farm Size m2 per Person</td>
<td>-0.0592**</td>
<td>-0.0663***</td>
<td>-0.0548*</td>
</tr>
<tr>
<td></td>
<td>(0.0295)</td>
<td>(0.0255)</td>
<td>(0.0281)</td>
</tr>
<tr>
<td>Fixed Productive Assets (000000)</td>
<td>-21.99***</td>
<td>-22.08***</td>
<td>-21.90***</td>
</tr>
<tr>
<td></td>
<td>(6.624)</td>
<td>(6.642)</td>
<td>(6.613)</td>
</tr>
<tr>
<td>Financial Productive Assets (000000)</td>
<td>0.196</td>
<td>0.140</td>
<td>0.327</td>
</tr>
<tr>
<td></td>
<td>(2.545)</td>
<td>(2.530)</td>
<td>(2.504)</td>
</tr>
<tr>
<td>Avg Edu Among Working Age in HH</td>
<td>0.0610*</td>
<td>0.0614*</td>
<td>0.0603*</td>
</tr>
<tr>
<td></td>
<td>(0.0326)</td>
<td>(0.0326)</td>
<td>(0.0324)</td>
</tr>
<tr>
<td>Indicator for HH with Non-Farm Biz</td>
<td>-0.194**</td>
<td>-0.194**</td>
<td>-0.193**</td>
</tr>
<tr>
<td></td>
<td>(0.0757)</td>
<td>(0.0757)</td>
<td>(0.0754)</td>
</tr>
<tr>
<td>Indicator for HH Has Tap Water</td>
<td>-0.0588</td>
<td>-0.0603</td>
<td>-0.0566</td>
</tr>
<tr>
<td></td>
<td>(0.0524)</td>
<td>(0.0518)</td>
<td>(0.0525)</td>
</tr>
<tr>
<td>Household Head Gender</td>
<td>0.103</td>
<td>0.0995</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>(0.148)</td>
<td>(0.148)</td>
<td>(0.147)</td>
</tr>
<tr>
<td>Household Head Age</td>
<td>-0.000652</td>
<td>-0.000565</td>
<td>-0.000733</td>
</tr>
<tr>
<td></td>
<td>(0.00233)</td>
<td>(0.00230)</td>
<td>(0.00231)</td>
</tr>
<tr>
<td>Household Head Communist Party Member</td>
<td>-0.162***</td>
<td>-0.162***</td>
<td>-0.161***</td>
</tr>
<tr>
<td></td>
<td>(0.0471)</td>
<td>(0.0472)</td>
<td>(0.0471)</td>
</tr>
<tr>
<td>Number of Young Children</td>
<td>-0.0513</td>
<td>-0.0523</td>
<td>-0.0507</td>
</tr>
<tr>
<td></td>
<td>(0.0472)</td>
<td>(0.0472)</td>
<td>(0.0472)</td>
</tr>
<tr>
<td>Number of Working Age</td>
<td>0.360***</td>
<td>0.362***</td>
<td>0.359***</td>
</tr>
<tr>
<td></td>
<td>(0.0286)</td>
<td>(0.0277)</td>
<td>(0.0284)</td>
</tr>
<tr>
<td>Prop of HH with Migrant by Coun</td>
<td>3.164***</td>
<td>3.172***</td>
<td>3.148***</td>
</tr>
<tr>
<td></td>
<td>(0.169)</td>
<td>(0.169)</td>
<td>(0.171)</td>
</tr>
<tr>
<td>Prop of HH with Non-Farm Biz by Vill</td>
<td>-1.650***</td>
<td>-1.643***</td>
<td>-1.660***</td>
</tr>
<tr>
<td></td>
<td>(0.466)</td>
<td>(0.464)</td>
<td>(0.466)</td>
</tr>
<tr>
<td>Prop of Economic Crops by Vill</td>
<td>-0.297**</td>
<td>-0.297**</td>
<td>-0.296**</td>
</tr>
<tr>
<td></td>
<td>(0.141)</td>
<td>(0.141)</td>
<td>(0.141)</td>
</tr>
<tr>
<td>Distance from nearest township govt</td>
<td>0.0155**</td>
<td>0.0155**</td>
<td>0.0155**</td>
</tr>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Mountainous Area</td>
<td>0.0468</td>
<td>0.0461</td>
<td>0.0475</td>
</tr>
<tr>
<td></td>
<td>(0.0661)</td>
<td>(0.0659)</td>
<td>(0.0667)</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.823***</td>
<td>-2.840***</td>
<td>-2.809***</td>
</tr>
<tr>
<td></td>
<td>(0.270)</td>
<td>(0.265)</td>
<td>(0.267)</td>
</tr>
<tr>
<td>Province Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>County-Level Clustered Standard Errors</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
Bank Consolidation and Systemic Risk: M&A During the 2008 Financial Crisis

GREGORY D. MASLAK
GONCA SENEL
Bowdoin College

ABSTRACT
This paper analyzes the relationship between bank consolidation and the stability of the financial system within the United States. In particular, we compare mergers and acquisitions of banks during the 2008 financial crisis with those that occurred during stable market conditions in order to determine whether the effects of bank consolidation on the overall economy at all differ depending upon the macroeconomic climate. The systemic risk measures of MES, SRISK, NSRISK, and ΔCoVaR are calculated both before and after a merger so as to capture the consequent merger-related change in an acquirer's exposure as well as contribution to systemic risk. Difference-in-differences analysis is subsequently conducted with a non-merging control group to determine whether the change in these risk metrics can accurately be attributed to the merger. For MES, NSRISK, and ΔCoVaR the results indicate that mergers during the defined crisis period contained a distinct diminishing effect on the acquiring bank's risk. Furthermore, when examining the underlying characteristics driving this observable difference through the use of a logit model, we find that mergers during the 2008 financial crisis tended to involve acquirers that possessed more diversified sources of income and less leverage than their stable market counterparts. Meanwhile, the targets of these transactions often were less profitable banks with lower tier 1 capital ratios, but possessed a greater share of deposits making them more vulnerable to the downturn, but also attractive acquisition targets. Overall, the findings of this paper suggest that during the 2008 financial crisis healthy banks acquired poorly performing target banks and successfully integrated them into their own operations.
An Application of the Phillips Curve to the Indian Economy

SHAMBHAVI TIWARI
Columbia University

ABSTRACT
Using data from the Indian economy spanning January 1980 to August 2016, I find that the tradeoff between inflation and unemployment implied by the Phillips Curve is nonexistent. Using supply shock variables for liberalization, droughts, and oil prices, I fit a New-Keynesian Phillips Curve (NKPC) to the Indian economy. The NKPC proves to be of little applicability to the Indian case, despite the inclusion of supply shocks, and despite the application of both OLS and GMM estimators.

1 My sincere thanks to my professor, Dr. Jushan Bai, for his guidance and feedback throughout the research process.
1. Introduction

The Phillips Curve is often used as a macroeconomic rule, and helps shape policy decisions made by many countries’ central banks. Can the Reserve Bank of India, which is the Indian central banking institution, use this macroeconomic rule as a policy determinant? Indeed, the relationship between inflation and unemployment is particularly important considering that RBI considers price stability to be one of its major goals. If so, can the curve be made applicable by certain modifications, such as the addition of supply shocks specific to the Indian economy? Finally, one may ask, why should we attempt to fit the Phillips Curve to the Indian economy?

My results are a departure from the findings of recent literature on the NKPC in the Indian economy, demonstrating a need for reevaluation of traditional macroeconomic models, and more specifically, their applicability to developing economies. Perhaps, in line with historical trends, Indian WPI inflation has returned to its previous unpredictability. The inapplicability of the NKPC to the Indian economy in my research may imply that the unemployment inflation tradeoff only applies to advanced economies, although further research on the NKPC in developing economies is necessary before reaching this conclusion. This is in fact a troubling suggestion, since the Phillips Curve has long been the backbone of macroeconomic policy-making.

My paper is organized as follows. The next section will cover the theoretical foundations of the Phillips Curve. Section 3 is a brief literature review that carries onto Section 4, covering data-related issues. Section 5 provides the results of my estimation, and Section 6 concludes.

2. Theoretical foundations

In its broadest application, the Phillips Curve provides basis for the existence of a tradeoff between inflation and unemployment. Further work on the Phillips Curve helped derive it from its microeconomic foundations in the product market. As explained by Sahu (2013), this specific model of the Phillips Curve, known as the New Keynesian Phillips Curve (NKPC), assumes that individual firms are monopolistically competitive and set prices by accounting for their current marginal costs and maximizing their expected future profits. This makes prices sticky in the NKPC: making firms set prices based on expectations of future prices as well as their current marginal costs. Firms who are experiencing high marginal costs will reduce the number of workers they employ. Thus, the NKPC is modeled as below:

\[ \pi_t = -\theta(u_a - u_n) + E_t \pi_{t+1} \]

Where \( \pi_t \) is the inflation at time \( t \), \( u_a \) is actual unemployment at time \( t \), and \( u_n \) is the natural rate of unemployment at time \( t \). Thus \( \theta \) measures the effect of cyclical unemployment on inflation \( \pi_t \), \( E_t \pi_{t+1} \) is agents’ expectation of inflation at time \( t+1 \) when they are in time \( t \).

Using Okun’s Law, one can substitute the term \( u_a - u_n \) with the output gap. This is also a relevant transformation for my problem because monthly unemployment data for the Indian economy is rarely found in freely available datasets on the Web. The above equation is transformed to the following:

\[ \pi_t = \beta(y_a - y_n) + E_t \pi_{t+1} \]

Where \( y_a \) denotes actual output, \( y_n \) denotes potential output, and thus \( \beta \) becomes the coefficient for the output gap.

Gordon (1984) points to the usefulness of adding supply shocks to the Phillips Curve. Indeed, supply shocks, which exogenously cause shifts in the aggregate supply curve, affect both inflation and output levels, and are included in the NKPC as below:

\[ \pi_t = \beta y_t + E_t \pi_{t+1} + S_t \]  

(1)
Where \( y_t \) denotes the output gap and \( S_t \) represents the various country-specific supply shocks.

### 3. Data

My data spans from January 1980 till August 2016, which makes it quite recent as compared to past literature. Inflationary expectations data for India is inconsistently found and only available for the most recent years. Due to this data availability problem, I assume rational expectations. With rational expectations, agents have perfect foresight and thus I replace \( E_t\pi_{t+1} \) in equation (1) above with \( E_t\pi_{t+1} = \pi_{t+1} \).

However, as explained by Jain (2015), the rational expectations-based NKPC has been heavily criticized, as it does not factor in past inflation, and the hybrid NKPC has been proposed instead by Galí and Gertler (1999) and Galí and López-Salido (2005). In the hybrid NKPC, present inflation is dependent on both future and past inflation, which is a sensible proposal considering that the New Keynesian model assume prices to be sticky and firms to have imperfect information, which leads to firms setting prices based on future expectations of profits as well.

\[
\pi_t = b_0 + \beta (L^k) y_t + \alpha (L^p) \pi_{t-1} + \mu (L^n) S_t + \nu \pi_{t+1} + \epsilon_t
\]

Where \((L)\) represents lags of order \( k, p, n \) on output gap, past inflation, and supply shocks respectively.

The inclusion of the term \( \pi_{t+1} \) is contentious in literature on the Phillips Curve. I will also make estimates in my results section omitting the term \( \pi_{t+1} \) in equation (2), making my new estimation of the form:

\[
\pi_t = b_0 + \beta (L^k) y_t + \alpha (L^p) \pi_{t-1} + \mu (L^n) S_t + \epsilon_t
\]

Sahu has a legitimate concern particularly to developing countries: modeling only \( y_t \) is faulty because agriculture is a significant portion of the Indian GDP (17% of the 2015 Indian GDP according to World Bank estimates). Paul explains that such a specification would be lead to multicollinearity, and suggests instrumenting agricultural output by using droughts as supply shocks, since a nationwide drought would be most harmful in reducing agricultural output.

Dua and Gaur claim that using output gap for demand side factors is insufficient for developing countries, and propose using both industrial production output gap and real money output gap in the hybrid NKPC:

\[
\pi_t = b_0 + \beta (L^k) y_t + \lambda (L^m) m_t + \alpha (L^p) \pi_{t-1} + \mu (L^n) S_t + \nu \pi_{t+1} + \epsilon_t
\]

Where \( y_t \) denotes the IIP (Index of Industrial Production) output gap, \( m_t \) is the real money output gap, calculated as the difference between actual and potential monetary aggregate (using M3 measures of money in the economy).

As Jain (2015) recognizes, data on inflation in India is found in the Wholesale Price Index (WPI) and the Consumer Price Index (CPI). Due to the very recent nature of aggregated CPI estimates for the Indian economy, I will use WPI data in my regressions.

I choose to use monthly data so I can record the effect of month-on-month changes in the regressors included in my NKPC estimation, as opposed to year-on-year changes. My calculations help me to avoid the base effect.

Data used for WPI is found from the Office of the Economic Advisor’s website. Data is unified to a single base year using the ‘arithmetic conversion’ method recommended by the Indian government; I rebase all the WPI to the base year of 2004-05.

I proxy output using monthly industrial production data, found from the Index of Industrial Production (IIP) from IMF’s International Finance Statistics (IFS) database. The entire data is based on the year 2010. To calculate the output gap, I transform IIP data as recommended by Mazumder (2011), Sahu (2013), Kapur (2013), and the majority of the literature on the Indian Phillips Curve. I first deseasonalize my data using the STATA moving average filter, and I then de-trend it.
using the Hodrick-Prescott (HP) filter. I also use data for broad money (M3) from the OECD and calculate the output gap identical to the calculation for the IIP output gap above.

In accordance with the NKPC estimations of Mazumder (2011) that augment Paul’s, I use the supply shock of global oil price levels in Indian Rupees (INR) instead of global oil price dummies, by multiplying crude oil prices by the US dollar to Indian rupee exchange rate. This particular calculation allows my supply shock of global price levels to also reflect any shocks to the US Dollar/INR exchange rate, which is a significant supply shock in both Sahu and Kapur’s estimations(2013).

Using Paul’s specification of the liberalization dummy variable, I create a dummy variable \( lib \) that takes the value of 1 for the years 1992-95 and 0 otherwise. During this period, the Indian economy underwent “industrial delicensing, exchange rate reform, partly liberalizing capital flows, and current account convertibility,” according to Paul. One cannot expect the shock of \( lib \) to be represented through the other variables included in the NKPC specification (2) because the shocks that the economy experienced during this liberalization period were multiple and widespread through the economy.

