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

VOLUME II

Editor-in-Chief:
Amy M. Osekowsky

Assistant Editors:
Stephen A. Brinkmann, Héber M. Delgado-Medrano, Laura E. Stephenson

A Carroll Round Publication
The Fifth and Sixth Annual Carroll Round Steering Committees
Georgetown University, Washington, DC
2007
CONTENTS

Acknowledgements vii
A Brief History of the Carroll Round ix
Introduction: Why I Support the Carroll Round xiii

Papers from the 2006 Keynote Lecturers

The Political and Economic Dimensions of Democratic Governance 3
KEMAL DERVIS
Administrator of the United Nations Development Programme

It's Getting Warmer 9
THOMAS C. SCHELING
University of Maryland, Distinguished Professor of Economics and Public Policy
2005 Nobel Laureate in Economics

Carroll Round Proceedings

Knocking from Within: Contemporary Social Unrest and its Consequences for a Stable China 15
MIKE MONTELEONE
University of Chicago

The Effect of Immigration on Growth in Texas Counties 18
AUSTIN VEDDER
Darouth College

The Senegalese Experience: Rethinking Fertility Theory for Highly Religious Societies 30
CORINNE S. LOW
Duke University

Effects of Minimum Wage Legislation on Employment and Inequality: Evidence from Eastern 67
and Central Europe
DUBRAVKA COLIC
Wellesley College

Can Development Cause AIDS?: Economic Growth and HIV Prevalence in the 92
Developing World
SHANTHI MANIAN
Georgetown University
# Table of Contents

1. Discrimination as a Factor of Development in Ecuador  
   **Natasha Nguyen**  
   *University of California--Berkeley*  
   p. 99

2. Economic Determinants of Dropout Rates in Mexico: A Cross-State Analysis  
   **Heber M. Delgado-Medrano**  
   *Georgetown University*  
   p. 105

   **Vikram Shankar**  
   *Georgetown University*  
   p. 118

4. Divorce Norms, Intrahousehold Bargaining, and Household Outcomes in Rural Ethiopia  
   **Sarah Carroll**  
   *Stanford University*  
   p. 120

5. Economic Development and the Gender Wage Gap  
   **Sherri Haas**  
   *Illinois Wesleyan University*  
   p. 126

   **Maira Emy Reimao**  
   *Georgetown University*  
   p. 137

7. Multilateral vs. Bilateral Aid: Differing Effects on Growth?  
   **Jennifer Hardy**  
   *Georgetown University*  
   p. 143

8. A Natural Experiment on Reference-Dependent Model  
   **Xiaoti Zhang**  
   *University of Warwick*  
   p. 152

9. The Determinants of Participation and Success at Olympic Games: A Cross-Country Analysis  
   **Oyebanke Oyeyinka**  
   *Carleton College*  
   p. 156

10. The Impact of Implicit Pension Debt on Sovereign Credit Ratings  
    **David G. Wiczer**  
    *Carleton College*  
    p. 181
## Table of Contents

Impact of Workers’ Remittances on Developing Countries’ Growth  
**Salifu Issoufou**  
*University of Wisconsin-Madison*  

214

Forecasting Mexico's Inflation: the Effects of an Inflation Targeting Regime  
**Juan Carlos Suárez Serrato**  
*Trinity University*  

237

Measuring the Wealth Effect of Housing Prices: A Regional Approach  
**John Nesbitt**  
*Georgetown University*  

241

The Winter Fuel Allowance and Pensioners: Warming the Home or Cold Comfort?  
**Pratik Dattani**  
*University of Warwick*  

257

Does Money Buy Happiness?  
**Ruth Coffman**  
*Georgetown University*  

277

Appendix A: Fifth Annual Carroll Round Presentation Schedule  

282

Appendix B: Past Speakers  

284

Appendix C: Former Carroll Round Steering Committees  

285

Appendix D: Members of the Advisory Panel  

286

Appendix E: Past Participants  

287

Note on Paper Submissions  

292
ACKNOWLEDGEMENTS

Over the past five years, the Carroll Round has quickly matured and has become recognized as one of the world’s finest forums for intellectual discourse at the undergraduate level. In the field of International Economics, it remains one of a kind. To date, over 100 students have had the chance to present their research as conference participants, and this publication is a testament to their significant scholarly contributions to the field. However, neither the Carroll Round Proceedings nor the Round itself would have thrived so remarkably without the steadfast support of a multitude of individuals and institutions, several of which we would like to acknowledge specifically.

The Sallie Mae Corporation has supported the Carroll Round for five years since its infancy, allowing it to establish its current status.

The Calvin K. Kazanjian Economics Foundation has provided the initial grant for the publication of the Carroll Round Proceedings.

Donors have made generous financial contributions to the Carroll Round Fund in Academic Year 2005-2006, as well as offered their strategic vision for the evolution of the Conference. We give special thanks to Ms. Victoria E. Bembenista, Mr. Mario F. Espinosa, Ms. Meredith L. Gilbert, Mr. and Mrs. John and Anne Kelly, Mr. and Mrs. Kenneth K. Kunkel, Mr. and Mrs. James E. Moore, Mr. J. Brendan Mullen, Mr. Scott E. Pedowitz, and Mr. and Mrs. Yunho Song. In addition, we thank Ms. Marianne Keler for being an advocate for the Carroll Round throughout its infancy. We also owe a great deal of gratitude to Dr. Venilde Jeronimo, Ms. Katerina Kulagina, and Ms. Elizabeth Franzino for their tireless advocacy of our cause.

Carroll Round participants have gone to great lengths to travel to Washington, D.C. and sharing the pinnacle work of their undergraduate education with their peers. We also thank the faculty members at the dozens of universities from which the Carroll Round recruits its participants for their assistance in steering their very best students our way. In particular, we appreciate the extraordinary efforts of Professor Nancy Marion of Dartmouth College and Professor Michael Seeborg of Illinois-Wesleyan University.
ACKNOWLEDGEMENTS

Speakers have shared their wisdom with our conference participants throughout the years. We thank Dr. Kemal Dervis of the United Nations Development Programme and Dr. Thomas Schelling of the University of Maryland for joining us for the 2006 conference.

We thank the offices and individuals at Georgetown and elsewhere who got the Carroll Round off the ground or have sustained it since, contributing their time, expertise, and creativity, especially:

- SFS Office of the Dean: Dean Robert L. Gallucci, SFS
- The Mortara Center: Dr. John Glavin
- Dr. Bernie Cook: Dr. Daniel Powers
- Ms. Meghan Hogge: Ms. Kendra Baity
- Ms. Jennifer Willis

Georgetown faculty members have directly influenced Carroll Round participants by chairing presentation sessions:

- Dr. Douglas Brown: Dr. Jonathan Heathcote
- Dr. Adhip Chaudhuri: Dr. Robin King
- Dr. Phil Cross: Dr. Arik Levinson
- Dr. Raj Desai: Dr. Anders Olofsgard
- Dr. Garance Genicot

Alumni, both past conference participants and steering committee members have formed a community that extends far beyond one weekend in Washington, D.C. We are particularly grateful to those that have remained in touch with the current steering committee, especially Mr. Jonathan Prin, offering their guidance and assistance in recruiting new students to the Round.

Lastly, we thank Dean Mitch Kaneda, the Carroll Round Faculty Advisor, without whom this endeavor would not be possible.
A BRIEF HISTORY OF THE CARROLL ROUND

Whenever I am asked about the history of the Carroll Round, stories about Oxford and the Radcliffe Arms pub usually abound. While there is truth to this aspect of the narrative, the conference’s roots extend firmly and unambiguously to the Georgetown University campus. For it was there that a remarkable team of friends and colleagues assembled and launched the Carroll Round five years ago.

During 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 ever-expanding knowledge of theory and their boundless enthusiasm for learning. Over time, however, I realized the extent to which I was learning from their unique perspectives; their insights often proved more valuable than the content of weekly lectures. I also became acquainted with a group of talented 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 to the London School of Economics, Brendan to the University of Bristol, and Josh, our resident Slavophile, to 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, Andrew, Ryan and I 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 Inklings called the Eagle and Child their intellectual home away from home, we adopted the Radcliffe Arms as our haven. Over pints and pub food, Andrew’s twin passions for game theory and philosophy emerged. The future of monetary policy and development began to vex Ryan’s thoughts, while I hoped to better understand the mechanisms of cooperation, or conflict, underlying international trade institutions.

Meanwhile at Pembroke College, I encountered a group of students from universities
INTRODUCTION

across the country also spending their junior years at Oxford. Although I befriended the other economists in our contingent, I also developed close relationships with the physicists, biologists, literary scholars, and art historians. In the Junior Common Room or over traditional English dinners in the dining hall, we shared stories about life at our respective home 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 these alternative experiences and approaches to scholarship. As the year progressed, I worried that my eventual return to the United States would sever these exciting connections.

One evening at the start of my final term in Oxford, I thought about the importance of this dialogue and my growing affinity for international economics. I harbored a distressing feeling that undergraduates, especially in economics, were not afforded adequate opportunities to present their work in a serious research setting. After all, I always felt privileged when Andrew, Ryan, and my fellow Pembrokians shared their original ideas with me. Thus, I reasoned that undergraduate economists from around the country deserved an event in which they could interact significantly with each other and the professional academic community. In March 2001, I composed the 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.

Unsure of their likely reactions, I invited Andrew and Ryan to join me in this endeavor over pints at the Radcliffe Arms. 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 was finally 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 brought onboard four underclassmen: 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 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 stewardship of the Carroll Round. Also during the fall, Deans Robert Gallucci and Betty Andretta extended moral and financial support, which cemented our institutional sponsorship at Georgetown.

The 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 influence students’ decision to attend? 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 Christine Smith and Jim Patti shared our ideas with generous alumni who responded favorably and pledged individual donations. Little by little, our initial concepts materialized into reality. When the Sallie Mae Fund contributed $10,000 to the Carroll Round, we both gained a lead sponsor and secured the long-term future of the conference.

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 32. Seniors traveled to Washington from as near as the University of Virginia and as far as Stanford University. The Committee was stunned by the enthusiasm expressed by the participants and their home departments. 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 inaugural 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 then 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 then Vice Chairman Roger W. Ferguson and current Governor 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. Lindsey’s speech marked another first—the inaugural Ibrahim Oweiss Lecture in honor of our beloved Georgetown economics professor. Dr. John Williamson of the Institute for International Economics spoke about development issues over a splendid dinner, 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 28 papers were presented over the weekend, including the impressive work of MIT’s Maria Jelescu in “The Role of Hedge Funds in World Financial Crises,” the noteworthy “The World Food Economy to 2050: A Nonlinear Dynamic General Equilibrium, Two sector, Three Factor Endogenous Growth Approach to Long-Term World-Level Macroeconomic Forecasting” from Stanford’s Benn Eifert, Carlos Gálvez, Avinash Kaza,
INTRODUCTION

and Jack Moore, and “The Global Integration of Stock Markets” by Yale’s Fadi Kanaan. Georgetown professors who served as panel discussants later remarked that the quality of some presentations met and surpassed the sophistication of recent graduate-level dissertations. Judging by their comments, the conference brought together some of the best young minds in economics as they approached the frontiers of research in the discipline.

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. Over the past six years, the event has grown in size and scope beyond my initial hopes. The participation of Nobel Laureates John F. Nash, Jr. in 2004 and Thomas Schelling in 2006 marked special peaks in the evolution of the conference, and I hope that over time students from the developing world will be able to attend. I continue to enjoy meeting participants and learning about their research interests. As they share in the excitement of presenting their work and the occasional trepidation of fielding questions, I feel humbled to be among such gifted individuals. In fact, alumni from previous years have advanced to graduate study at Berkeley, Chicago, MIT, Michigan, Oxford, Princeton, Yale, and Wisconsin as well as top government and finance positions around the country. This group of former conference participants has truly grown into a professional and academic network unlike any other for young economists.

I would like to thank the Kazanjian Foundation for their generous support, which made the publication of these Carroll Round Proceedings possible. I also would like to extend my unwavering gratitude to the members of the inaugural Carroll Round Steering Committee without whom this history would have remained fiction. I have great respect and admiration for successive Chairs Seth Kundrot, Meredith Gilbert, Erica Yu, Marina Lafferriere, and Stephen Brinkmann as they assumed leadership of the conference. Other past and present Committee members have tirelessly ensured the success of the conference each year and deserve our appreciation. Finally, 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 join the ranks of distinguished alumni. With his continued collaboration and the eagerness of future Georgetown students, I expect that the next chapter in the history of the Carroll Round will far surpass its exciting beginning.

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

xii
INTRODUCTION

WHY I SUPPORT THE CARROLL ROUND

Over the last several years, a number of my Dartmouth students have participated in the Carroll Round at Georgetown University. Their experience has been uniformly positive. They have returned to the Dartmouth campus re-energized about economics and economic research.

At the Carroll Round, they have the opportunity to present their research on current international economic issues to a group of fellow undergraduates drawn from around the country. They get practice responding to a discussant's critique of their work and fielding questions about their hypotheses and methodologies in the group discussions that follow presentations. They also have the chance to observe the research and presentation skills of other talented students and engage them in further discussion. Some are further honored to have their research papers selected for publication in the Carroll Round Proceedings.

While at the Carroll Round, my students also have the good fortune to meet with important leaders in public service and academics. They are so excited to meet such distinguished economists, such as the now chairman of the U.S. Federal Reserve and two Nobel Prize winners in economics. These formal and informal meetings with key economic players help them gain new insight into the linkages between good economic research and public policy.

Over the three-day conference, my students have fun interacting with interesting, enthusiastic students from other colleges and universities. For a few days, they are able to leave behind the New Hampshire mud season and relish spring in the nation's capital. A number of my returning students also start seeing themselves doing graduate work in economics and public policy. The Carroll Round has thus been an unqualified success on so many levels.

I am enthusiastic about promoting the Carroll Round. I believe in its mission—to provide a forum for academic discourse among undergraduate students in economics and to recognize and reward high-quality undergraduate research. I encourage all college and university professors who teach undergraduate international economics and supervise independent undergraduate research projects to nominate their best students for the Carroll Round.

Students learn economics in a different way when they conduct hands-on, original research—when they form a hypothesis, collect relevant data, test their hypothesis rigorously, and write up their results. But too often the experience ends when the paper is
turned in to the professor at the end of the term. The Carroll Round gives students the chance to experience another important aspect of what economists actually do—to present their work, learn from constructive criticism, and discuss alternative viewpoints with other engaged scholars. Moreover, all future Carroll Round participants will experience the rewards of having their work published in either synopsis or full-length format in the Carroll Round Proceedings, a volume dedicated to the promotion of excellence in undergraduate international economics research.