Paul explains that instrumenting agricultural output using a dummy variable for droughts is a complicated process since only a certain level of drought will harm agricultural output. I take a more rudimentary approach than Paul by investigating recent rainfall deficiencies. I calculate the average of annual rainfall levels from 1901-2013, with data obtained from the Indian Meteorological Department. I then calculate the percentage deviation of each year’s rainfall from its average value. I find that the percentage deviations of rainfall in the years 2009 and 2012 are just as excessively negative as the ones in the drought years of 1982, 1987, and 2002. I thus extend the dummy \( drought \), and assign it the value 1 for the years 2009 and 2012 (using Paul’s reasoning to assign \( drought = 1 \) for each post-drought year as well), and 0 otherwise.

Although I have followed similar data collection and processing methods to the past literature, my data still has several shortcomings that should be noted. Firstly, the WPI does not account for any services provided. This is a major shortfall, since Jain (2015) recognizes the services sector to be the largest in the Indian economy. The calculation of the WPI according to base year 2004-05 is also faulty as this base year is dated and needs to be updated as well. Additionally, the HP filter has been critiqued by Mohanty and Klau (2001). Although the HP filter has been conventionally used to detrend data, the selection of an arbitrary \( \lambda \) has little theoretical foundation). Finally, the assumption of perfect foresight and rational expectations is very unlikely to be true in most economies, and especially in developing countries where literacy rates are low and the majority of the population is poorly educated, in comparison to their counterparts in developed economies.

4. Estimation results

Mazumder (2011) adopts an OLS estimation of the Indian NKPC, with heteroskedasticity autocorrelation robust (HAC) standard errors. The results of my estimation are provided in Table 2 below, and in order to be concise I choose to only provide the significant coefficients. I estimate the backwards-looking NKPC, consistent with equation (2) above, using only IIP output gap in specification 1 and using both IIP and M3 output gap in specification 3. In specification 2 and 4, I estimate the hybrid NKPC that is consistent with equation 1 above.

My results are surprising given the consistency of past literature after Paul in finding evidence for the Phillips Curve for the Indian case. All of my OLS regressions show a negative relationship between output gap (both M3 and IIP) and WPI with certain lags. Looking specifically at the twice-lagged IIP output gap, specifications 1-4 tell me that the effect of a percentage increase in the output gap (which implies that actual output is above its potential) would cause a decrease in the WPI two periods later ranging between 0.08-0.14 percent. This result is counterintuitive to the economic theory.
Table 1. Abbreviated OLS coefficients with HAC standard errors estimation of the Indian NKPC

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Spec. 1 Backwards looking NKPC</th>
<th>Spec.2 Hybrid NKPC</th>
<th>Spec.3 Backwards looking NKPC with IIP + M3 output gap</th>
<th>Spec. 4 Hybrid NKPC with IIP + M3 output gap</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-</td>
<td>0.31*** (0.04)</td>
<td>-</td>
<td>0.30*** (0.04)</td>
</tr>
</tbody>
</table>

p<0.10, ** p<0.05, *** p<0.010

However, the usage of OLS estimation for the NKPC is problematic, as explained by Baum (2006, p200), “since both variables [inflation and unemployment] are determined within the macroeconomic environment, we cannot consider either as exogenous.” Indeed, the endogeneity of both inflation and output gap make Hayashi et al. (2015) and Dua and Gaur (2009) adopt the usage of instrumental variables for their estimations.

Conventionally, the Phillips Curve is estimated using the Generalized Method of Moments (GMM) technique. GMM presents several advantages for estimating the NKPC over OLS. More specifically, the OLS assumption of conditional exogeneity of the error term $E(u | x) = 0$, that the error term and the regressors are uncorrelated, is violated by the existence of endogenous regressors. With the GMM estimation, I can specify which of my regressors are endogenous and how they will be instrumented, even when faced with heteroskedasticity of unknown form. Valid instruments, of course, have to be significantly correlated with the endogenous regressors and uncorrelated with the error process.

The results of my GMM estimations with HAC standard errors are presented in tables 3 and 4 below. In Table 3, I estimate specification 5, with the assumption that once-lagged WPI and IIP output gap are endogenous and instrumented by their lags, oil prices, and the drought variable. I leave the dummy variable for liberalization as an additional exogenous variable. In specification 6, I assume the hybrid NKPC and instrument $\pi_{t+1}$ WPI inflation and IIP output gap on their lags, oil prices, and the drought dummy variable. Like before, I leave the dummy variable for liberalization as an additional exogenous variable. In specification 7, I assume the hybrid NKPC and instrument $\pi_{t+1}$ WPI inflation and IIP output gap on their lags only, leaving oil prices, the drought and liberalization dummy variables as additional exogenous variables in the regression.

As seen below, due to collinearity issues, Stata chooses to drop a certain number of variables for each regression.

Table 2. Abbreviated GMM estimation with HAC standard errors of the Indian NKPC

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient (Standard Errors)</td>
<td>Coefficient (Standard Errors)</td>
<td>Coefficient (Standard Errors)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0.44*** (0.06)</td>
<td>0.47*** (0.07)</td>
</tr>
<tr>
<td></td>
<td>0.35*** (0.07)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.15</td>
<td>0.15</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Underidentification test (LM statistic)

<table>
<thead>
<tr>
<th></th>
<th>Specification 5</th>
<th>Specification 6</th>
<th>Specification 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-sq. P-val</td>
<td>86.525</td>
<td>104.387</td>
<td>95.286</td>
</tr>
<tr>
<td>Chi-sq. P-val = 0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Hansen J statistic

<table>
<thead>
<tr>
<th></th>
<th>Specification 5</th>
<th>Specification 6</th>
<th>Specification 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-sq. P-val</td>
<td>101.619</td>
<td>85.642</td>
<td>83.976</td>
</tr>
<tr>
<td>Chi-sq. P-val = 0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

p<0.10, ** p<0.05, *** p<0.010
Although the results for the IIP output gap are insignificant for specification 5, for specifications 6 and 7 I find that a percentage increase in the output gap corresponds to a 0.04 percentage decrease in WPI inflation, which is a significant coefficient. Once again, this result is counterintuitive, and indeed the fact that the one-step forward lagged WPI inflation accounts for the majority of this regression is alarming because it suggests that the Indian WPI is unpredictable using this specification. Although my under identification test tells me that my endogenous regressors are adequately correlated with their instruments, the strong rejection of the Hansen J-statistic also tells me that the instruments are not uncorrelated with the error process, a key requirement for an appropriate instrument.

In Table 4 below, I estimate the GMM NKPC with HAC standard errors for specifications 8-10. In the specifications below, I choose to include M3 output gap as suggested by Dua and Gaur (2009). In specification 8 I assume that once-lagged WPI inflation, IIP and M3 output gaps are endogenous and instrumented by their lags only. I leave oil prices, the drought and liberalization dummy variables as additional exogenous variables in the regression. In specification 9, I assume the hybrid NKPC and instrument $\pi_{t+1}$ WPI inflation, IIP output gap and M3 output gap on their lags only, leaving oil prices, the drought and liberalization dummy variables as additional exogenous. Finally, in specification 10, I assume the hybrid NKPC and instrument $\pi_{t+1}$ WPI inflation and IIP output gap on their lags and oil prices. I assume the drought and liberalization dummy variables as additional exogenous variables in the specification. Disappointingly, my results continue to show a negative coefficient for output gap. The rejection of the Hansen J-statistic along with the low R-squared show that I still have weak instruments which do not satisfy the exogeneity requirement, although the model's rejection of the underidentification test shows that my endogenous regressors continue to be adequately correlated with the instruments.

| Table 3. Abbreviated GMM estimation with HAC standard errors of the Indian NKPC |
|-----------------------------------|-----------------------------------|-----------------------------------|
| Regressor                         | Coefficient (Standard Errors)     |                                  |
|                                  | Specification 8                  | Specification 9                  | Specification 10                  |
|                                  |                                  | 0.47*** (0.07)                   | 0.44*** (0.06)                   |
| ✗                                 |                                  | 0.38*** (0.07)                   |                                  |
| R-squared                         | 0.18                             | 0.18                             | 0.15                             |
| Underidentification test (LM statistic) | 92.967                          | 102.122                          | 105.690                          |
|                                  | Chi-sq. P-val = 0.0000           | Chi-sq. P-val = 0.0000           | Chi-sq. P-val = 0.0000           |
| Hansen J statistic               | 90.029                           | 88.434                           | 90.650                           |
|                                  | Chi-sq. P-val = 0.0000           | Chi-sq. P-val = 0.0000           | Chi-sq. P-val = 0.0000           |

p<0.10, ** p<0.05, *** p<0.010

With the consistent misspecification of the model as explained by the rejection of the Hansen J-statistic, I consider testing whether my endogenous regressors of WPI inflation and IIP and M3 output gap are indeed endogenous. I estimate the C-Statistic of 0.434 for specification 5, which means that I cannot reject the null that IIP output gap and WPI inflation are exogenous variables. This suggests, as explained by Baum (2006), that a linear OLS regression cannot be considered a worse option than GMM. Indeed, since Baum also explains that using GMM estimation methods results in a loss of efficiency as the asymptotic variance of the instrumental variables estimator is often much larger than that of the OLS estimator, perhaps OLS is a better estimation. Although one cannot reject the claim that the OLS specifications of Table 2 were the best estimations of the Indian NKPC, the question of the negative coefficient on output gap still remains.
5. Conclusion
My OLS and GMM estimations for the NKPC in the Indian case found that the relationship between output gap and WPI inflation is counter intuitively negative. Although my results are not in line with previous literature on the Indian NKPC, my usage of monthly and more recent data that extends from January 1980 till August 2016 represents a more novel approach.

Indeed, my estimations’ rejections of the unemployment-inflation tradeoff, the low R-squared of my GMM estimations and the rejection of the Hansen J-Statistic provide insight to some serious misspecification issues. Perhaps, in line with historical trends, Indian WPI inflation has returned to its previous unpredictability. Certainly, the inapplicability of the NKPC to the Indian economy as shown in my research may imply the unemployment-inflation tradeoff only applies to advanced economies, further application to developing economies is necessary before reaching this conclusion.
References
Baum, C. 2006. An Introduction to Modern Econometrics using Stata, StataCorp LP.

Data obtained from:
Reserve Bank of India (RBI). https://www.rbi.org.in
OECD Data. https://data.oecd.org
Petroleum Planning and Analysis Cell, Government of India. http://ppac.org.in
Indian Meteorological Department. http://www.imd.gov.in/Welcome%20To%20IMD/Welcome.php
**Goodbye Lenin, Hello Murat?¹**

*The Effect of Communism on Individual Attitudes Toward Immigration*

MATTHEW CARL²

*Washington & Lee University*

**ABSTRACT**

This paper argues conceptually and demonstrates empirically that individual attitudes toward immigration are deeply affected by a country’s politico-economic legacy. Drawing on data from the German Socio-Economic Panel and exploiting a quasi-natural experiment arising from the historic division of Germany into East and West, I show that former East Germans are, because of their exposure to communism, notably more likely to be very concerned about immigration than former West Germans that were never exposed to the communist regime. The effect of living in East Germany is driven by those former East Germans that were actually born during, and not prior to the imposition of, the communist rule. I further find that differences in attitudes held by former East versus West Germans converge over time and that the level of trust in strangers represents one salient channel through which communism affects individual preferences toward immigration.

1 Inspired by the popular film “Good Bye Lenin” (2003). “Murat” is a common Turkish name.
2 I would like to thank first and foremost Professor Peter Grajzl for his mentorship and unconditional support of this project over the past year. I would also like to acknowledge Professors Joseph Guse, Art Goldsmith, and Neils-Hugo Blunch for their incredibly helpful comments and advice, whether in conversations or coursework, all of which have inexorably woven their way into the pages of this paper. Without the assistance of these individuals, this paper would not have been possible; I am profoundly grateful to each of you.
On the Effects of Intellectual Property Rights on Foreign Direct Investments in Developing Economies: A System GMM Approach

VINCENT JERALD R. RAMOS
University of the Philippines

ABSTRACT
This paper aims to examine the effects of Intellectual Property Rights quantity (IPRPC) and protection quality (IPRI) on Foreign Direct Investment (FDI) inflows in a panel of 79 low- and middle-income countries with average per capita real GNIs of not more than $12,475 over the period 2000-2014, also accounting for the existence of nonlinearities. Using the semiparametric Blundell and Bond’s [1998] and Windmeijer’s [2005] two-step System Generalized Method of Moments (SGMM) to obtain more asymptotically efficient estimates, address endogeneity issues, and factor in the inclusion of nonstationary macroeconomic variables, this paper finds that for at least half of the countries in the sample, an improvement in IPR protection quality has a positive marginal effect on FDI inflows. Meanwhile, per capita IPR has a negative marginal impact on FDI inflows in more than 95% of the countries in the sample. The nonlinearity in the per capita IPR-FDI nexus is such that increases in IPR applications would not necessarily translate into higher FDI inflows until a minimum scale of IPR applications is reached.
1. Introduction
Foreign Direct Investment Inflows (FDIs), measured locally, is the cumulative value of all investments in the receiving country made directly by residents of other countries primarily to companies within a specified period (World Development Indicators, 2016). It is an aggregation of equity capital, reinvestment of earnings, other long-term capital, and short-term capital as shown in the balance of payments (World Development Indicators, 2016). Among the types of financial or capital inflows, FDIs are the most associated with physical and human capital building and employment generation in the domestic economy. Moreover, the literature identifies FDIs as a vehicle for transfers of technology; thus helping create an environment that promotes competition in industries. As such, FDI inflows are crucial to the industrial development of emerging economies.

Among the many factors that affect FDIs, Intellectual Property Rights (IPR) protection is purported to be an important, but relatively less explored determinant of FDI. The sparsely existing IPR-FDI literature can be divided into three strands: a positive relationship (Mansfield, 1995; Maskus, 1998; Smarzynska, 2004; Adams, 2010; and Khan and Samad, 2010), an ambiguous or insignificant relationship (Braga and Fink, 2001; Fink and Maskus, 2005), and a negative relationship (Maskus, 2000; Yang and Maskus, 2001; Gathii, 2015). This tension in the literature might be explained by the two opposing effects that stronger IPR protection can have on FDIs: A stronger IPR protection can be a signal of a better overall institutional quality that is necessary for innovations and thus FDIs to thrive in the domestic economy; however, stronger IPR protection, in terms of longer patents and trademarks, might also deter competition and innovation by hindering new entrants into the market – both local and foreign (Branstetter and Saggi, 2009).