Nancy Marion
Professor of Economics
Dartmouth College
THE 2006 KEYNOTE LECTURES

Carroll Round Proceedings
THE POLITICAL AND ECONOMIC DIMENSIONS OF DEMOCRATIC GOVERNANCE

Kemal Dervis
Administrator of the United Nations Development Programme

Strong institutions are essential for economic growth and human development. Democratic governance can be understood as the process of building and sustaining those institutions. Holding free and fair elections and supporting democratically elected governments are some of the examples of that democracy-building process. But these constitute only one side of governance—the political side. Democratic governance has an economic dimension, one less visible than ballots and voting machines, but every bit as important.

Democracy is a powerful driver of development, but as is widely recognised, it does not automatically bring effective economic governance and growth. Indeed, history reveals no systematic differences in economic growth between authoritarian and more-democratic systems. Some authoritarian regimes have managed successful growth, while other countries have pursued democratic reforms and been disappointed by their economic development.

Of course, the real question is whether well-performing countries could have grown even faster—and seen more progress in human development—with stronger democratic institutions. Addressing this question requires a counter-factual analysis that is at the heart of much of the current research and academic debate on the relationship between democracy and development. The extensive literature on the subject makes clear that democracy must be accompanied by a broader strengthening of institutions conducive to growth and development. Here, the economic aspects of democratic governance are key.

This analysis does not, however, diminish the importance of the political dimensions of governance and their relationship to economic issues. While there is today wide agreement that efficient allocation of resources (including the mobilization of talent and finance for innovation and entrepreneurial activity) requires markets, it is also clear that markets alone will not suffice. To function well, markets need an enabling state that addresses market failures and more widely distributes market benefits—one that produces outcomes more acceptable to all. Establishing that enabling state and ensuring that it is responsive and accountable to the people is one of the crucial political aspects of governance.
Democratic governance’s political and economic dimensions are very much mutually interdependent: One cannot progress—at least not very far—without the other.

The Latin American experience typifies this symbiotic relationship. Countries in the region have set in motion significant economic and democratic-governance reforms in the last two decades, a wave of change that appears to be gaining momentum still. But there is wide discontent among citizens in a number of countries about the lack of economic benefits and broad-based development progress this has brought, resulting in a backlash against both political and economic reforms.

On the political front, recent opinion polls suggest that slightly over half of Latin Americans—about 55 percent—said that they would accept an authoritarian government if it could solve their economic problems (UNDP 2004a, p.33). Surveys have also indicated that there are more young people who hold views against democracy than who don’t (UNDP 2004b, p.52).

The backlash against economic reforms follows suit. Average support for a market economy in Latin America has dropped markedly in recent years. Recent analysis by Panizza and Yanez (2006) suggests that the discontent with market-oriented reforms in Latin America is driven primarily by disappointment with economic performance particularly among lower income groups. The evidence does not support other possible explanations such as a change in political orientation by the electorate or more activism by opponents of reform. Persistent inequality is an important explanation. In most countries market reforms have failed to be real reforms in the sense of altering the basic socio-economic structure in favour of lower income groups. On the contrary, market-oriented reforms unaccompanied by social reforms have further strengthened income concentration in the hands of the privileged.

Democratic governance, therefore, must deliver political benefits—fair representation, a political voice—as well as economic benefits—notably tangible economic progress for the majority of citizens. This also supports the idea that political and economic governance reforms cannot outpace one another; sustained progress on one front most often depends in large measure on success on the other. In their recent empirical analysis of a sample of 30 developing countries in Eastern Europe and Latin America, Desai and Olofsgard (2006) examined public sentiment about market reforms and found evidence that reforms that limit the discretionary authority of public officials and enable voters to monitor, reward and sanction politicians can boost the legitimacy of market reforms with the public. This finding is consistent with broader analysis of the reforms now taking place in many developing countries.

Reforms generally come in two stages. First-generation reforms like tariff and budget cuts or privatisation depend on the stroke of a pen. Second-generation reforms, meanwhile, require wider and deeper changes in areas like judicial and regulatory systems, as well as ongoing, and often lengthy, technical and political discussions. Many also run into stiff opposition from entrenched interest groups. Strong institutions of democracy can be powerful actors in legitimizing and sustaining such reforms (see for example Krueger 2000; Navia and Velasco 2003).
Thus the crux of the challenge of advancing democratic governance lies in this interaction between politics and economics. Policymakers need to manage a constant balance. This balance appears to be particularly critical in two dimensions: striking a balance between the market and the state; and between technocracy and accountability.

Balancing the Market and the State

When we look at the advance of liberal democracy we see that the development of competitive markets went hand in hand with the growth of strong democratic institutions. The most advanced and many of the rapidly advancing economies today have vibrant markets embedded in strong social and political institutions—“embedded liberalism,” as John Ruggie has referred to it (Ruggie 1982), a grand social bargain in which societies agreed to be subject to markets, as long as the less-desirable outcomes of markets were contained and shared. Thus, economic liberalization was linked inseparably to social and political goals (Ruggie 2003).

Some more extreme advocates of free markets argue that it is not really the business of government to regulate or meddle in economic issues. Most would agree that this is both unrealistic and naive. Privatization has brought mixed results, often failing or under-delivering when market-embedding is absent or inadequate. Efforts to privatize and liberalize the electricity sector within this context has sometimes ended in disaster; the blackouts and power failures in Brazil and California being just two instructive examples. The practice of bringing in private agents seems to be on the rise, and those agents often have the technology and management capacity that is required to implement many projects. But unless the markets in which they operate are regulated, and unless that regulation is really responsive to citizens, privatization probably will not work. And that failure will naturally unleash the kinds of backlash we are seeing in many parts of Latin America, which is echoing all over the world.

Every successful market system tends to be embedded in strong institutions of economic governance. Developing countries can run into trouble if they see an explosion of markets (and a consequent accumulation of wealth in private hands) while the agencies of government remain weak and under-resourced. Such imbalances not only undermine democracy, but also can prove fatal for the markets themselves: A grossly enfeebled state will not be able to uphold the conditions that sustain free and productive economic activity.

Balancing Technocracy and Accountability

From the most advanced to the least developed country, there is then a constant tension between the need to enable policymakers to set economic policies driven by technical knowledge and the obligation of being accountable to the people. Democratic governance requires striking a balance between technocracy and accountability across a range of policy issues, including, but not limited to, trade and industrialization policy, monetary and exchange rate policy, fiscal and budgetary policy, and regulatory policies.

Sound economic policy-making often requires the ability to credibly pre-commit to
certain stability- and growth-enhancing approaches. That is why the most successful modern economies, in addition to a professional civil service, have non-partisan and largely autonomous regulatory bodies and central banks.

We see the soundness of this approach in central banks’ political autonomy—that is, rules that include preventing the bank governor from holding government office and the ability to implement monetary policy without political approval. This autonomy has improved markedly across a sample of developed and developing countries between 1991 and 2003 (Arnone, Laurens and Segalotto 2006). In addition, the laws in all the countries analyzed in this study stipulate that lending to the government, when allowed, would be non-automatic, temporary, and subject to quantitative limits, thus effectively protecting the central bank from having to finance unsustainable and inflationary deficit spending.