This study takes off particularly from the study of Adams (2010) which asks “What is the impact of Intellectual Property Rights (IPR) Protection on Foreign Direct Investment?” His study uses a panel data involving 75 developing countries over 19 years (1985-2003) and concludes that those with stronger IPR protection benefits from more foreign direct investment inflows. The variables IPR and FDI are significantly and positively correlated.

This paper, however, contributes to the literature in two ways. One, it examines the IPR-FDI relationship, accounting for the possible existence of nonlinearities, using two approaches—first, using a qualitative indicator of IPR protection which is the traditional approach in studies dealing with IPR-FDI relationship and second, using a scaled quantitative measure of the stock of IPR in the form of per capita IPR. The latter approach has not been done previously in studies dealing with IPR-FDI relationship. Two, it restricts the sample to low- and middle-income economies, focusing on more recent years to capture new developments in intellectual property protection and to ascertain the fact that the determinants of FDI in developing economies might be distinct from those in developed economies.

This research has the following objectives. The first is to test both the qualitative (IPR protection index) and quantitative (per capita IPR) effects of IPR on FDI. Hopefully, this research can be a useful contribution in further research and policymaking related to FDI and reforms in IPR regimes. The second is to ascertain the findings of existing literature using a dataset that contains more recent years and dropping the pre-TRIPS (Agreement on Trade-Related Aspects of Intellectual Property Rights) era years since the effect of the TRIPS agreement on the IPR-FDI link has already been established in various studies.

The sample is a balanced panel dataset consisting of 79 low and middle income countries with average per capita real GNI of not more than $12,475 over the period 2000-2014. The estimation technique of this paper uses Blundell and Bond’s (1998) and Windmeijer’s (2005) two step System Generalized Method of Moments (SGMM) to obtain more asymptotically efficient estimates and to address endogeneity issues.

To preview the results, this paper confirms the existence of nonlinearities in the IPR-FDI nexus, in line with the findings

---
1 Adams (2010) finds that the impact of IPR protection on FDI inflows is far greater in the post-TRIPS era than the pre-TRIPS era.
2 The latest World Bank (July 2016) update reports that middle income countries are countries with average per capita real GNI of not more than $12,475.
of Asid (2004). In terms of the IPR protection index—a proxy measure of IPR quality—IPR has a positive marginal effect of FDI inflows for at least half of the countries in the sample. This supports existing literature on positive IPR-FDI relationship (Mansfield, 1995; Maskus, 1998; Smarzynska, 2004; Adams, 2010; and Khan and Samad, 2010). Meanwhile, per capita IPR has a negative marginal impact on FDI inflows. The nonlinearity in the IPR-FDI nexus, in this instance, implies that increases in the stock of patents, trademarks, and copyrights would not necessarily translate into higher FDI inflows unless this stock reaches a certain threshold, which is only attained between the 95th and 99th percentile observations of per capita IPR. These results suggest that stronger IPR regimes are additional incentives for foreign investors but have to be accompanied by bureaucratic reform that will minimize red tape and improve the overall investment climate of a country. Figure 1 in the annex shows the relationship of FDIs, Transfer of Technology, and Economic Development, as pursued in this paper.

The rest of the paper is organized as follows: Section 2 features a review of theoretical and empirical literature on IPR-FDI relationship. Section 3 presents the methodology—data sources, estimation technique, and preliminary data analysis. Section 4 discusses the results of the study. Section 5 offers policy implications, recommendations, and concluding remarks.

2. Review of Related Literature

Conceptually, strengthening IPR regimes alone is an insufficient incentive for firms to invest in the receiving country. The receiving country must ensure that there is a properly implemented legal framework to protect the IPRs by which the investors’ advantage is obtained. Therefore, mutual recognition and protection of IPRs must be established. Empirical literature is divided on whether or not stronger IPR regimes are likely to positively affect FDI inflows. There are three strands by which existing literature can be classified. The table below summarizes the literature included.

<table>
<thead>
<tr>
<th>Source (Year)</th>
<th>Relationship</th>
<th>Findings and Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maskus (1998)</td>
<td>Positive</td>
<td>IPR is part of many interlinked components that attract FDI</td>
</tr>
<tr>
<td>Adams (2010)</td>
<td>Positive</td>
<td>IPR is positively and significantly correlated with FDI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-TRIPS era associated with better IPR regimes contributing to FDI</td>
</tr>
<tr>
<td>Smarzynska (2004)</td>
<td>Positive</td>
<td>IPR affects probability of investment in high technology sectors that heavily rely on IPR</td>
</tr>
<tr>
<td>Mansfield (1995)</td>
<td>Positive</td>
<td>The strength or weakness of a country’s system of IP protection seems to have a substantial effect in relatively high-technology industries</td>
</tr>
<tr>
<td>Khan and Samad (2010)</td>
<td>Positive</td>
<td>In the case of East Asia, stronger IPR regimes increase the likelihood of FDI not only in the manufacturing sector but also in retail and distribution networks</td>
</tr>
<tr>
<td>Fink and Maskus (2005)</td>
<td>Ambiguous</td>
<td>The welfare implications of a particular regime are theoretically ambiguous due to limited empirical evidence</td>
</tr>
<tr>
<td>Braga and Fink (2001)</td>
<td>Ambiguous</td>
<td>The effects of IPR on bilateral trade flows are theoretically ambiguous.</td>
</tr>
<tr>
<td>Yang and Maskus (2001)</td>
<td>Negative</td>
<td>IPR has a negative significant effect (1 percent level) in licensing receipts</td>
</tr>
</tbody>
</table>

3 The results obtained in this study cannot be used to explain the IPR-FDI nexus in developed economies.
Gathii (2015) | Negative | - | Strong IP protection least developed economies is unlikely to yield economic benefits of stronger FDI flows

Maskus (2000) | Negative | - | stronger IPR protection has short term net welfare losses

### 2.1 Positive IPR-FDI relationship

IPR is just one of the many interlinked components that attract FDI including tax system, competition regimes, corruption perception, and infrastructure development, among others (Maskus, 1998).

The finding of Maskus (1998) is consistent with the conclusion by Adams (2010). The study uses a panel data involving 75 developing countries over 19 years (1985-2003) and boldly concludes that those with stronger IPR protection benefits from more foreign direct investment inflows. The variables IPR and FDI are significantly and positively correlated. This particular study takes a closer look at the pre- and post- TRIPS agreement era. The Trade Related Aspects of Intellectual Property Rights (TRIPS) agreement era came into effect in January of 1995 to heed stronger calls for reforms in IPR regimes. Under the terms of TRIPS, current and future members of WTO must adopt and enforce strong non-discriminatory minimum standards of intellectual property protection in each of the areas commonly associated with IPRs including patents, copyrights, trademarks, and trade secrets (Adams, 2010).

Smarzynska (2004) has a different observation of the effect of IPR on investments in high-technology sectors relying heavily on IPRs (such as manufacturing enterprises). In five of the six regressions, IPR protection affects the probability of investment in the said sectors. They bear positive and significant (five percent level) coefficients. This study utilizes a probit model using firm-level data based on the EBRD Foreign Investment Survey conducted in 1995 and the Ginarte-Park Index of IPR Protection (1997). Smarzynska (2004) gives an even bolder conclusion based on the results of her regressions—all firms may be affected by IPR protection because IPR regimes play a signaling role.

The finding of Smarzynska (2004) is consistent with the findings of Mansfield (1995) where he concludes that the “strength or weakness of a country’s system of IP protection seems to have a substantial effect in relatively high-technology industries like chemicals, pharmaceuticals, machinery, and electrical equipment on the kinds of technology transferred to that country and the amount of FDI in that country by Japanese and German firms.” This study utilized Linear OLS regressions. Meanwhile, a study on the effect of IPR on inward FDI in the case of 14 Asian and South East Asian countries concludes that stronger IPR regimes increase the likelihood of FDI not only in the manufacturing sector but also in retail and distribution networks (Khan and Samad, 2010). This study uses a panel data from 1970-2005 and empirical estimates were derived using pooled OLS techniques.

### 2.2 Insignificant or Ambiguous IPR-FDI relationship

An argument for a relaxed IPR regime is that the owner of IPR abuses its monopoly power over the first placement of a protected good or service by preventing parallel trade of the said good or service by third parties since no other firm or person can replicate their product until such time when the IPR expires. A solution was applied in Europe and the US and was called the principle of exhaustion in the former and first sale doctrine in the latter. Under this principle, IPR owners were no longer entitled to control the subsequent marketing strategies of the protected products beyond what is legitimately necessary to protect the subject matter of the rights (Transfer of Technology, 2001).

Theoretically, the purpose of the exhaustion principle is justifiable. However, it is one of the most complicated regulations of international business. Fink and Maskus (2005) conclude that the welfare implications of a particular exhaustion regime are theoretically ambiguous due to limited empirical evidence and are likely to differ based on the form of IPR, and industry- and product-specific considerations.

Braga and Fink (2001) provide evidence regarding the effect of patent protection on international trade by using a gravity model of bilateral trade flows and estimating the effects of increased patent protection on a cross-section of 89x88
countries. The aggregates used are limited to total non-fuel trade and high technology based on expectations that the effect of stronger patent protection is more eminent in knowledge-intensive traded goods and services. This study concludes that the effects of IPR on bilateral trade flows are theoretically ambiguous. However, the estimation of the gravity model provides empirical evidence that higher levels of protection have a significantly positive effect on non-fuel trade. The effect of IPR on high technology trade is statistically insignificant. To further widen the knowledge on the IPR-FDI link, empirical research focusing on industry and firm-specific variables should be done.

2.3 Negative IPR-FDI relationship
The study by Yang and Maskus (2001) provides a counterintuitive result where they find that IPR has a negative significant effect (1 percent level) in licensing receipts whereas its squared term has a positive significant effect (10 percent level) under a pooled OLS model using data from 1988, 1993, and 1998. This result is counterintuitive since if a country protects intellectual property better, then more investors will be attracted to enter into licensing agreements. Yang and Maskus (2001) provides a conjecture. Consider a small nation with a limited skilled labor endowment where imitation risk is slight. Improving IPR protection will further minimize this risk thus lowering licensing costs to the benefit of the investing firm. The monopoly power effect would dominate the economic returns effect. Under this condition, there is less incentive for the firm to enter into more licensing agreements as it can just exploit the monopoly power it has.

The findings of Yang and Maskus (2001) are consistent with that of Gathii (2015) which concludes that extending strong IP protection – with a particular focus on patents – on least developed economies is unlikely to yield the positive economic benefits of stronger FDI flows or higher growth. The author cites the case of China—it experienced massive industrialization and FDI inflows under a less than satisfying intellectual property regime. Therefore, this experience of China argues against the conclusions that strong IP protection is a prerequisite for higher FDI inflows. Strong IP protection is crucial for maintaining the competitive advantages of early industrializers but may not be a crucial determinant for the emergence of new ones (Gathii, 2015).

Meanwhile, an earlier study of Maskus (2000) suggests that stronger IPR protection has short term net welfare losses but dynamic benefits in the longer run. However, these expected benefits in the long run are best likely small and can be easily overrun by short term losses and high cost of implementing laws and policies.

3. Methodology
3.1 Data Sources

The data for this study comes from various sources. The list of developing countries included in the sample took off from the sample used by Adams (2010). Most sources of data are obtained from World Development Indicators Databank of the World Bank. Variables coming from this source include FDI Inflows, GDP growth, Trade Openness, Inflation, Population, Return on Investment, Telephone Subscriptions, and Total IPR. The net FDI inflows are investments to acquire a lasting management interest in an enterprise other than that of the investor. This is obtained by summing equity capital, reinvestment of earnings, other long term capital, and short term capital. GDP growth is the average real GDP growth rate.

Total IPR is obtained by adding all the patents and trademark applications in the country within a certain year made by both residents and non-residents. IPR is scaled by population in order to adjust for the size of the economy since the number of IPR applications may suffer a “home advantage bias” thus; IPR per capita is used as the quantitative measure of IPR (Falk, 2004).

Trade Openness is obtained by getting the value of total trade (exports plus imports) as percent of GDP. Data on inflation
is based on Consumer Price Index (CPI) and is used as an indicator of macroeconomic stability. The natural logarithm of population is used as a proxy for market size. The inverse of GDP per capita is the proxy for return on investment and was used by studies including Asiedu (2002), Adams (2010), and Quazi (2007). The rationale is that the return on investment is a measure of profitability and it should be positively correlated with the marginal product of capital, which is expected to be high in capital-scarce developing countries where per capita income is low (Quazi, 2007). Telephone Subscriptions (per 100 people) is indicative of infrastructure development and access to technology.

The Risk variable is a composite indicator of the overall investment climate of a country. It is obtained from the Political Risk Services’ Country Risk Guide and is made up of three subcomponents—political, financial, and economic risk. It is rated on a scale of 0 to 100 with 100 pertaining to the lowest risk.

The variables IPR Index, Foreign Ownership Restriction Index, and Business Regulation Index came from the Economic Freedom of the World Dataset. The EF index broadly reflects the extent to which a country is pursuing free market principles. The index is constructed by incorporating 50 independent variables. The IPR index used in this study is a sub-index of the composite EF index. Meanwhile, ownership is a measure of the stringency of restrictions on foreign ownership of domestic assets. It is obtained from the Global Competitiveness Report Questions “How prevalent is foreign ownership of companies in your country?” and “How restrictive are regulations in your country relating to international capital flow?” A score of 10 implies that the country is not restrictive at all and a score of 1 implies that the country is highly restrictive. Regulation is an alternative of the “Ease of Doing Business Index” which had only started in 2013. This variable has six subcomponents in the EFW dataset—Administrative Requirements, Bureaucracy Costs, Starting a business, Red tape, Licensing Restrictions, and Cost of Tax Compliance. A score of 10 implies less restrictive regulations and a score of 1 implies stricter regulations are enforced.

It should be noted that most of the countries in the sample (with the exception of only two countries) experienced lower Ownership and Regulation scores from 2000-2014. This trend suggests that most developing countries implemented more restrictive regulations and foreign ownership policies. A possible explanation, as mentioned earlier, is that this period captures the aftermath of the 2008-2009 Global Financial Crisis. Therefore, countries voluntarily enforced tighter capital controls. This limitation of the sample could affect the empirical results of the study (i.e. how they affect FDI inflows).

IPR Index used in the study is another component of EFW. This EFW component came from the Global Competitiveness Report question: “Property rights, including over financial assets, are poorly defined and not protected by law or are clearly defined and well protected by law.” A higher score (as score approaches to 10) implies better property rights protection whereas a lower score (as score approaches to 0) implies a weaker IP regime.

3.2 Estimation Technique
Building upon the IPR-FDI study of Adams (2010) and various researches on FDI dynamics, the relationship between Intellectual Property Rights and Foreign Direct Investment Inflows is an explicit function of the form:

\[
FDI_{it} = f(GDP_{it}, TRADE_{it}, ROI_{it}, LPOP_{it}, RISK_{it}, OWN_{it}, REG_{it}, INF_{it}, TEL_{it}, IPR_{it})
\]
To be able to estimate the parameters of the model, the main qualitative econometric model using IPR protection index is:

$$\begin{align*}
FDP_{it} &= \alpha + \beta_1 FDP_{it-1} + \beta_2 GDP_{GRit} + \beta_3 TRADE_{it} + \beta_4 ROI_{it} + \beta_5 LOP_{it} + \beta_6 RISK_{it} \\
&+ \beta_7 OWNERSHIP_{it} + \beta_8 OWNERSHIPSQ_{it} + \beta_9 REGULATIONSQ_{it} + \beta_{10} REGULATIONS_{it} + \beta_{11} INF_{it} + \beta_{12} TEL_{it} + \beta_{13} IPRI_{it} + \beta_{14} IPRISQ_{it} + \epsilon_{it}
\end{align*}$$

(2)

where: $i$ indicates the country in year $t$, $\alpha$ refers to the constant term, and $\epsilon$ is the composite error term.

The main quantitative econometric model using per capita IPR is:

$$\begin{align*}
FDP_{it} &= \alpha + \beta_1 FDP_{it-1} + \beta_2 GDP_{GRit} + \beta_3 TRADE_{it} + \beta_4 ROI_{it} + \beta_5 LOP_{it} + \beta_6 RISK_{it} \\
&+ \beta_7 OWNERSHIP_{it} + \beta_8 OWNERSHIPSQ_{it} + \beta_9 REGULATIONSQ_{it} + \beta_{10} REGULATIONS_{it} + \beta_{11} INF_{it} + \beta_{12} TEL_{it} + \beta_{13} IPRI_{it} + \beta_{14} IPRP_{it} + \beta_{15} IPRPCS_{it} + \epsilon_{it}
\end{align*}$$

(3)

where: $i$ indicates the country in year $t$, $\alpha$ refers to the constant term, and $\epsilon$ is the composite error term.

The variables ownership, regulation, per capita IPR, and IPR index contained their squared terms to capture any nonlinear relationship with FDI inflows.

Dynamic panel data estimators have been used in many studies dealing with FDIs. Examples of these estimators include Arellano-Bond (1991) and Arellano-Bover (1995)/Blundell-Bond (1998) where dynamic models of the first-differenced equations estimated by Generalized Method of Moments (GMM) are used. Taking-off from the weaknesses of the Arellano-Bond (1991) estimator, Arellano-Bover (1995) and Blundell-Bond (1998) developed a system of regressions in differences and levels. The lagged levels of explanatory variables are used as instruments for regressions in differences while the lagged differences of explanatory variables are used as instruments for regressions in levels.

System GMM is designed for panel data meeting the following characteristics: (1) “Small T, large N” panels (few time periods and many observations), (2) a linear functional relationship, (3) one left-hand side variable that is dynamic, (4) independent variables are not strictly exogenous, (5) fixed individual effects, and (6) heteroskedasticity and autocorrelation within but not across individual observations (Roodman, 2009). Moreover, system GMM is an appropriate technique for handling non-stationary time series, which most macroeconomic variables, particularly those included in this study, are.

Previous studies that use SGMM estimates observe that it is an effective method to correct for the problems of endogeneity of the regressors, measurement error, and omitted variables. This method of estimation can eliminate biases that may arise from ignoring dynamic endogeneity and provides theoretically based and powerful instruments that account for simultaneity while also eliminating unobservable heterogeneity (Alege & Ogundipe, 2014; Davidson & Mackinnon, 2004). SGMM can be estimated in either one step or two-step. Two-step SGMM gives more asymptotically efficient estimates since it uses the consistent variance–co-variance matrix from first step GMM. While it is true that both one-step and two-step estimates are consistent, the latter is more asymptotically efficient. Any bias in the two-step standard errors are corrected by Windmeijer’s (2005) small sample correction incorporated in xtabond2.

It is with these arguments that two step System GMM is selected as the appropriate estimation technique for this study. The Stata syntax to be used for regressions is xtabond2. We transform the above equations (2) and (3) into GMM form. Transforming the main quantitative econometric equation (2) into a more parsimonious form, we obtain:

$$\begin{align*}
FDP_{it} &= \alpha FDP_{it-1} + \delta IPRI_{it} + \gamma IPRISQ_{it} + \beta X_{it} + \beta Z_{it} + \epsilon_{it}
\end{align*}$$

Likewise, transforming the main qualitative econometric equation (3) into a more parsimonious form, we obtain:
where $X_{it}$ is a matrix of control variables including:

- GDP Growth ($GDP_{it}$)
- Trade Openness ($TRADE_{it}$)
- Return on Investment ($ROI_{it}$)
- Natural Logarithm of Population ($LPOP_{it}$)
- Risk: Investment Climate ($RISK_{it}$)
- Ownership and Ownership Squared ($OWNERSHIP_{it}$, $OWNERSHIPSQ_{it}$)
- Regulation and Regulation Squared ($INF_{it}$)
- Inflation ($INF_{it}$)
- Telephone lines per 100 people ($TEL_{it}$),

$Z_{it}$ is a matrix of strictly exogenous regressors, which include time dummies, and $\epsilon_{it}$ is a composite error term including: $u_{it}$, which is time invariant and accounts for any unobservable individual country-specific effects not included in the regression and the stochastic term $v_{it}$.

The marginal effect of IPR on FDI determines how incremental changes in IPR regimes affect FDI inflows. Total marginal effects are obtained by:

\[
\frac{\partial FDI}{\partial \text{IPRI}} = \delta + 2\gamma \ast \text{IPRI} \quad \text{(6)}
\]

\[
\frac{\partial FDI}{\partial \text{IPRPC}} = \hat{\delta} + 2\hat{\gamma} \ast \text{IPRPC} \quad \text{(7)}
\]

where IPRI and IPRPC are the median observations of the variables$^5$

A positive marginal effect implies that an increase in the quantity or an improvement in the quality of protection of IPR increases FDI inflows and a negative marginal effect implies that an increase in the quantity or an improvement in the quality of protection of IPR decreases FDI inflows. While the quantity of IPR generally affects the quality of the IPR regime, there are other factors involved in determining the strength or weakness of IPR regimes. In other words, higher IPRPC values do not necessarily imply higher IPR index scores. A scenario where the marginal effects of equations (6) and (7) are different is possible.

In case nonlinearities are observed, it would be better to obtain the critical point of the equation. To do so, we set equations (6) and (7) to zero and solve for the variable of interest. Considering the possibility that the median observations used in obtaining the total marginal effects of both equations are not its turning points, a different notation is used.

Obtaining the critical value of IPRI from equation 6 yields

\[
\text{IPRI}^* = -\frac{\delta}{2\gamma} \quad \text{(8)}
\]

$^5$ The marginal effects are the values we are mostly interested with since it tells us if a nonlinearity can be observed and where most developing countries are situated vis-à-vis the “U”-shaped curve. We use two measures of central tendency in Section 5 to prove consistency but the median is used as the main benchmark. The median observation is the middlemost entry when observations are arranged in an increasing array.
Likewise, obtaining the critical value of IPRPC from equation (7), we have

$$I_{PRPC}^* = -\frac{\delta}{2\varphi}$$  \hspace{1cm} (9)

### 3.3 Preliminary Data Analysis
The 79 low and middle income countries with average per capita GNI of not more than $12,475 included in the sample are listed below:


Table 2 reports the summary statistics of the variables used in the study including the mean, standard deviation, minimum, and maximum value for all 79 countries within the time period. It can be observed that there are significant variations in variables such as FDI inflows, inflation, and GDP growth where the minimum and maximum values have different signs. Meanwhile, Table 3 reports the correlation among selected variables. An overall consideration of the correlation coefficients suggests that multicollinearity is not a problem. (See Appendix A for Tables 2 and 3).

### 4. Discussion of Results
The models above are estimated using dynamic panel data estimation, more specifically two-step SGMM (Arellano-Bover, 1995; Blundell-Bond, 1998; Windmeijer, 2005) to correct for the problems of endogeneity of the regressors and omitted variables. This method of estimation can eliminate biases that may arise from ignoring endogeneity and provides theoretically based and powerful instruments, using lagged values of the dependent and independent variables, while also allowing the elimination of unobservable heterogeneity (Alege & Ogundipe, 2014; Davidson & Mackinnon, 2004).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Qualitative Model</th>
<th>Quantitative Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreign Direct Investment</td>
<td>.572*</td>
<td>.803*</td>
</tr>
<tr>
<td>Inflows (Lagged)</td>
<td>(.107)</td>
<td>(.071)</td>
</tr>
<tr>
<td>Per Capita IPR</td>
<td>-</td>
<td>-1201.967***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(705.029)</td>
</tr>
<tr>
<td>Per Capita IPR Squared</td>
<td>-</td>
<td>89105.52*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(42821.85)</td>
</tr>
<tr>
<td>IPR Index</td>
<td>2.714*</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(1.279)</td>
<td></td>
</tr>
<tr>
<td>IPR Index Squared</td>
<td>-.234**</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(.119)</td>
<td></td>
</tr>
</tbody>
</table>

---

6 Chile, for the most part of the period of interest, was an upper middle income economy. It was only in mid-2010 when it was classified as a high income OECD country. Its GINI coefficient suggests that economic inequality is still high and its inequality-adjusted HDI is still comparable to other middle income economies in the sample. Uruguay and Venezuela are high income non OECD countries but other indicators suggest that its environment is comparable to that of an emerging economy.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient 1</th>
<th>Coefficient 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP Growth</td>
<td>.257</td>
<td>.617*</td>
</tr>
<tr>
<td></td>
<td>(.173)</td>
<td>(.095)</td>
</tr>
<tr>
<td>Population (in natural logarithm)</td>
<td>-.381</td>
<td>-.136</td>
</tr>
<tr>
<td></td>
<td>(.484)</td>
<td>(.479)</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>.057*</td>
<td>.033*</td>
</tr>
<tr>
<td></td>
<td>(.020)</td>
<td>(.010)</td>
</tr>
<tr>
<td>Return on Investment</td>
<td>8.944*</td>
<td>9.463*</td>
</tr>
<tr>
<td></td>
<td>(2.408)</td>
<td>(3.513)</td>
</tr>
<tr>
<td>Risk</td>
<td>.313*</td>
<td>.271*</td>
</tr>
<tr>
<td></td>
<td>(.110)</td>
<td>(.110)</td>
</tr>
<tr>
<td>Telephone subscriptions (per 100 people)</td>
<td>-.026</td>
<td>.124</td>
</tr>
<tr>
<td></td>
<td>(.086)</td>
<td>(.105)</td>
</tr>
<tr>
<td>Foreign Ownership Restriction Index</td>
<td>-5.381*</td>
<td>-8.648*</td>
</tr>
<tr>
<td></td>
<td>(2.103)</td>
<td>(3.579)</td>
</tr>
<tr>
<td>Foreign Ownership Restriction Index Squared</td>
<td>.388*</td>
<td>.611*</td>
</tr>
<tr>
<td></td>
<td>(.166)</td>
<td>(.295)</td>
</tr>
<tr>
<td>Business Regulation Index</td>
<td>-18.191*</td>
<td>-4.716</td>
</tr>
<tr>
<td></td>
<td>(6.571)</td>
<td>(5.399)</td>
</tr>
<tr>
<td>Business Regulation Index Squared</td>
<td>1.620*</td>
<td>.384</td>
</tr>
<tr>
<td></td>
<td>(.583)</td>
<td>(.491)</td>
</tr>
<tr>
<td>Inflation (GDP Deflator)</td>
<td>.121</td>
<td>-.223*</td>
</tr>
<tr>
<td></td>
<td>(.122)</td>
<td>(.090)</td>
</tr>
<tr>
<td>Period 1</td>
<td>.119</td>
<td>-.061</td>
</tr>
<tr>
<td></td>
<td>(.922)</td>
<td>(1.191)</td>
</tr>
<tr>
<td>Period 2</td>
<td>-.262</td>
<td>-1.08</td>
</tr>
<tr>
<td></td>
<td>(.669)</td>
<td>(.978)</td>
</tr>
<tr>
<td>Period 3</td>
<td>.251</td>
<td>.207</td>
</tr>
<tr>
<td></td>
<td>(.839)</td>
<td>(.975)</td>
</tr>
<tr>
<td>Period 4</td>
<td>-.817</td>
<td>-.430</td>
</tr>
<tr>
<td></td>
<td>(.387)</td>
<td>(.550)</td>
</tr>
<tr>
<td>Period 5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Instruments</td>
<td>40</td>
<td>44</td>
</tr>
<tr>
<td>Countries</td>
<td>46</td>
<td>45</td>
</tr>
<tr>
<td>Observations</td>
<td>211</td>
<td>183</td>
</tr>
<tr>
<td>AR (2) p-value</td>
<td>0.424</td>
<td>0.622</td>
</tr>
<tr>
<td>Hansen test (p-value)</td>
<td>0.575</td>
<td>0.342</td>
</tr>
</tbody>
</table>

**p<0.10. *p<0.05 Standard errors are enclosed in parentheses.
Table 4 above presents the results for both qualitative (1) and quantitative (2) models. Intellectual Property Rights, the variable of interest, is significant in both quantitative and qualitative models but its marginal effect on FDI varies. Per capita IPR (see quantitative model), which is a proxy for the quantity of IPR of a country, has a negative value which implies that countries with more patents, trademarks, and copyrights have lower FDI inflows. In contrast, the IPR protection index (see qualitative model) has a positive value which indicates that strong property rights protection encourages the entry of foreign direct investment in the host country. Moreover, it is important to note that the study observed nonlinearities between IPR and FDI in both models. This result supports the findings of Asid (2004) and contradicts that of Adams (2010). A deeper discussion on the IPR variables is conducted in the subsections below.