As Turkey’s Minister of Economic Affairs and the Treasury, I introduced a new law in Parliament making our central bank autonomous in the same way as many advanced economies have done. However, many of my colleagues found this undemocratic, and of course, I understood this reaction. One has to be careful not to separate the bodies entrusted with formulating and implementing economic policy from the (ideally) democratic and legitimizing political process. Good economic governance must be politically legitimate. Competent and independent economic policymakers should have the power to make decisions based on technical considerations, but there must be some means of holding them to account without curtailing the autonomy that enables them to do their jobs. The developed countries of Europe, North America, and the Pacific Rim have all found ways to balance popular rule with autonomous and semi-technical organs of economic governance. The board members of the U.S. Federal Reserve, for instance, are appointed by the President and confirmed by the Senate, but serve for 14-year terms and can be removed, for cause, only by the President.

There is also growing recognition of the need for balance between independent regulatory agencies—which may or may not be public entities—and democratically elected bodies. Regulatory agencies play a critical role in stabilising overall monetary and exchange-rate policy and banking rules, so sudden political shifts cannot disturb them. Clear rules and independent institutions also often help prevent “state capture” or “regulatory capture”—that is, preventing incumbent groups from raiding the coffers of the state or engaging in rent-seeking behaviour. Equally important, these rules and institutions can help temper the vicissitudes and potential volatility of markets themselves, notably financial markets. Regulatory institutions can also help ensure a level playing field for private actors, and help guarantee that social “goods” (e.g. access for the poor to electricity, water, and telecommunications) are achieved even in a free-market environment.

Politics changes all policy-making and regulatory institutions eventually, if only through the normal replacement of personnel, but that process should be gradual, and it must preserve the credibility of those institutions. The crucial balance between technocracy and accountability, readily taken for granted in developed democracies, is in reality a delicate affair. Those of us who work with much poorer countries should keep this in mind and be on the lookout for two dangers in particular: exposing regulatory bodies too rea-
ly to day-to-day contention, thereby threatening the stability in basic economic policy that development requires, and adopting an overly technocratic approach that too heavily insulates institutions such as central banks from democratic accountability.

**Conclusion**

In essence, democratic governance requires a symbiotic approach: political and economic considerations taken at equal value, mutually supporting and depending on one another in an environment that nurtures both political stability and economic growth—the lynchpins of sustained development.

Advancing democratic governance also involves striking a critical balance on these axes: first, between the market and the state; and second, between technocracy and accountability. In achieving this balance, it is clear that one size does not fit all. How best to proceed will depend on the specific characteristics, stage of development, and aspirations and capacity of individual countries.

Here, we find UNDP’s contribution: connecting countries to knowledge, best practice and resources, and helping them to also develop the long-term capacity needed to follow their chosen course. UNDP also assists policy-makers in explaining to civil society and the general public why a particular policy is chosen, and ensuring transparency in carrying out that policy. Indeed, institutional development in support of democratic governance has become the most important part of UNDP’s work, taking up to 40 to 45 percent of our worldwide programme budget. Much effort is already going into the political side of democratic governance, and it is bringing gratifying results. Our next challenge is to look much more intensely at the economic aspects of democratic governance and the way in which they interact with the political dimensions to change for the better people’s lives.

**References**


DEMOCRATIC GOVERNANCE


It's Getting Warmer

Thomas C. Schelling

_Distinguished Professor of Economics and Public Policy at the University of Maryland_

_2005 Nobel Laureate in Economics_

We’ve had “global warming” for more than a decade—the hottest decade on record worldwide. Is this the “greenhouse effect” that scientists have been warning about, a response to increased carbon dioxide in the atmosphere, or is it some natural, not man-made, climatic change?

The Intergovernmental Panel on Climate Change (IPCC) has cautiously proposed a “discernable” human influence. The IPCC is a cautious body not disposed toward outright conclusions. And actually most of the climate models do not predict the sudden increases in temperatures of recent years.

Something is going on. What does it tell us about the need to curtail, drastically, carbon emissions during the coming century?

The popular guessing game—do we see a greenhouse “signature,” can we identify a clear “signal” in the “noise”?—is probably premature. The history of climate shows that sudden changes of global atmospheric temperature have occurred. There are random or “chaotic” influences on climate. El Nino is an example, volcanic emissions are another. There are anthropogenic (man-made) influences besides greenhouse gases: aerosols of dust and, especially, sulfur emissions can block incoming sunlight. Urbanization can produce “heat islands” that affect temperature estimates. Finally, most of the globe is ocean; the specific heat of water is great relative to air, and the oceans act as a huge cooling reservoir that delays by perhaps decades the appearance of atmospheric warming. The issue is not whether we can discern a signal in the noise, but discern a signal among other signals.

So the recent temperature record is unlikely to be conclusive on the cause of the warming. Greenhouse warming is not clearly established by the temperature record nor is it in any way ruled out. We may see the greenhouse signal clearly in another decade or two. Meanwhile we have to rely on what science can tell us.

Earth is unique in our solar system for its temperature range, and greenhouse gases are to be thanked. Venus has so much greenhouse atmosphere that water can’t exist as liquid, Mars so little that water can’t exist as liquid. That carbon dioxide molecules absorb infrared radiation has been known for a century. The earth’s atmosphere is transparent to most solar radiation; but as the earth, warmed by daylight sun, radiates energy back into space it does so in the infrared part of the spectrum, and the carbon dioxide absorbs some of it
and gets warm.

Carbon dioxide is only one of the gases that have that property. The most important is water vapor, and part of the estimated enhancement of temperature is the positive feedback of warming on absolute global humidity. (Incidentally, actual greenhouses do not produce the "greenhouse effect; they mainly trap the air that is warmed by contact with the ground that is warmed by the sun.)

I find the case for prospective greenhouse warming to be almost beyond doubt. The uncertainties are not mainly about whether it is going to be real, but about the magnitude and speed of warming, about the variegated climate effects—not just "warming" but all the changes in precipitation, humidity, sunlight and clouds, storms, and variations between night and day, summer and winter, polar regions and tropical, mountains and planes, east coasts and west.

In the two major unspecialized scientific journals, Science and Nature, one has to go back a decade or two to find serious doubts about the basic science. Rarely is there such scientific consensus as there is on whether the greenhouse effect is real, even though it cannot yet be incontrovertibly detected in the climate record.

But the uncertainties are daunting: the best the IPCC can do is give us a range of possible warmings for any given increase in carbon dioxide concentration. And the upper bound of that range has been, for two decades, three times the lower bound—an enormous range of uncertainty. On top of that are the uncertainties of what the change in temperature will do to climates around the world. And on top of that are the uncertainties of what those climate changes may do to the worlds we live in, and what people will be able to do to adapt successfully to what change is allowed to occur.

The uncertainty makes it nearly impossible to establish an international target for greenhouse gas containment over the century, but looser international arrangements might be effective. Although economic theory suggests nothing will be effective without "binding" commitments and some enforcement mechanism, there has never been any regime of enforceable commitments of the magnitude required to cope with the greenhouse problem. There has been one international regime in which commitments were undertaken and largely carried out, with no pretense at any enforcement mechanism. That was NATO, to which nations committed troops, money, and real estate on a major scale over several decades. It is the only model I can identify.

Such a regime would have to depend initially on the nations best able to afford serious conversion of their energy sources and uses, namely the already-developed nations. Later, once such a regime had demonstrated its seriousness, major developing nations could be urged to participate, with financial and technological assistance from the nations that can afford to help.

Patience will be required, not despair.