4.1 Results using IPR Protection Index
Following all available studies discussing the IPR-FDI link, this paper estimates the IPR-FDI relationship using the conventional approach—that is, using a qualitative index of IPR protection. Table 6 below shows estimation result of the model using IPR Index (IPRI) as a proxy for IPR. This data is a sub-component of the Economic Freedom of the World Composite Index. For comparison purposes, three different techniques—fixed effects (FE), random effects generalized least squares (GLS), and SGMM are used. The qualitative model builds on the model used by Adams (2010). The marginal effect of the IPR protection index on FDI using median and mean is presented first in Table 5 followed by the results comparing three different models in Table 6 (See Appendix A for Table 6).

<table>
<thead>
<tr>
<th>Marginal Effect of IPRI on FDI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using the median IPRI value (=4.50)</td>
</tr>
<tr>
<td>Using the mean IPRI value (=4.521)</td>
</tr>
</tbody>
</table>

The variable of interest, IPR index, has a positive marginal effect on FDI inflows. This is the intuitively expected relationship between IPR and FDI and is consistent with the studies of Adams (2010), Smarzynska (2004), Khan & Samad (2010), Maskus (1998), and Mansfield (1995). The empirical results provide substantial evidence that improvements in the formulation and enforcement of IPR laws and the overall bureaucratic climate of IPR regimes will positively and significantly affect FDI inflows.

Smarzynska (2004) suggests a positive relationship between IPR and foreign investments in a firm-specific sample from high technology sectors. This result is used to induce the argument that IPR protection benefits not just high technology sectors but also firms from other industry because IPR plays a signaling role. Mansfield (1995) also reports a sector-specific study on IPR-FDI link and finds that industries such as chemicals, pharmaceuticals, electrical, etc. benefit most from transfer of technology brought about by FDIs. A region-specific study is conducted by Khan and Samad (2010) where they find that in 14 Asian and South East Asian countries, stronger IPR regimes increase the likelihood of FDI in manufacturing and retail sectors. Meanwhile, Seymour (2006) uses both developed and developing countries and reports a significant positive effect of patent protection on FDI. This study uses developing countries and the empirical results are consistent with those mentioned.

Furthermore, nonlinearity is also observed in the relationship between IPR and FDI and can be validated by the results of the regression in Table 6 (See Appendix A). While Adams (2010) fails to observe this non-linear relationship, Asid (2004) suggests that the patterns of patent protection follow a u-shaped analysis but is unsure about the break point level of the IPR. However, the results of the estimation in this paper suggest that the nonlinearity is such that as the level of intellectual property protection increases, the level of FDI inflows would also increase until it reaches a maximum (inverted “U” shaped). Any further marginal improvements beyond this turning point would be a futile attempt to increase FDI inflows as this policy approach can be considered by investors as being too “restrictive.”

Dr. Anna Klis of Northern Illinois University in the 16th Carroll Round conference suggested that instead of qualitative and quantitative, we use “structure” and “saturation” as alternative model names. However, these terms are unused in previous empirical literature as of this writing which makes the initial names preferably safe for this study.
4.2 Results using Per Capita IPR

Of the available literature on the relationship of IPR-FDI inflows, no study deals with the effect of the sheer quantity of intellectual property right applications of a country on FDI inflows. All studies available use indices such as the Ginarte-Park (GP) index or the Economic Freedom of the World—IPR protection index (EFW). This is under the assumption that if indeed IPR is a factor considered by foreign investors, what matters is how well these rights are protected rather than how much IPR applications there are. However, the quantity of IPR applications can also be indicative of institutional quality since investors would not file applications in countries with weak and underdeveloped IPR regimes.

As Table 4 shows, the coefficient of IPRPC is negative and significant, while the coefficient of its squared value is positive and significant, indicating the existence of a U-shaped relationship between per capita IPR and FDI inflows. This implies that a minimum value of IPRPC has to be reached before per capita IPR can result in higher FDI inflows. Indeed, evaluated at either the median or mean values of IPRPC, we find a negative marginal impact of IPRPC on FDI inflows. See Table 7 below.

<table>
<thead>
<tr>
<th>Table 7. Marginal Effect of Per Capita IPR on FDI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marginal Effect of Per Capita IPR on FDI</td>
</tr>
<tr>
<td>Using the median Per Capita IPR value (=0.00035)</td>
</tr>
<tr>
<td>Using the mean Per Capita IPR value (=0.00081)</td>
</tr>
</tbody>
</table>

Simply put, as patents, trademark, and copyright applications in a country increase, the less FDI inflows are observed. This is true for at least half of the countries in the sample. Again, the nonlinear relationship of per capita IPR and FDI implies that a minimum value must be obtained for the former to positively affect the latter.

The threshold value of IPRPC implied by equation (9) is

$$IPRPC^* = 0.00674,$$

indicating that this is the minimum level of IPRPC that has to be attained for IPRPC to have a nonnegative marginal impact on FDI inflows. While this may seem small an absolute number, this number implies that to positively affect FDI inflows, the saturation of IPR applications has to meet a minimum scale and cannot remain insignificantly small. To further contextualize this critical value, Table 8 below presents the different percentile values of IPRPC. It is only in countries that have IPRPC levels above the 95th percentile that are, on average, able to experience a positive a marginal effect of per capita IPR on FDI inflows.

<table>
<thead>
<tr>
<th>Table 8. Per Capita IPR across different percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per Capita IPR</td>
</tr>
<tr>
<td>IPRPC* (Critical Point)</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Median (50th percentile)</td>
</tr>
<tr>
<td>75th percentile</td>
</tr>
<tr>
<td>90th percentile</td>
</tr>
<tr>
<td>95th percentile</td>
</tr>
<tr>
<td>99th percentile</td>
</tr>
</tbody>
</table>

Figure 2 below situates the obtained negative marginal effect of IPRPC on FDI relative to the median IPRPC observation using a U-shaped curve. From Table 8, all countries with per capita IPR up to the 95th percentile have yet to hit the threshold value. These countries experience lower FDI inflows as per capita IPR increases. However, observations along the 99th percentile begin experiencing higher FDI inflows as per capita IPR increases.
Meanwhile, Table 9 below shows estimation result of the model using per capita IPR (IPRPC) as a proxy for IPR. For comparison purposes, three different techniques—fixed effects (FE), random effects generalized least squares (GLS), and SGMM are used (See Appendix A for Table 9). The coefficients of per capita IPR in these three techniques consistently show its negative effect on FDI inflows with the presence of nonlinearity. In this model, the GMM estimates of the coefficient of the lagged dependent variable satisfies the Fixed Effects < Generalized Method of Moments < Random Effects condition for GMM estimators (Blundell and Bond, 1998).

**4.3. Other Determinants of FDI inflows**

In both models, Lagged FDI Inflows, Risk, Return on Investment, and Trade Openness has positive and significant effects on FDI. Lagged FDI is consistently positive and significant in both models across three different estimation techniques used. The risk variable is significant at 5% (at 1% in the qualitative model) level, which indicates that the political and economic conditions and the overall investment climate of a country is an important determinant of FDI inflows. It should be noted that all countries in the sample are classified as developing countries which makes an improvement in its investment climate all the more significant in attracting FDIs.

The Return on Investment (ROI) coefficient is positive and significant. Theoretically, ROI should be positively correlated with the marginal product of capital which is high among capital scarce developing countries where GDP per capita is low. Therefore, the inverse of per capita GDP, which is ROI, should positively and significantly affect FDI inflows. This result is consistent with studies that are concerned with the determinants of FDI inflows such as Quazi (2007) and Asiedu (2002). Trade openness is another positive and significant variable in both models. As mentioned before, trade openness is the volume of trade as a percentage of GDP. Many developing countries in the sample geared towards export oriented industrialization strategies in the years covered by this study. More specifically, many South American countries embraced import substitution industrialization strategies until the late 1980s (Quazi, 2007). After decades of sluggish economic conditions, they began shifting their philosophy towards export oriented industrialization. Therefore, its full swing effects on FDI inflows is part of what was captured in the positive and significant trade openness coefficient. Indeed, countries that generally change their economic philosophy and turned outwards toward globalization appear favorable and attractive to foreign investors.

Typically, foreign investors refrain from investing in countries that have not had favorable amounts of FDI inflows in the past and, instead, choose to invest their resources on countries that are capable recipients of financial and capital investments. Companies or individuals who plan to test relatively unknown territories will stagger their investment levels until it reaches the desired amount of investment after some time (Quazi, 2007). Therefore, positive incremental lagged changes in FDI levels positively affect FDI inflows—and this is true for both quantitative and qualitative models.
Population is negative but insignificant in both models, which implies that the absolute number of people in a country does not affect FDI inflows. However, studies such as Khan's (2010) which used population growth as a proxy for size of the labor force, showed that population growth positively and significantly affected FDI inflows. GDP growth is positive but is only significant in the quantitative model. A possible explanation is that GDP growth rate becomes a less important determinant of FDI inflows when indices of how property rights are protected in the host country are included in the equation. This result contradicts the findings of Adams (2010) and Nunnenkamp (2003) which both used qualitative indices of IPR protection where it was found that GDP growth positively and significantly affects FDI inflows.

The variable telephone per 100 people is insignificant in both models. This is consistent with the findings of Adams (2010). A significant number of countries in the sample are in the Sub-Saharan African region where the average telephone line subscriptions are 1.5 per 100 people. While insignificant, the summary statistics of this variable is indicative of underdeveloped technology and infrastructure in developing countries.

The coefficient of foreign ownership restriction index is negative and significant. This result is true for both models across three different estimation techniques (See Tables 5 and 8). This is counterintuitive and contradicts the theoretical models that suggest that countries where foreign ownership of companies is allowed and is more prevalent are likely to attract more FDI. This can be seen in the case of China where its gradual opening up from State Owned Enterprises (SOE) to Joint Ventures (JV) and finally to Wholly Owned Foreign Enterprises (WOFE) significantly increased its FDI inflows. A significant coefficient of Ownership suggests that it indeed is a factor considered by foreign investors. Its negative value can be explained by taking a closer look at the trend of the ownership observations. In all but 2 (Sri Lanka and Vietnam) of the 79 countries in the sample, the scores of ownership in the first period (2000-2002) are higher than the last period (2012-2014). Therefore, it can be inferred that the relaxation of foreign ownership environment of the countries in the sample is not captured between 2000 and 2014. Simply put, in all but 2 of the 79 countries in this study, foreign ownership of domestic companies is generally more relaxed and prevalent in 2000-2002 than in 2012-2014. This can explain why the coefficient on ownership is counterintuitive. However, a consideration of the correlation matrix in Table 3 implies that ownership is positively correlated with FDI inflows.

Meanwhile, the value for business regulation index is significant only in the qualitative model and is negative in both. This result is consistent with the findings of Quazi (2007) where he finds that there is a significantly positive correlation between FDI inflows and more repressive regulation. In the study of Quazi, of the nine countries in the sample, only Mexico and Argentina experience less repressive regulations whereas the seven remaining countries in the sample are heavily regulated. An explanation the author provides is that even the seven countries received far higher FDI inflows and that is what is captured by the estimation. Moreover, seven countries scored higher in other FDI-inducing measures which may have offset the heavy regulation scores. Similar to the case of Quazi (2007), majority of the countries in the sample of this study are heavily regulated but the more heavily regulated countries received more FDI inflows. In fact, less than five percent of the sample experienced relaxed business regulation over the fifteen year period. It is possible that the effect of heavy regulations is offset by other FDI inducing incentives. In other words, for both Ownership and Regulation variables, the lack of variation in data must have played a role.

4.4. Diagnostic Tests
The general rule that the number of instruments must be less than or equal to the number of groups is satisfied in both models. Moreover, another important diagnostic test in Dynamic Panel Data estimation is the Arellano-Bond Test for Autocorrelation of the Residuals (AR(x)) where the null hypothesis is "no autocorrelation of order x". The AR (1) results suggest that if the residuals are uncorrelated, their first differences are expected to be correlated so in this case, AR (2) test will better identify serial correlation of the residuals. Based on the figures reported in Table 3, both quantitative and qualitative model showed no signs of first order serial autocorrelation, thus the instruments are good.
Meanwhile, the Sargan and Hansen tests are indicative of the exogeneity of the instruments with the null hypothesis being: "the instruments as a group are exogenous" which should be rejected by obtaining a high p value. In both tests, failure to reject the null hypothesis gives support to the model. An overall consideration of the Sargan and Hansen p-values suggests that the instruments are effective and valid.
Again, the coefficients for both quantitative and qualitative System GMM estimations satisfy the Fixed Effects < Generalized Method of Moments < Random Effects condition for effective GMM estimators (Blundell and Bond, 1998).

5. Policy Implications and Concluding Remarks
The empirical results of the study suggest that strengthening IPR regimes has a positive marginal effect on FDI inflows. However, the relationship of IPR and FDI is not direct. This study finds sufficient evidence that strengthening IPR protection alone does not automatically positively affect FDI inflows. An overall consideration of the investment climate, ownership restrictions, business regulations, and macroeconomic indicators of a country all affect the decision making process of foreign investors. As a matter of fact, FDI inflows surged in China after gradually opening up their foreign business ownership restrictions while maintaining a stably weak IPR regime and a repressive government intervention in corporations.

Moreover, this paper observes a negative marginal effect of per capita IPR on FDI with the presence of nonlinearity. Therefore, the stock of IPR quantity in the countries in the sample has not yet reached the necessary threshold to positively affect FDI inflows. This means that even if the sheer quantity of IPR applications is steadily increasing, it does not translate to higher FDI inflows unless the stock of IPR applications reaches a certain threshold. This critical point is found to be a point between the 95th and 99th percentile observations. Therefore, policy should not only aim to increase IPR applications by attracting more investors to apply for patents, trademarks, and copyrights, but also focus on making sure that these applications are protected against infringement. The strategic policy direction is a simultaneous bureaucratic and institutional reform so that implementing agencies do not have overlapping responsibilities and jurisdictions. FDIs will significantly increase when red tape is eliminated, business set up is simplified, and property rights are well protected. Another policy direction is to ensure that other reforms that generate spillover effects be prioritized. For instance, judicial reform which is badly needed in the Philippines would have spillover effects on the quality of IPR protection. Effectively pursuing this will improve the country’s IPR index without the policy directly intending to do so.