Reprinted from The Wall Street Journal © 2006 Dow Jones & Company. All rights reserved.
CARROLL ROUND PROCEEDINGS
The Fifth Annual Carroll Round
Undergraduate Economics Conference
Knocking from Within: Social Unrest and Its Consequences for a Stable China

Mike Monteleone

University of Chicago

The ancient Chinese philosopher Confucius is known to have said that “a man who does not plan long ahead will find trouble right at his door.” Indeed, from the desks of political scholars to the closed-door sessions of major governments across the globe, many are attempting to plan for the future of East Asia. Recently, the People’s Republic of China has been lauded for recently experiencing unprecedented economic growth, most notably in the past few years. Due to market-oriented reforms instituted by Mao Zedong’s successor Deng Xiaoping in 1978, China’s GDP has quadrupled in the past 25 years. With it, however, a disturbing pattern of rising social unrest and instability can be seen. According to statistics recently released from China’s Ministry of Public Security (MPS), there were approximately 58,000 protests and other incidents of social unrest in 2003, up from a seemingly meager 8,700 in 1991. These incidents have not only grown in number but also in size; the average size of demonstrations reached 52 in 2003, compared to 8 in 1993. Clearly a good part of the Chinese public is agitated – but why? My paper answers this question as well as examines possible consequences on China’s stability and growing international power.

Many China scholars will argue that there is a firm economic basis for the upward trend of social unrest. A myriad of problems associated to China’s ongoing reform of state-owned enterprises has led to massive layoffs, withheld wages, and missing housing and healthcare payments. China’s highly publicized economic growth hides the true nature of development in the Middle Kingdom; in addition to an overall slowing of growth, many coastal provinces have benefited greatly from both domestic and foreign investment, whereas several western provinces on the interior of the country have been neglected. This difference in developmental pace can also been seen through China’s growing economic and social gap between the rural and the urban. Many of those in rural communities have no choice but to become migrant workers in urban areas, working for extremely low pay in grueling conditions. It is not surprising, therefore, that the China as seen from the eyes of, say, a peasant in Xinjiang is quite different than the China of a budding entrepreneur in Shanghai.

Yet after examining similar research and analyzing the MPS data in comparison to
some basic economic indicators, I conclude that economic explanations alone cannot account for the growing unrest. In fact, the number of incidents continued to rise even during the 1993-1995 period, where the rate of economic growth was over 10%. As a result, I argue that social unrest in China is due to a complex matrix of several different factors. In my paper I choose to focus on the aforementioned economic explanations as well as several sociopolitical factors, such as aging political and legal institutions. Some examples include the now-controversial hukou system and the overwhelmed State Letters and Complaints Bureau, which in the first quarter of 2004 reported a 95% increase in complaints. In addition, I look at the rampant corruption found throughout contemporary China, from village governments to factory owners, that leaves many Chinese outraged and subjected to the whims of ambivalent or abusive overseers. I argue that these institutions are no longer efficient (if they ever were) for the masses to express their concerns, and thus more and more dissatisfied Chinese citizens have taken to the streets.

With this in mind, one may place equal importance on the new political atmosphere of China. While adapting market reforms over the years, the Chinese Communist Party (CCP) has gradually allowed progressive political reform. In contrast to Deng’s China of 1989, the China of today is much more open to public protest. The CCP no longer fears “foreign forces” taking advantage of the Chinese populace as much as before, and instead acknowledges the justified complaints of its citizens against corruption and wage inequalities. Positive and engaging government response has, in turn, shown the Chinese people that peaceful protest is an effective channel to air grievances without the fear of Tiananmen-like retribution from the government. All of these factors play a role in the increase of public unrest.

I have already described social unrest in today’s China and have offered both economic and sociopolitical explanations for the phenomenon. The other half of my paper discusses possible implications of rising unrest on China’s stability both internally and in the global community. I argue that the effect on China’s stability is dependent on three main factors: (1) the methods and strategies of the Chinese government’s response to incidents of unrest; (2) the international community’s reaction to the citizens’ plight and the CCP’s response; and (3) the role of the Chinese intellectual community and its involvement in the political campaigns of the lower class.

The appropriateness of the government’s response is crucial to maintaining stability in China. While most incidents are handled peacefully, there are documented incidents where severe force has been used by the People’s Armed Police and strict controls have sometimes been placed on entire villages. Oppressive censorship and radical moves by the government to try to quell protests may set off a chain reaction of unrest, quickly leading to a spiral instability.

Likewise, any Chinese government reaction deemed inappropriate by the West would substantially hurt China in the international community. After Tiananmen, strong criticism from abroad destroyed China’s credibility and resulted in several disabling sanctions that are still in place today, such as the EU’s oft-contested arms ban. Many countries in the West also sympathized with the pro-democracy students, further igniting objection to the
Chinese government. A similar show of force against peaceful protests and/or a sympathetic cause could potentially cause outrage in the international community, damaging China at a time when it is campaigning to show its “peaceful rise” to the world.

Perhaps most importantly, however, is the role of the Chinese intellectual community. I argue that while there may be a sizeable unrest from peasants and the working class, these incidents are relatively easy to control from the government’s standpoint; poor education and the marginalization of their class in Chinese society leaves little threat to Beijing. Yet if students and intellectuals across China took up the working class’ causes, their abilities to effectively disseminate information and write intelligently on theoretical justifications for political or social change would lead to a nationwide campaign that the CCP would most likely struggle to control. This scenario draws historical parallels to the unrest leading to Tiananmen in 1989.

In conclusion, my paper demonstrates that extensive unrest has the potential for debilitating effects on Chinese political and economic stability. The causes for contemporary unrest in China are varied and not easily solved; in some cases, widespread reform is needed for crucial government services. Furthermore, the extent of unrest’s effect on overall stability is determined by three main factors: the Chinese government’s response, international reaction, and the role of influential intellectuals.

The role of unrest and its influence on China’s stability is an important variable in determining the outlook for China and subsequently East Asia. Scholars studying such issues may find that the future of China is indeed in the hands – and protesting mouths – of its people.
THE EFFECT OF IMMIGRATION ON GROWTH IN TEXAS COUNTIES

Austin Vedder
Dartmouth College

ABSTRACT

The study uses Census and BEA data for the 254 counties in Texas to estimate the impact of international immigration on per capita economic growth using a reduced-form model typically used to test for economic convergence across regions. Results from a series of heteroskedasticity-robust OLS regressions suggest that counties with higher levels of international immigration ceteris paribus experienced slower per capita growth from 1990 to 2003. Estimates from a series of two-stage least squares regressions using the proportion of Hispanics in a county’s population and the minimum distance from the county to the most-traversed points of entry along the Texas-Mexico border as instrumental variables confirm these findings, providing evidence against the null hypothesis that international immigration has no material impact on growth. However, due to unique characteristics of the specific region under analysis, the results are not sufficient to make the general claim that immigration has a negative impact on growth.

I. Introduction

Debate over the impact of international immigration on the lives of U.S. citizens has greatly intensified over the past few decades. As legal and illegal immigration into the U.S. from Mexico has exploded, politicians such as former governor of California Pete Wilson and former Presidential candidate Pat Buchanan have staked their careers on decrying the alleged economic and social ills wrought by illegal Mexican immigration.
These politicians have ridden a wave of anti-immigration momentum manifested in the implementation of policies such as the Immigration Reform and Control Act (IRCA) of 1986 and Operation Gatekeeper of 1994 aimed at “militarizing” the U.S.-Mexico border to stem flows of illegal immigrants.