Apart from these determinants, this study suggests that stronger IPR regimes are additional incentives for foreign investors. To strengthen IPR regimes, this study suggests a two-pronged solution: reach the threshold and at the same time improve protection. Improving IP protection without incentivizing individuals and firms to register for intellectual property will not positively affect FDI inflows, as shown in this study. In contrast, simply reaching the minimum stock of IPR would not translate to higher FDI inflows if not accompanied by improvements in institutions that protect these IP rights. Note that the minimum stock of IPR computed in this study will be an ever changing indicator depending on developments in IPR regimes and population growth rates. Further researches should be devoted to looking at industry-specific and firm-specific data (i.e. how IPR quantity and protection quality affect investments in that specific industry).

Again, the countries included in the sample of this study are limited to low and middle income countries with average per capita real GNI of not more than $12,475 and the data gathered covered a total of 15 years from 2000-2014. This sample limit could have contributed to the counterintuitive results on ownership and regulation variables.

Encouraging FDI should not be the end goal of policy, as shown in the conceptual framework. FDI is a tool to generate jobs, make use of idle resources, and benefit from transfers of more efficient technology. All these contribute to economic development. This research provides sufficient empirical evidence that if IPR regimes are strengthened and overall economic climate is improved, developing countries will benefit.
References


World Bank Open Dataset

World Development Report 2004


### APPENDIX

#### Table 2: Summary Statistics of Select Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreign Direct Investment Inflows</td>
<td>4.08988</td>
<td>2.771</td>
<td>5.26432</td>
<td>-5.3214</td>
<td>43.6672</td>
</tr>
<tr>
<td>IPR Index</td>
<td>4.52145</td>
<td>4.5</td>
<td>1.42116</td>
<td>0.95</td>
<td>8.1667</td>
</tr>
<tr>
<td>IPR Index Squared</td>
<td>22.6583</td>
<td>20.25</td>
<td>13.1764</td>
<td>0.905</td>
<td>66.7233</td>
</tr>
<tr>
<td>Per Capita IPR</td>
<td>0.00081</td>
<td>0.00035</td>
<td>0.00165</td>
<td>2.58E-07</td>
<td>0.0123</td>
</tr>
<tr>
<td>Per Capita IPR Squared</td>
<td>3.37E-06</td>
<td>1.21e-07</td>
<td>1.64E-05</td>
<td>6.66E-14</td>
<td>0.00015</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>4.50537</td>
<td>4.571</td>
<td>3.17193</td>
<td>-10.308</td>
<td>21.8945</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>76.4785</td>
<td>68.236</td>
<td>38.6126</td>
<td>0.3327</td>
<td>303.375</td>
</tr>
<tr>
<td>Return on Investment</td>
<td>0.44356</td>
<td>0.2488</td>
<td>0.46562</td>
<td>0.0412</td>
<td>2.46677</td>
</tr>
<tr>
<td>Risk</td>
<td>64.823</td>
<td>64.889</td>
<td>7.43567</td>
<td>37.5208</td>
<td>83.6042</td>
</tr>
<tr>
<td>Foreign Ownership Restriction Index</td>
<td>6.18903</td>
<td>6.3</td>
<td>1.22134</td>
<td>2.75</td>
<td>8.63333</td>
</tr>
<tr>
<td>Foreign Ownership Restriction Index Squared</td>
<td>39.7903</td>
<td>39.69</td>
<td>14.4643</td>
<td>7.5625</td>
<td>74.5344</td>
</tr>
<tr>
<td>Business Regulation Index</td>
<td>5.58283</td>
<td>5.633</td>
<td>0.90839</td>
<td>3.26667</td>
<td>8.4</td>
</tr>
<tr>
<td>Business Regulation Index Squared</td>
<td>31.9907</td>
<td>31.734</td>
<td>10.1622</td>
<td>10.6711</td>
<td>70.56</td>
</tr>
<tr>
<td>Inflation (GDP Deflator)</td>
<td>10.96852</td>
<td>6.588</td>
<td>47.85115</td>
<td>-3.646698</td>
<td>911.634</td>
</tr>
<tr>
<td>Telephone Lines (per 100 people)</td>
<td>7.46043</td>
<td>3.399</td>
<td>8.44707</td>
<td>0.0175</td>
<td>38.3427</td>
</tr>
</tbody>
</table>

#### Table 3: Correlation Matrix of Select Variables

<table>
<thead>
<tr>
<th></th>
<th>FDI</th>
<th>IPRI</th>
<th>IPR_PC</th>
<th>GDP_GR</th>
<th>TRADE</th>
<th>ROI_1000</th>
<th>LPOP</th>
<th>ICRG</th>
<th>OWN</th>
<th>REG</th>
<th>INFGDP</th>
<th>TEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDI</td>
<td>1</td>
<td>0.1468</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>IPRI</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>IPR_PC</td>
<td>0.107</td>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>GDP_GR</td>
<td>0.3352</td>
<td>0.2537</td>
<td>0.1347</td>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TRADE</td>
<td>0.2538</td>
<td>0.3388</td>
<td>0.2321</td>
<td>0.1655</td>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ROI_1000</td>
<td>0.1145</td>
<td>-0.3588</td>
<td>-0.2611</td>
<td>0.058</td>
<td>-0.208</td>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>LPOP</td>
<td>-0.3012</td>
<td>0.0457</td>
<td>-0.5037</td>
<td>-0.0471</td>
<td>-0.2485</td>
<td>-0.0884</td>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ICRG</td>
<td>0.0472</td>
<td>0.5886</td>
<td>0.1325</td>
<td>0.2339</td>
<td>0.4046</td>
<td>-0.4894</td>
<td>0.109</td>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>OWN</td>
<td>0.1694</td>
<td>0.2476</td>
<td>0.036</td>
<td>0.2346</td>
<td>0.1029</td>
<td>0.0288</td>
<td>-0.098</td>
<td>0.2522</td>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>REG</td>
<td>0.2557</td>
<td>0.6228</td>
<td>0.1161</td>
<td>0.1494</td>
<td>0.3307</td>
<td>-0.2197</td>
<td>-0.165</td>
<td>0.3389</td>
<td>0.097</td>
<td>1</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>INFGDP</td>
<td>0.0580</td>
<td>0.019</td>
<td>0.063</td>
<td>0.0311</td>
<td>0.0189</td>
<td>0.4163</td>
<td>-0.150</td>
<td>-0.1141</td>
<td>0.0573</td>
<td>-0.1047</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>TEL</td>
<td>-0.1366</td>
<td>0.3996</td>
<td>0.0684</td>
<td>-0.1633</td>
<td>0.1428</td>
<td>-0.517</td>
<td>0.251</td>
<td>0.4958</td>
<td>0.0265</td>
<td>0.1085</td>
<td>-0.0019</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 6. IPR Protection Index Model Estimation—(1) Fixed Effects Model, (2) Lagged Model, (3) Generalized Least Squares Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreign Direct Investment Inflows (Lagged)</td>
<td>(0.305^*)</td>
<td>(0.572^*)</td>
<td>(0.796^*)</td>
</tr>
<tr>
<td></td>
<td>((0.101))</td>
<td>((0.107))</td>
<td>((0.081))</td>
</tr>
<tr>
<td>IPR Index</td>
<td>(-3.658^*)</td>
<td>(2.714^*)</td>
<td>(0.123)</td>
</tr>
<tr>
<td></td>
<td>((1.138))</td>
<td>((1.279))</td>
<td>((0.860))</td>
</tr>
<tr>
<td>IPR Index Squared</td>
<td>(0.405^*)</td>
<td>(-0.234^{**})</td>
<td>(-0.040)</td>
</tr>
<tr>
<td></td>
<td>((0.123))</td>
<td>((0.119))</td>
<td>((0.089))</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>(0.135)</td>
<td>(0.257)</td>
<td>(0.190^*)</td>
</tr>
<tr>
<td></td>
<td>((0.087))</td>
<td>((0.173))</td>
<td>((0.083))</td>
</tr>
<tr>
<td>Population</td>
<td>(-15.205^*)</td>
<td>(-0.381)</td>
<td>(-0.059)</td>
</tr>
<tr>
<td>(in natural logarithm)</td>
<td>((6.063))</td>
<td>((0.484))</td>
<td>((0.162))</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>(0.035^{**})</td>
<td>(0.057^*)</td>
<td>(0.010)</td>
</tr>
<tr>
<td></td>
<td>((0.018))</td>
<td>((0.020))</td>
<td>((0.007))</td>
</tr>
<tr>
<td>Variable</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Foreign Direct Investment Inflows (Lagged)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return on Investment</td>
<td>-20.132*</td>
<td>8.944*</td>
<td>1.924*</td>
</tr>
<tr>
<td>(3.107)</td>
<td>(2.408)</td>
<td>(.694)</td>
<td></td>
</tr>
<tr>
<td>Risk</td>
<td>-0.037</td>
<td>.313*</td>
<td>-0.008</td>
</tr>
<tr>
<td>(.066)</td>
<td>(.110)</td>
<td>(.042)</td>
<td></td>
</tr>
<tr>
<td>Telephone subscriptions (per 100 people)</td>
<td>.200**</td>
<td>-.026</td>
<td>.015</td>
</tr>
<tr>
<td>(.104)</td>
<td>(.086)</td>
<td>(.039)</td>
<td></td>
</tr>
<tr>
<td>Foreign Ownership Restriction Index</td>
<td>-3.269**</td>
<td>-5.381*</td>
<td>-.561</td>
</tr>
<tr>
<td>(1.957)</td>
<td>(2.103)</td>
<td>(1.514)</td>
<td></td>
</tr>
<tr>
<td>Foreign Ownership Restriction Index Squared</td>
<td>.302*</td>
<td>.388*</td>
<td>.066</td>
</tr>
<tr>
<td>(.149)</td>
<td>(.166)</td>
<td>(.124)</td>
<td></td>
</tr>
<tr>
<td>Business Regulation Index</td>
<td>-1.027</td>
<td>-18.191*</td>
<td>-1.287</td>
</tr>
<tr>
<td>(3.405)</td>
<td>(6.571)</td>
<td>(3.333)</td>
<td></td>
</tr>
<tr>
<td>Business Regulation Index Squared</td>
<td>.147</td>
<td>1.620*</td>
<td>.164</td>
</tr>
<tr>
<td>(3.26)</td>
<td>(.583)</td>
<td>(.321)</td>
<td></td>
</tr>
<tr>
<td>Inflation</td>
<td>.046</td>
<td>.121</td>
<td>-.024</td>
</tr>
<tr>
<td>(GDP Deflator)</td>
<td>(.039)</td>
<td>(.122)</td>
<td>(.031)</td>
</tr>
<tr>
<td>Period Dummies</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Instruments</td>
<td>-</td>
<td>40</td>
<td>-</td>
</tr>
<tr>
<td>Groups</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Observations</td>
<td>211</td>
<td>211</td>
<td>211</td>
</tr>
<tr>
<td>Overall R-squared</td>
<td>9.21</td>
<td>-</td>
<td>53.16</td>
</tr>
</tbody>
</table>

**p<0.10. *p<0.05 Standard errors are enclosed in parentheses.

Table 9. Per Capita IPR Model Estimation— (1) Fixed Effects Model, (2) Lagged Model, (3) Generalized Least Squares Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreign Direct Investment Inflows (Lagged)</td>
<td>.383*</td>
<td>.803*</td>
<td>.736*</td>
</tr>
<tr>
<td></td>
<td>(.110)</td>
<td>(.071)</td>
<td>(.093)</td>
</tr>
<tr>
<td>Per Capita IPR</td>
<td>-1188.468</td>
<td>-1201.967**</td>
<td>-713.682</td>
</tr>
<tr>
<td></td>
<td>(984.801)</td>
<td>(705.029)</td>
<td>(642.407)</td>
</tr>
<tr>
<td>Per Capita IPR Squared</td>
<td>22819.93</td>
<td>89105.52*</td>
<td>12929.6</td>
</tr>
<tr>
<td></td>
<td>(78162.08)</td>
<td>(42821.85)</td>
<td>(54568.54)</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>.262**</td>
<td>.617*</td>
<td>.392*</td>
</tr>
<tr>
<td></td>
<td>(.157)</td>
<td>(.095)</td>
<td>(.128)</td>
</tr>
<tr>
<td>Population (in natural logarithm)</td>
<td>-16.768*</td>
<td>-.136</td>
<td>-.306</td>
</tr>
<tr>
<td></td>
<td>(8.131)</td>
<td>(.479)</td>
<td>(.265)</td>
</tr>
<tr>
<td></td>
<td>Estimate 1</td>
<td>Estimate 2</td>
<td>Estimate 3</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>0.056*</td>
<td>0.033*</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.010)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Return on Investment</td>
<td>-2.066*</td>
<td>9.463*</td>
<td>0.961</td>
</tr>
<tr>
<td></td>
<td>(4.168)</td>
<td>(3.513)</td>
<td>(1.016)</td>
</tr>
<tr>
<td>Risk</td>
<td>-0.044**</td>
<td>0.271*</td>
<td>-0.052</td>
</tr>
<tr>
<td></td>
<td>(0.075)</td>
<td>(0.110)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Telephone subscriptions</td>
<td>0.129</td>
<td>0.124</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>(0.099)</td>
<td>(0.105)</td>
<td>(0.050)</td>
</tr>
<tr>
<td>Foreign Ownership Restriction Index</td>
<td>-2.77</td>
<td>-8.648*</td>
<td>-1.163</td>
</tr>
<tr>
<td></td>
<td>(2.245)</td>
<td>(3.579)</td>
<td>(1.833)</td>
</tr>
<tr>
<td>Foreign Ownership Restriction Index Squared</td>
<td>2.633</td>
<td>0.611*</td>
<td>0.108</td>
</tr>
<tr>
<td></td>
<td>(0.173)</td>
<td>(0.295)</td>
<td>(0.151)</td>
</tr>
<tr>
<td>Business Regulation Index</td>
<td>-9.649*</td>
<td>-4.716</td>
<td>-1.594</td>
</tr>
<tr>
<td></td>
<td>(4.355)</td>
<td>(5.399)</td>
<td>(3.856)</td>
</tr>
<tr>
<td>Business Regulation Index Squared</td>
<td>0.963*</td>
<td>0.384</td>
<td>0.180</td>
</tr>
<tr>
<td></td>
<td>(0.421)</td>
<td>(0.491)</td>
<td>(0.366)</td>
</tr>
<tr>
<td>Inflation</td>
<td>0.028</td>
<td>-0.223*</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.090)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>Period Dummies</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Instruments</td>
<td>-</td>
<td>44</td>
<td>-</td>
</tr>
<tr>
<td>Groups</td>
<td>45</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td>Observations</td>
<td>183</td>
<td>183</td>
<td>183</td>
</tr>
<tr>
<td>Overall R-squared</td>
<td>10.37</td>
<td>-</td>
<td>54.43</td>
</tr>
</tbody>
</table>