Economists have also proven increasingly more apt to weigh in on the controversial immigration debate. George Borjas has studied the skill levels of immigrants and their impact on the labor market, concluding in a 1999 study (1999, Chapter 4) that recent immigration has led to a decrease in the relative wages of native-born, lesser-skilled workers in the U.S. In contrast, Butcher and Card (1991) conclude that immigration had little impact on cross-sectional wages in the 1980s, even for presumably unskilled workers at the 10th percentile of the wage distribution.

In this study I hope to add to the international immigration (henceforth simply “immigration”) debate not by examining the impact of immigration on the labor market, but its effect on economic growth. Namely, I estimate the impact of immigration on per capita income growth to examine whether immigration has a material effect on average living standards. I test the null hypothesis that immigration does not have a significant effect on growth, implicitly assuming that any negative effects of immigration on growth exactly counterbalance any positive effects. However, after controlling for observable characteristics, I find evidence that immigration might actually negatively impact growth, evidence that becomes stronger when I use instrumental variables estimation to control for the reverse causality possibility that growth impacts immigration.

I first proceed with a brief theoretical justification for why immigration might impact economic growth, explaining potential positive effects and negative effects alike. I then describe the cross-sectional data set of Texas counties that I use to test my hypothesis before explaining the growth regression I use to actually make the estimates. A section discussing the results comes next, followed by a concluding section that draws conclusions, explains shortcomings of the study, and proposes areas for further research.

II. Theoretical Justification for Possible Effect of Immigration on Growth

There are a number of reasons why immigration might be thought to have a positive effect on economic growth. First, labor market competition created by an influx of immigrants intending to work might induce all workers to work harder, increasing the marginal productivity of labor. Furthermore, as Borjas (1999, p. 88) notes, the immigration of less-skilled workers who can conduct less skill-intensive service tasks might free some native workers to devote their time to those activities at which they are most productive, thus exploiting the idea of comparative advantage. Finally, immigrants might bring with them a disproportionate amount of clever ideas and entrepreneurial spirit, possibly leading to economic growth through technological innovation. Borjas (1999, p. 88) lists a number of foreign-born leaders in U.S. business and cites the disproportionate representation of the foreign-born in the population of Nobel laureates at American universities as examples of entrepreneurial and intellectual contributions of immigrants.

Nevertheless, there is one reason in particular why immigrants might be thought to
negatively affect economic growth: lower levels of skill. Borjas’ famously controversial thesis of “declining immigrant quality” (Borjas, 1985) argues that the most recent waves of immigrants possess decreasingly lower levels of human capital relative to native workers, as measured by years of education. If immigrants come to the U.S. with lower skill or human capital levels relative to native workers, the typical immigrant will presumably not be able to add as much to overall economic output as the typical native worker. Since economic growth is typically measured in per capita terms, the typical immigrant will add relatively more to the denominator in the calculation than he will to the numerator, leading ultimately to lower levels of output per capita and growth per capita. Whether or not this potentially negative effect overwhelms the positive effects mentioned in the preceding paragraph is the central question of this study.

III. Data: Cross-Section of 254 Counties in Texas

To establish whether there exists a causal link between immigration and economic growth, I examine data collected between 1990 and 2003 for the 254 counties in the state of Texas. There is one particular disadvantage to limiting the study to the Texas counties in this manner: the limited scope for generalizing the findings. The counties of Texas are located in a particular region of a particular country and have unique demographic make-ups and social institutions relative to other places in the U.S. and the world. Even more importantly, a solid majority (65%) of the foreign-born in Texas counties are immigrants from Mexico, revealing an immigration mix that comes largely from one country. Due to these peculiarities, one would be hard-pressed to make the case that the finding in this paper that immigration negatively impacts economic growth can be generalized for all cases of immigration everywhere in the world.

Nevertheless, the unique features of the Texas geography and immigration mix are the very thing that makes addressing the inherent endogeneity in the problem possible using instrumental variables regression. It is likely that higher growth rates in particular counties spur immigrants to settle there to take advantage of growing economic opportunities in the labor market and elsewhere. This positive effect of growth on immigration makes it difficult to isolate the independent effect of immigration on growth using simple OLS regression. However, since Texas shares a large portion of its border with Mexico and the immigration mix in Texas comes largely from Mexico, it is possible to instrument for immigration between 1990 and 2003 using the proportion of a county’s 1990 population that was of Hispanic descent and the distance of a county from Mexico. It was this possibility that led to my limiting the study to Texas counties.

IV. Estimation Techniques and Explanation of Variables

To estimate the effect of immigration on growth, I employ a growth regression popularized by Robert Barro (see, for example, Barro, 1991). The underlying model extends the idea of the Solow neoclassical growth model. Rather than simply implying that poorer regions will tend to converge to richer ones due to their lower levels of capital stock and
consequent higher marginal products of capital, the model underlying Barro’s regression allows for the phenomenon of conditional convergence, the idea that poor regions will converge to rich ones conditional on the regions being similar in other respects. If the regions differ significantly in relevant areas, they will approach different steady-state levels of capital per worker and will grow at different rates dependant upon their relative distances from those steady-state levels.

I hypothesize that immigration does not significantly affect the steady-state level of capital per worker to which Texas counties converge and hence does not have a material impact on growth. To test this hypothesis, I estimate the equation

$$\text{gr9003}_i = \beta_0 * \text{lgdp90}_i + \beta_1 * \text{immigpop9003}_i + \beta * X_i + \epsilon_i$$

where $\text{gr9003}_i$ is the growth rate of nominal per capita GDP in county $i$ between 1990 and 2003, $\text{lgdp90}_i$ is the natural logarithm of the level of per capita GDP in county $i$ in 1990, $\text{immigpop9003}_i$ measures the average yearly flow of immigration into county $i$ between 1990 and 2003 relative to the average population, $X$ is a matrix of county control characteristics thought to affect either GDP growth or immigration, and $\epsilon_i$ is an error term. Due to the idea of conditional convergence explained in the previous paragraph, I expect $\beta_0$ to be significantly negative, reflecting the tendency for poorer counties to converge to richer ones ceteris paribus. My hypothesis that immigration does not have a material impact on growth suggests that $\beta_1$ should not differ significantly from zero.

The matrix of independent control variables is motivated largely by a county convergence study conducted by Higgins, Levy, and Young (2005). Tables 1 and 2 define and summarize the variables. Contained within the matrix is metro90, a dummy variable taking a value of 1 if a county contained a metropolitan statistical area in 1990 according to the U.S. Census Bureau and 0 otherwise. This variable is expected to carry a positive coefficient due to the network benefits often associated with the clustering of people in metropolitan areas.

Despite this predicted network effect of urbanization, the variables landpop90 and waterpop90, measurements of the land and water area in a county divided by its 1990 population, are also expected to carry positive coefficients due to the benefits associated with individual county citizens having a greater amount of economic resources at their disposal. While this prediction might seem counterintuitive given the previous prediction that counties containing metropolitan areas should grow faster, it is important to remember that the coefficients measure the ceteris paribus effect of per capita land and water area on growth. Individual counties could very well both contain a metropolitan statistical area and have a high level of per capita land area and thereby reap the benefits from both factors on growth.