**p<0.10. *p<0.05 Standard errors are enclosed in parentheses.
Appendices of the Carroll Round Proceedings
APPENDIX A:
Presentation Schedule

Friday, April 21, 9:00 – 11:50 am

Session 1A
Healy 103
Chair: Robert Cumby (Professor, Georgetown University)

Rafael Schwalb (Columbia University)
The Effects of Large Scale Asset Purchases on Corporate Bond Yields – Drivers & Channels
Discussant: Stefani Kostadinova

Robert Garcia (Universitat Pompeu Fabra)
Why Uncertainty Matters: An Empirical Approach for Spain
Discussant: Gregory Maslak

Stefani Kostadinova (American University in Bulgaria)
Financial Fragility and Contagion in Emerging Economies: A Panel Analysis
Discussant: Robert Garcia

Gregory Maslak (Bowdoin College)
Bank Consolidation and Systemic Risk: M&As During the 2008 Financial Crisis
Discussant: Rafael Schwalb

Session 1B
Healy 104
Chair: Marko Klasnja (Assistant Professor, Georgetown University)

Alyson Matthews (Georgetown University)
Working Women: The Impact of Financial Crises on Female Labor Force Participation Rates in Developing Countries
Discussant: Marcel Schlepper

David Henning (University of Warwick)
Exploring conflict in Nigeria: Does the rapacity effect dominate across local government areas in Nigeria?
Discussant: Jonathan Kaufmann

Marcel Schlepper (University of Warwick)
"Comrade, did you see the latest polls?" A quantitative analysis of the impact of national election polls on vote defection in the European Parliament
Discussant: David Henning

Jonathan Kaufmann (American University)
The Economic Efficacy of Reintegration Assistance for Former Child Soldiers
Discussant: Alyson Matthews

Friday, April 21, 2:00 – 4:00 pm

Session 2A
Healy 103
Chair: Anna Maria Mayda (Associate Professor, Georgetown University)

Michelle Knesbach (Dartmouth College)
The Roots of ISIS Foreign Fighters: Systemic Disadvantages for Immigrants in the Labor Markets of OECD Nations
Discussant: Matthew Carl

Gabrielle Cohen (University of Cape Town)
The Costs of Conflict: Refugee Flows in Sub-Saharan Africa
Discussant: Michelle Knesbach

Matthew Carl (Washington and Lee University)
Goodbye Lenin, Hello Murat? The Effect of Communism on Individual Attitudes Towards Immigration
Discussant: Gabrielle Cohen

Session 2B
Healy 104
Chair: Anna Klis (Assistant Professor, Northern Illinois University)

Ty Greenberg (Georgetown University)
Voting for Credibility: The Political Economy of Sovereign Debt Ratings
Discussant: Grace Kim

Vincent Ramos (University of the Philippines)
On the Effects of Intellectual Property Rights and Foreign Direct Investments in Developing Economies: A System GMM Approach
Discussant: Ty Greenberg

Grace Kim (Georgetown University)
The Effect of Scandal on the Value of Political Connections in South Korean Chaebol
Discussant: Vincent Ramos

Saturday, April 22, 9:00 – 11:00 am

Session 3A
Healy 103
Chair: Olga Timoshenko (Assistant Professor, George Washington University)

Duncan Hobbs (Georgetown University)
What are the Effects of Specific Trade Agreement Provisions on Bilateral Trade Flows?
Discussant: Olivia Bisel and Yuou Wu

Amy Yang (Dartmouth College)
Trade Creation and Diversion in the U.S.-Colombia Trade Promotion Agreement: A Preliminary Analysis
Discussant: Duncan Hobbs

Olivia Bisel and Yuou Wu (Georgetown University)
Transparency and the Effect of Natural Resources on Economic Growth
Discussant: Amy Yang

Session 3B
Healy 104
Chair: Dan Cao (Assistant Professor, Georgetown University)

Alexandra Colyer (Georgetown University)
Pharmaceutical Royalty Licensing Transaction Information Content on Shareholder Wealth: An Event Study
Discussant: Christian Fernando

Shambhavi Tiwari (Columbia University)
An application of the Phillips Curve to the Indian Economy
Discussant: Alexandra Colyer

Christian Fernando (Georgetown University)
A Performance Analysis of Managed Futures
Discussant: Shambhavi Tiwari

Saturday, April 23, 1:00 – 3:00 pm

Session 4A
Healy 103
Chair: Charles Udomsaph (Assistant Professor, Georgetown University)

Tim Maecker (University of Warwick)
Do language and cultural barriers enable firms to price discriminate between countries in the European Union?
Discussant: Grady Killeen

Thomas Segerstrom (Gettysburg College)
County-Level Crime Rates and Sanctuary Jurisdictional Status
Discussant: Tim Maecker

Grady Killeen (Georgetown University)
Attribute-Based Regulations: The Case of Corporate Average Fuel Economy (CAFE) Standards
Discussant: Thomas Segerstrom

Session 4B
Healy 104
Chair: Christopher Griffin (Research Director, Access to Justice Lab, Harvard Law School)

Mathison Clore (Georgetown University)
Why don't countries sue more often at the WTO? An Analysis of Specific Trade Concerns in the SPS Committee
Discussant: Griffin Cohen

Mohammad Ahmad (Georgetown University in Qatar)
Relative Deprivation, Rural-Urban Migration and Rural Inequality in China
Discussant: Mathison Clore

Griffin Cohen (Georgetown University)
Power and Petropolitics: How Oil Prices Affect Clout
Discussant: Mohammad Ahmad
APPENDIX B:
Past Speakers

First Annual Carroll Round
(April 5-7, 2002)
Roger Ferguson, Federal Reserve Board of Governors
Donald L. Kohn, Federal Reserve Board of Governors
Lawrence 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, 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
Francois 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)
Susan C. Athey, Harvard University
Philip I. Levy, American Enterprise Institute
Steven Radelet, Senior Fellow at the Center for Global Development

Eighth Annual Carroll Round
(April 16-19, 2009)
Eric S. Maskin, Princeton University (2007 Nobel Laureate)
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 (2001 Nobel Laureate)

Eleventh Annual Carroll Round
(April 19-22, 2012)
Jonathan Levin, Stanford University
Gene Sperling, Director of 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

Fourteenth Annual Carroll Round
(April 16-19, 2015)
Rajiv Shah, United States Agency for International Development
George Akerlof, Georgetown University (2001 Nobel Laureate)

Fifteenth Annual Carroll Round
(April 21-24, 2016)
Daniel Kaufmann, President and CEO of Natural Resource Governance Institute
Rodney Ludema, Chief Economist at the Department of State and Georgetown University

Sixteenth Annual Carroll Round
(April 20-23, 2017)
Jason Furman, Harvard University Kennedy School of Government
Nobuhiro Kiyotaki, Princeton University
APPENDIX C:  
Former Carroll Round Steering Committees

First Annual Carroll Round  
(April 5-7, 2002)  
Christopher L. Griffin, chair (SFS ’02)  
William Brady (SFS ’02)  
Cullen Drescher (COL ’04)  
Meredith L. Gilbert (COL ’04)  
Joshua Harris (SFS ’02)  
Andrew T. Hayashi (SFS ’02)  
Mark Longstreth (SFS ’04)  
Kathryn Magee (SFS ’02)  
Ryan Michaels (SFS ’02)  
J. Brendan Mullen (SFS ’02)  
Scott E. Pedowitz (SFS ’02)  
Waheed Sheikh (SFS ’04)  

Second Annual Carroll Round  
(April 11-13, 2003)  
Seth Kundrot, chair (SFS ’03)  
Nada Abdelnour (SFS ’03)  
Maria Arhancet (SFS ’04)  
Victoria Bembenista (SFS ’03)  
Michael Callen (SFS ’05)  
Eric Fischer (SFS ’03)  
Daphney Francois (SFS/GRD ’04)  
Meredith L. Gilbert (COL ’04)  
Jeffrey Harris (COL ’03)  
Robert S. Katz (COL ’04)  
Marina Lafferriere (SFS ’06)  
Lu Shi (SFS ’03)  
Stacey Tsai (SFS ’03)  
Robert Wrobel (SFS ’03)  
Erica Yu (COL ’05)  

Third Annual Carroll Round  
(April 15-18, 2004)  
Meredith L. Gilbert, chair (COL ’04)  
Héber Delgado-Medrano (SFS ’06)  
Ryan Fraser (SFS ’04)  
Tetyana Gaponenko (SFS ’07)  
Yunjung Cindy Jin (SFS ’05)  
Sarah Knupp (SFS ’04)  
Robert S. Katz (COL ’04)  
Marina Lafferriere (SFS ’06)  
Alia Malik (SFS ’04)  
Susan Work (SFS ’04)  
Beatka Zakrzewski (SFS ’04)  

Fourth Annual Carroll Round  
(April 22-24, 2005)  
Erica Yu, Chair (COL ’05)  
Jasmina Beganovic (SFS ’05)
Lucia Franzese (SFS ’07)
Dennis Huggins (SFS ’05)
Yun Jung Cindy Jin (SFS ’05)
Jonathan Kirschner (SFS ’05)
Susan Kleiman (SFS ’05)
Yousif Mohammad (SFS ’06)
Amy Osekowsky (SFS ’07)
Daniel Schier (SFS ’05)

Fifth Annual Carroll Round
(April 27-30, 2006)
Marina Lafferriere, chair (SFS ’06)
Irmak Bademli (SFS ’06)
Stephen Brinkmann (SFS ’07)
Héber Delgado-Medrano (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 Brinkmann, chair (SFS ’07)
Lucia Franzese (SFS ’07)
Nicholas Hartman (SFS ’07)
Ian Hinsdale (COL ’09)
Alexander Kostura (SFS ’09)
Jennifer Noh (SFS ’07)
Amy Osekowsky (SFS ’07)
Allison 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)
Lijia 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 Heide, chair (SFS ’09)
James Arnold (SFS ’11)
Amanda B. Delp (SFS ’12)
Henry Gillam (SFS ’10)
Tom Han (SFS ’10)
Anna Klis (SFS ’10)
Daniel Leonard (SFS ’09)
Daniel Lim (SFS ’11)
Arjun Pant (SFS ’09)
Benjamin Simmons (COL ’09)
Ariell Zimran (SFS ’10)

Ninth Annual Carroll Round
(April 22-25, 2010)
Ariell Zimran, chair (SFS ’10)
Michael Counihan (SFS ’11)
Amanda B. Delp (SFS ’12)
Katherine Donato (SFS ’10)
Tom Han (SFS ’10)
Michael Karno (SFS ’11)
Allison Kern (SFS ’10)
Anna Klis (SFS ’10)
Daniel Lim (SFS ’11)
H. Jess Seok (SFS ’12)
Matthew Shapiro (SFS ’11)

Tenth Annual Carroll Round
(April 14-17, 2011)
Amanda B. Delp, chair (SFS ’12)
James Arnold (SFS ’11)
Albert Chiang (SFS ’13)
Malin Hu (SFS ’11)
Katrina Kosner (SFS ’12)
Nancy Lee (SFS ’11)
Doug Proctor (SFS ’12)
Vivek Sampathkumar (SFS ’11)
Monica Scheid (SFS ’11)
Matthew Shapiro (SFS ’11)

Eleventh Annual Carroll Round
(April 19-22, 2012)
Katrina Kosner, chair (SFS ’12)
Albert Chiang (SFS ’13)
Amanda B. Delp (SFS ’12)
Nhaca Le (SFS ’13)
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 ’12)

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, chair (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)

**Fourteenth Annual Carroll Round**
**(April 16-19, 2015)**
Kristen Skillman, chair (SFS ’16)
Natalie Nah (SFS ’15)
Thomas Christiansen (SFS ’16)
Geeva Gopalkrishnan (SFS ’15)
Elle Kang (SFS ’15)
Grace Kim (SFS ’17)
Cheryl Lau (SFS ’16)
Eve Lee (SFS ’15)
Shom Mazumder (SFS ’15)
Morgan Snow (COL ’16)
MaryAnne Zhao (SFS ’16)

**Fifteenth Annual Carroll Round**
**(April 21-24, 2016)**
MaryAnne Zhao, chair (SFS ’16)
Olivia Bisel (SFS ’18)
Audrey Chambers (COL ’19)
Felicia Choo (SFS ’16)
Alexander Colyer (SFS ’17)
Serena Gobbi (SFS ’16)
Elizabeth Johnson (COL ’16)
Grace Kim (SFS ’17)
Duy Mai (SFS ’18)
Harry Rosner (SFS ’18)
Kristen Skillman (SFS ’16)

**Sixteenth Annual Carroll Round**
**(April 21-24, 2017)**
Grace Kim, chair (SFS ’17)
Olivia Bisel (SFS ’18)
Audrey Chambers (SFS ’19)
Tina Cheesman (SFS ’19)
Seventeenth Annual Carroll Round
(April 19-22, 2018)
Olivia Bisel, chair (SFS ’18)
Shine Aung (SFS ’21)
Grady Killeen (SFS ’18)
Arjun Krishnan (SFS ’18)
Joshua Levy (SFS ’20)
Duy Mai (SFS ’18)
Parker Malarkey (SFS ’19)
Meggie Underwood (SFS ’19)
Elinor Walker (SFS ’20)
Jacob Witt (SFS ’20)
Crystal Zhu (SFS ’19)
Appendix D:  
Members of the Advisory Panel  