The matrix also includes variables describing the ages and educational attainment levels of the citizens in each county. age90 and age90sq represent the median age in a given county in 1990 and the square of the median age, respectively. The relationship between age and economic growth is expected to be non-linear since having a greater proportion of workers in their most productive years is expected to translate into higher growth for a given county. Thus, the coefficient on age90 is expected to be positive while that on...
age90sq is expected to be negative, reflecting an expectation of eventually diminishing returns to age. Turning to educational attainment, hsgradpop90 and bachpop90 measure respectively the percentage in 1990 of people 25 years or older in the population with a high school degree or higher and the percentage with a bachelor’s degree or higher. Since human capital attainment is assumed to make people more productive through skill acquisition, both coefficients are expected to carry positive coefficients.

Finally, the matrix contains totgovpop90, a variable measuring the percentage of the population employed by either the local, state, or federal government in 1990. Since economies with large government sectors are thought to be less dynamic and less entrepreneurial due to increased job stability and a consequent decreased incentive to innovate, totgovpop90 is expected to carry a negative coefficient.

V. OLS Results

The estimates from a series of OLS specifications with standard errors corrected for heteroskedasticity are shown in Table 3. Column (1) shows a positive but insignificant correlation between growth and average per capita immigration. However, this relationship quickly becomes negative in column (2) when lgdp90 is added to the regression to control for the initial level of county GDP. As additional control variables sufficient to explain as much as 32% of the cross-county variation in economic growth are added in columns (3) – (5), the negative relationship between growth and immigration becomes even stronger. The coefficient on immigpop9003 in the regression with the most significant results (column (3)) implies that a one percentage-point increase in average net immigration flows relative to the average population (a very large move given that the standard deviation of immigpop9003 is just over a quarter of a percentage-point) is associated with a -0.38 percentage-point decrease in average annual growth per capita, a relationship that is significant at the five-percent level.

Besides the immigration variable, all of the control variables are correlated with gr9003 in the anticipated direction except for waterpop90, whose coefficient is negative but statistically insignificant. The coefficient on lgdp90 is negative and statistically significant at the one-percent level in each specification, providing support for the theory of conditional convergence. Also significant at least at the five-percent level across multiple specifications are metro90 and landpop90, whose positive coefficients suggest that counties with metropolitan areas and counties with abundant land resources available per capita grow faster than other counties. Finally, the marginal effect of totgovpop90 on growth is negative and significant at the one-percent level in the one specification in which it is included, lending some credence to the theory that counties with large government sectors relative to their population grow less dynamically.

VI. Instrumental Variables Method and Results

Since immigration flows are likely to be larger to counties that are growing more rapidly, there is a strong possibility that immigpop9003 is an endogenous variable. If this “reverse causality” scenario holds, the coefficient on immigpop9003 should pick up the
positive relationship between growth and immigration and thus should be biased in the positive direction using standard OLS regression. Moreover, since illegal immigration is unobserved, it is very likely that measurement error is present in immigpop9003, further biasing the results. Isolating the independent effect of immigration on growth requires more advanced regression techniques.

To attempt to account for the endogeneity of immigpop9003, I estimate a two-stage least squares specification. I use the proportion of Hispanics in a county’s 1990 population (hispanic90) and the logged minimum point-to-point distance from a county’s geographic center to any of the eleven most-traversed points of entry on the Texas-Mexico border (ldistance) to instrument for immigpop9003.

hispanic90 is a valid instrument to the extent that it is uncorrelated with unpredicted disturbances to gr9003 and relevant to the extent that it is correlated with immigpop9003. Since the ethnic composition of a region is generally not thought to have a material impact on economic growth, it is reasonable to assume that hispanic90 operates on growth only through its influence on subsequent levels of immigration and is therefore a valid instrument. Furthermore, there is a strong positive correlation of 0.476 between hispanic90 and immigpop9003, suggesting that hispanic90 is a relevant instrument as well. This strong correlation meshes with contemporary network migration theory, which suggests that immigrants tend to cluster in areas where they have strong ethnic ties to the resident population (see Frey and Liaw, 2005). Since over 65% of international immigrants to Texas come from Mexico, it makes sense that counties with higher proportions of Hispanics in their population would receive higher immigration flows.

Similarly, the distance of a county from Mexico is generally not thought to have an independent effect on economic growth, so it is reasonable to assume that ldistance is a valid instrument. Moreover, immigration flows are likely to attenuate with distance from the Mexican border, as initial migration costs and the cost of subsequent travel to Mexico generally increase with distance from the border. The strong negative correlation of -0.358 between ldistance and immigpop9003 suggests that ldistance is a relevant instrument as well. Figure 1 illustrates this negative correlation for the level of Mexican immigrants present in a given county rather than the flow of net immigration. Counties farther from the Texas-Mexico border generally have lower levels of Mexican immigrants in their populations.

Table 4 lists the estimates from a series of two-stage least squares specifications using the same independent variables as those used in the OLS specifications. While the coefficients on the significant variables are similar in magnitude to those from the OLS regressions, the estimated marginal effect of immigpop9003 on gr9003 becomes much larger in magnitude and much more significant using two-stage least square estimation. The estimates in columns (8) – (10) imply that a one percentage-point increase in per capita average immigration flows is associated with anywhere from a 1.4 to 3.2 percentage-point decrease in the growth rate of per capita county GDP. This result suggests that there may indeed have been positive bias on the intimmigpop9003 coefficients from the OLS regressions due to the impact of economic growth on immigration and provides evidence against
the null hypothesis that the impact of immigration on growth is not significant.

**VII. Conclusions and Possibilities for Further Investigation**

Estimates using two different regression techniques and multiple specifications imply that net immigration into Texas counties had a significant negative impact on county economic growth from 1990 to 2003. One might conclude from these results that the negative influences of immigration on growth, possibly stemming from lower levels of human capital attainment among immigrants, overwhelmed the positive influences in Texas over this time period.

Nevertheless, the results are nowhere near sufficient to make the general claim that immigration negatively impacts growth. As stressed earlier, the counties in Texas are located in a specific region and have a particular immigration mix. About 65% of immigrants to the state hail from Mexico, a country whose immigrants to the U.S. have substantially lower levels of educational attainment than the average immigrant (Borjas, 1999, p. 43). It could very well be that the impact of immigration on growth is positive in states or countries whose immigrants have higher levels of educational attainment. Future research could examine a broader scope of geographic areas, looking at other states or countries that might have a mix of immigrants with higher levels of educational attainment. However, such a study would require rethinking how to instrument for immigration to control for reverse causality, as the use of the proportion of the county’s population that is Hispanic and the distance of the county from Mexico is predicated on the fact that a solid majority of Texas immigration comes from Mexico.

It would also be very informative to somehow use the average educational attainment levels of immigrants upon their arrival as a control variable. If this information were available, it would be possible to examine whether levels of human capital drive the effect of immigration on growth. It might even make sense to interact the immigration variable and this human capital variable to examine whether the marginal effect of immigration on growth depends on the human capital attainment levels of immigrants.

Another possibility would be to compare the effect of international migration on growth with the effect of domestic migration, movement from somewhere else in the U.S. to the county. If domestic migrants are thought to have higher levels of human capital than international migrants and the estimated effect of domestic migration on growth is less negative or even positive, then the hypothesis that the level of human capital attainment drives the effect of immigration on growth becomes more defensible.