Meredith Gilbert Ballotta, UCSF Medical Center  
Christopher L. Griffin, College of William & Mary  
Andrew T. Hayashi, The University of Virginia  
Mitch Kaneda, Georgetown University  
Robert S. Katz, Amazon  
J. Brendan Mullen, American College of Cardiology  
Scott E. Pedowitz, Libby Garvey for Arlington County Board  
Erica Yu Wright, Bureau of Labor Statistics
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  
Karla Campbell, The 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 Jelescu, 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 Slezak, 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 Doerspike, Georgetown University  
Benn Eifert, Stanford University  
Adam Engberg, Georgetown University  
Alexandra Fiorillo, Connecticut College  
Eric Fischer, Georgetown University  
Zlata Hajro, Wellesley College  
Samina Jain, Georgetown University  
Avinash Kaza, Stanford University
Eric Kim, The George Washington University
Seth Kundrot, Georgetown University
Lada Kyi, Rice University
Lee Lockwood, Northwestern University
Sunil Mulani, New York University
Holly Presley, Vanderbilt University
Duncan Roberts, University of California, Berkeley
Lu Shi, Georgetown University
Shanaz Taber, Barnard College
Jiang Wei, University of Michigan

Third Annual Carroll Round
(April 15-18, 2004)
Jeffrey Arnold, Dartmouth College
Julia Berazneva, Mt. Holyoke College
Mehmet Cangul, Georgetown University
Richard Carew, The University of Virginia
Ashley Coleman, Vanderbilt University
Dilyana Dimova, Stanford University
Fernando Galeana, Stanford University
M. Blair Garvey, Emory University
Meredith Gilbert, Georgetown University
Adam Greeney, Oberlin College
Asim Gunduz, The University of Virginia
Marc Hafstead, Northwestern University
Andrew Hayashi, University of California, Berkeley
Katherine Howitt, McGill University
Sohini Kar, Columbia University
Josh Lewis, Illinois-Wesleyan University
Alexis Manning, Illinois-Wesleyan University
Sara Menker, Mt. Holyoke College
Elizabeth Mielke, Vanderbilt University
Stratos Pahis, Dartmouth College
Alicja Pluta, Georgetown University
Adam Raymakers, Dalhousie University
Caroline Schmutte, Dartmouth College
Matt Sekerke, John Hopkins University
John Soleanico, Columbia University
Kai Szakmary, Columbia University
Brandon Wall, Yale University
Kenneth Ward, The 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 Lichter, Washington University in St. Louis
Wee Lee Loh, Cornell University
Alice Luo, Duke University
Katharine Mullock, University of Western Ontario
Jose Mustre del Rio, The Ohio State University
Leah Nelson, Georgetown University
Ee Cheng Ong, Wellesley College
Matthew Phan, Columbia University
Nina Rendelstein, Washington University in St. Louis
David Rogier, Washington University in St. Louis
Ana Maria Romero, Illinois-Wesleyan University
Nathan Saperia, Dartmouth College
Bogdan Tereshchenko, Georgetown University
Olga Timoshenko, University of Western Ontario
Tom Vogl, Princeton University
Kenneth Ward, The University of Chicago
Jonathan Wolfson, Washington University in St. Louis
Suzanne Zurkiya, Emory University

Fifth Annual Carroll Round
(April 27-30, 2006)
Sarah Carroll, Stanford University
Ruth Coffman, Georgetown University
Dubravka Colic, Wellesley College
Pratik Dattani, University of Warwick
Jennifer Dawson, Illinois-Wesleyan University
Heber Delgado-Medrano, Georgetown University
Sherri Haas, Illinois-Wesleyan University
Jen Hardy, Georgetown University
Lauren Iacocca, University of California, Los Angeles
Salifou Issoufou, University of Wisconsin - Madison
Stella Klemperer, Brown University
Daniel Kurland, Dartmouth College
Corinne Low, Duke University
Shanthi Manian, Georgetown University
Michael Monteleone, The University of Chicago
John Nesbitt, Georgetown University
Natasha Nguyen, University of California, Berkeley
Oyebanke Oyeyinka, Carleton College
Evgeniya Petrova, Dartmouth College
Emy Reimao, Georgetown University
Svetoslav Roussanov, Columbia University
Vikram Shankar, Georgetown University
Juan Carlos Suarez, Trinity University
Austin Vedder, Dartmouth College
David Wiczer, Carleton College
Geoffrey Yu, Carleton College
Xiaoti Zhang, University of Warwick
Sixth Annual Carroll Round
(April 19-22, 2007)
Matthew Adler, Oberlin College
Marion Aouad, Princeton University
Stephen Brinkmann, Georgetown University
Erik Eggum, University of Warwick
Lucia Franzese, Georgetown University
Tanja Groth, University of St. Andrews
Ashley Halpin, Dartmouth College
Nicholas Hartman, Georgetown University
Adrienna Huffman, Washington University in St. Louis
Abdulla Humaidan, University of Warwick
Mohammad Huq, Georgetown University
Nedko Kyuchukov, Dartmouth College
Zachary Mahone, New York University
R. Priya Mathew, Washington University in St. Louis
Yana Morgulis, The University of Chicago
Jennifer Noh, Georgetown University
Andrew O’Brien Penney, Georgetown University
Jessica Oliveri, Monash University
Matthew Pech, Dartmouth College
Allison Phillips, Georgetown University
Angelica da Rocha, University of Warwick
Sören Radde, University of Bayreuth
Heleri Rande, New York University
Elena Spatoulas, University of Michigan
Yi Sun, Georgetown University
Bennett Surajat, Carleton College
Freddy Tsai, University of British Columbia
David Wolff, Dartmouth College
Jennifer Xi, Dartmouth College
Cynthia Yim, Princeton University

Seventh Annual Carroll Round
(April 17-20, 2008)
Karl Andres, University of Warwick
Cecil Ang, The University of Virginia
Alaina Antonucci, The Pennsylvania State University
Sue Bai, Georgetown University
Marinella Boyadzhiev, Oberlin College
Quentin Brummet, Illinois-Wesleyan University
Brendan Cooper, Carleton College
Gerard DiPippo, Dartmouth College
Stacey Droms, Georgetown University
Varun Dutt, Macalester College
Yasmine Fulena, Georgetown University
Amish Gandhi, University of Warwick
Katherine Gordon, Mt. Holyoke College
Yi Kang, Wesleyan College
Michael Kunkel, Georgetown University
Han Youp Lee, Georgetown University
Claudio LoCascio, Dartmouth College
Olivia Lynch, Georgetown University
Amr Moubarak, The George Washington University
Simone Nitsch, University of Warwick
Saurabh Pant, New York University
Carson Sherwood, University of Western Ontario
Tadashi Shirai, University of Warwick
Dominique Shure, Georgetown University
William Slater, Vanderbilt University
Shyam Sundaram, Brown University
Poh Lin Tan, Princeton University
Dorothy Voorhees, Georgetown University
Kris Walsh, Georgetown University
Monica Yu, Dartmouth College

Eighth Annual Carroll Round
(April 16-19, 2009)
Jennifer Cairns, Calvin College
David Childers, Georgetown University
Vaska Dimitrova, American University in Bulgaria
Rebecca Freeman, Smith College
Georg Graetz, London School of Economics and Political Science
Markus Gstoettner, London School of Economics and Political Science
Arpit Gupta, The University of Chicago
Frederick Haney, New York University
Rebecca Heide, Georgetown University
Gregory Howard, The University of North Carolina at Chapel Hill
Jacqueline Iwata, The George Washington University
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
Hekuran Neziri, American University in Bulgaria
Casey Oswald, Georgetown University
Arjun Pant, Georgetown University
Caitlin Pierce, Dartmouth College
Isra Salim, Macalester College
Keval Sangani, University of Warwick
Pronita Saxena, University of California, Berkeley
Benjamin Simmons, Georgetown University
Maximilian Siriani, 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 Eapen 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 Hull, Wesleyan University
Michael Karno, Georgetown University
Todd Kawakita, Dartmouth College
Allison Kern, Georgetown University
Anna Klis, Georgetown University
Birgit Leimer, New York University
Daniel Lim, Georgetown University
Benjamin Morley, University of Warwick
In Un Flora Ng, Dartmouth College
Katharine Ng, University of San Francisco
Xing Cong Ong, London School of Economics and Political Science
Hang Qian, Dartmouth College
Paul Unanue, Princeton University
Ahmad Wahdat, Oberlin College
Ariell Zimran, Georgetown University

Tenth Annual Carroll Round
(April 14-17, 2011)
Dimitri Avramov, American University in Bulgaria
Daniel Boada, Harvard University
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 Gutman-Kenney, University of Warwick
Malin Hu, Georgetown University
Kilian Huber, London School of Economics and Political Science
Tomas Jagelka, Dartmouth College
Shorena Kalandarishvili, Smith College
Hideto 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 Schwedeber, Maastricht University
Matthew Shapiro, Georgetown University
Zane Silina, Stockholm School of Economics in Riga
Anusuya Sivaram, Georgetown University
David Thomas, University of Oxford
Maximilian Thormann, London School of Economics and Political Science

Eleventh Annual Carroll Round
(April 19-22, 2012)
Madara Bogdane, Stockholm School of Economics in Riga
Paul Byatta, Harvard University
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
Taras Ignashchenko, Lancaster University
Katrina Koser, Georgetown University
Nhaca Le, Georgetown University
Wanyi Li, Macalester College
Elzita Nacheva, American University in Bulgaria
Anastasija Oleinika, Stockholm School of Economics in Riga
Carlo Pizzinelli, Dartmouth College
Thomas Preston, University of Warwick
Doug Proctor, Georgetown University
Julian Richers, Columbia University
Christopher Roth, University of Warwick
Andrea Ruiz, The George Washington University
Kaivan Sattar, New York University
Mark Schmidt, Georgetown University
H. Jess Seok, Georgetown University
Kenichi Shimizu, Soka University of America
Anusuya Sivaram, Georgetown University
Shuo Yan Tan, Georgetown University
Anna Weber, Georgetown University
Edie Wu, Dartmouth College
Qianyi Yang, Macalester College

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 and Political Science
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
Weifen 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
Emily Oehlsen, 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 Trottner, London School of Economics and Political Science
Ilyas Zhukhenov, 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, The University of Virginia
Brian Goggin, Georgetown University
Heather Hedges, Georgetown University
Alyssa Huberts, Georgetown University
Johnny Huynh, Pomona College
Nikhil Kalathil, Oberlin College
Matthew Klein, The University of Chicago
Samsun Knight, Oberlin College
Vincent La, Dartmouth College
Nataliya Langburd, Yale University
Katherine Loosley, London School of Economics and Political Science
Soumyajit Mazumder, Georgetown University
Russell Morton, Princeton University
Jill Ni, Georgetown University
Jonathan Pedde, Dartmouth College
Viktoria Pilinko, Stockholm School of Economics in Riga
Andrei Romancenco, Stockholm School of Economics in Riga
Saugata Sen, London School of Economics and Political Science
Benjamin Shoesmith, The University of North Carolina at Wilmington
Meredith Strike, Georgetown University
Chris Stromeyer, Georgetown University
Rachel Syzmanski, Georgetown University
Josh Walker, Lancaster University

Fourteenth Annual Carroll Round
(April 16-19, 2015)
Levi Boxell, Taylor University
Hameem Races Chowdhury, University of Warwick
Thomas Christiansen, Georgetown University
Mathison Clore, Georgetown University
Chenbo Fang, University of California, Berkeley
Ryan Su-Shien Go, University of California, Berkeley
Aaron Goodman, Dartmouth College
Geeva Gopalkrishnan Georgetown University
Sankalp Gowda, Georgetown University
Thomas Gutierrez, Harvard University
Dora Heng, Cornell University
Samuel Huang, London School of Economics and Political Science
Kenan Jusufovic, London School of Economics and Political Science
Eve Lee, Georgetown University
Michael Lee, The University of Texas at Austin
Karlis Locmelis, Stockholm School of Economics in Riga
Shom Mazumder, Georgetown University
Jonathon McClure, Georgetown University
Michael McGrath, Georgetown University
John McKeon, Boston University
Virginia Minni, University of Warwick
Natalie Nah, Georgetown University
Emily Reeves, Dartmouth College
Lea Rendell, Vassar College
Daniel Roeder, Duke University
Raphael Small, Haverford College
Jack Wiloughby, Duke University
Nancy Wu, Dartmouth College
Yingtong Xie, Macalester College

Fifteenth Annual Carroll Round
(April 21-24, 2016)
Mihaly Abel, University of Warwick
Ahwaz Akhtar, Georgetown University in Qatar
Rachel Anderson, Duke University
Lukas Bolte, London School of Economics and Political Science
Lilia Chobanova, American University in Bulgaria
Felicia Choo, Georgetown University
Griffin Cohen, Georgetown University
Emily Corning, Columbia University
Dashnamjil Enkhbayar, Georgetown University
Michael Gill, University of Warwick
Serena Gobbi, Georgetown University
Elizabeth Johnson, Georgetown University
Anastasiya Kazhar, Stockholm School of Economics in Riga
Stephanie Kestelman, Swarthmore College
Olena Kuzan, Stockholm School of Economics in Riga
Jia Jun Lim, University of Warwick
Omeed Maghzian, Columbia University
Querida Qiu, The University of Chicago
Mason Reasner, Vanderbilt University
Tim Rudner, Yale University
Robert Scales, Dartmouth College
Joycelyn Su, The University of North Carolina at Chapel Hill
Yichuan Wang, University of Michigan
Matthew Waskiewicz, American University
Danny Watson, Georgetown University
Vanessa (Wenyie) Xiao, Vanderbilt University
MaryAnne Zhao, Georgetown University

Sixteenth Annual Carroll Round
(April 20-23, 2017)
Mohammad Ahmad, Georgetown University in Qatar
Olivia Bisel, Georgetown University
Matthew Carl, Washington and Lee University
Mathison Clore, Georgetown University
Gabriella Cohen, University of Cape Town
Griffin Cohen, Georgetown University
Alexandra Colyer, Georgetown University
Christian Fernando, Georgetown University
Robert Garcia, Universitat Pompeu Fabra
Ty Greenberg, Georgetown University
David Henning, University of Warwick
Duncan Hobbs, Georgetown University
Jonathan Kaufman, American University
Grady Killeen, Georgetown University
Grace Kim, Georgetown University
Michelle Knesbach, Dartmouth College
Stefani Kostadinova, American University in Bulgaria
Tim Maecker, University of Warwick
Gregory Maslak (Bowdoin College)
Alyson Matthews (Georgetown University)
Vincent Ramos, University of the Philippines
Marcel Schlepper, University of Warwick
Rafael Schwalb, Columbia University
Thomas Segerstrom, Gettysburg College
Shambhavi Tiwari, Columbia University
Yuou Wu, Georgetown University
Amy Yang, Dartmouth University