Finally, it might be prudent for a future scholar to account for time more systematically than does this study. A first-differences regression using a longitudinal data set could control for unobservable heterogeneity due to fixed time effects. Furthermore, specifying a model allowing immigration to have a lagged effect on economic growth might more accurately capture the true effect of immigration, as immigrants and their children might take time to integrate themselves into the economy. Controlling for time more systematically would also allow a researcher to estimate the impact of certain policy changes like IRCA and NAFTA to see whether their impact on immigration influenced the effect of...
immigration on growth.

**References**


### Table 1: Variable Definitions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Period</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>gr9003</td>
<td>Average yearly growth in per capita personal income from 1990-2003</td>
<td>1990-2003</td>
<td>BEA</td>
</tr>
<tr>
<td>gdp90</td>
<td>Level of per capita personal income in 1990 in dollars</td>
<td>1990</td>
<td>BEA</td>
</tr>
<tr>
<td>immigpop9003</td>
<td>Average yearly net international immigration from 1990-2003 divided by average population between 1990 and 2003</td>
<td>1990-2003</td>
<td>Census</td>
</tr>
<tr>
<td>metro90</td>
<td>Dummy Variable: 1 if the county included a metropolitan statistical area in 1990, 0 otherwise</td>
<td>1990</td>
<td>Census</td>
</tr>
<tr>
<td>landpop90</td>
<td>Land area in square miles divided by 1990 population</td>
<td>1990</td>
<td>Census</td>
</tr>
<tr>
<td>waterpop90</td>
<td>Water area in square miles divided by 1990 population</td>
<td>1990</td>
<td>Census</td>
</tr>
<tr>
<td>age90</td>
<td>Median county age in 1990 in years</td>
<td>1990</td>
<td>Census</td>
</tr>
<tr>
<td>hsgradpop90</td>
<td>Percentage of people 25 years or older in the population with a high school degree or higher</td>
<td>1990</td>
<td>Census</td>
</tr>
<tr>
<td>bachpop90</td>
<td>Percentage of people 25 years or older in the population with a bachelor's degree or higher</td>
<td>1990</td>
<td>Census</td>
</tr>
<tr>
<td>totgovpop90</td>
<td>Total number of people employed by the government in 1990 divided by 1990 population</td>
<td>1990</td>
<td>BEA</td>
</tr>
<tr>
<td>hispanic90</td>
<td>Total number of people of Hispanic descent in the population in 1990 divided by 1990 population</td>
<td>1990</td>
<td>Census</td>
</tr>
<tr>
<td>distance</td>
<td>Distance in miles from county centroid to nearest of eleven most-traversed points of entry along Texas-Mexico border</td>
<td>1990-2003</td>
<td>Census, Bureau of Transportation, Personal Calculations</td>
</tr>
<tr>
<td>Variable</td>
<td>Obs</td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----</td>
<td>-------</td>
<td>-----------</td>
</tr>
<tr>
<td>gr9003</td>
<td>254</td>
<td>3.7%</td>
<td>0.9%</td>
</tr>
<tr>
<td>immigrpop9003</td>
<td>254</td>
<td>0.2%</td>
<td>0.3%</td>
</tr>
<tr>
<td>gdp90</td>
<td>254</td>
<td>14793</td>
<td>3374</td>
</tr>
<tr>
<td>metro90</td>
<td>254</td>
<td>0.19</td>
<td>0.40</td>
</tr>
<tr>
<td>age90</td>
<td>254</td>
<td>34.4</td>
<td>4.5</td>
</tr>
<tr>
<td>hsgrad90</td>
<td>254</td>
<td>64.2%</td>
<td>9.1%</td>
</tr>
<tr>
<td>bach90</td>
<td>254</td>
<td>12.9%</td>
<td>5.3%</td>
</tr>
<tr>
<td>totgov90pop</td>
<td>254</td>
<td>8.0%</td>
<td>3.2%</td>
</tr>
<tr>
<td>hispanic90</td>
<td>254</td>
<td>23.8%</td>
<td>22.9%</td>
</tr>
<tr>
<td>distance</td>
<td>254</td>
<td>473.2</td>
<td>148.6</td>
</tr>
</tbody>
</table>
**Table 3: OLS Regression**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gr9003</td>
<td>0.105</td>
<td>-0.129</td>
<td>-0.377</td>
<td>-0.337</td>
<td>-0.39</td>
</tr>
<tr>
<td>immigpop9003</td>
<td>(0.217)</td>
<td>(0.169)</td>
<td>(0.154)*</td>
<td>(0.204)</td>
<td>(0.198)</td>
</tr>
<tr>
<td>lgdp90</td>
<td>-0.015</td>
<td>-0.02</td>
<td>-0.023</td>
<td>-0.024</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)**</td>
<td>(0.003)**</td>
<td>(0.004)**</td>
<td>(0.004)**</td>
<td></td>
</tr>
<tr>
<td>metro90</td>
<td>0.007</td>
<td>0.006</td>
<td>0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)**</td>
<td>(0.002)**</td>
<td>(0.002)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>landpop</td>
<td>0.006</td>
<td>0.006</td>
<td>0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)*</td>
<td>(0.003)*</td>
<td>(0.002)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>waterpop</td>
<td>-0.012</td>
<td>-0.01</td>
<td>-0.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.008)</td>
<td>(0.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>age90</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age90sq</td>
<td>-0.000023</td>
<td>-0.000016</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000018)</td>
<td>(0.000016)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hsgad90</td>
<td>0.013</td>
<td>0.014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.012)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bach90</td>
<td>0.001</td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.014)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>totgov90pop</td>
<td>-0.059</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.021)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.037</td>
<td>0.183</td>
<td>0.225</td>
<td>0.222</td>
<td>0.246</td>
</tr>
<tr>
<td></td>
<td>(0.001)**</td>
<td>(0.033)**</td>
<td>(0.025)**</td>
<td>(0.032)**</td>
<td>(0.031)**</td>
</tr>
<tr>
<td>Observations</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.28</td>
<td>0.28</td>
<td>0.29</td>
<td>0.32</td>
<td></td>
</tr>
</tbody>
</table>

* significant at 5%; ** significant at 1%

Notes: The time period is 1990-2003. The sample is the 254 counties in the state of Texas. All data come from the U.S. Census Bureau and the Bureau of Economic Analysis. gr9003 is the dependent variable in all regressions. See Tables 1 and 2 for detailed variable sources, definitions, and summaries.
**Table 4: IV Regressions**

<table>
<thead>
<tr>
<th></th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gr9003</td>
<td>0.515</td>
<td>-0.669</td>
<td>-1.416</td>
<td>-3.158</td>
<td>-2.864</td>
</tr>
<tr>
<td></td>
<td>(0.460)</td>
<td>(0.458)</td>
<td>(0.445)**</td>
<td>(1.180)**</td>
<td>(1.089)**</td>
</tr>
<tr>
<td>lgdp90</td>
<td>-0.016</td>
<td>-0.022</td>
<td>-0.018</td>
<td>-0.019</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)**</td>
<td>(0.003)**</td>
<td>(0.004)**</td>
<td>(0.004)**</td>
<td></td>
</tr>
<tr>
<td>metro90</td>
<td>0.009</td>
<td>0.007</td>
<td>0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)**</td>
<td>(0.002)**</td>
<td>(0.002)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>landpop</td>
<td>0.006</td>
<td>0.006</td>
<td>0.008</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)**</td>
<td>(0.001)**</td>
<td>(0.001)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>waterpop</td>
<td>-0.011</td>
<td>-0.015</td>
<td>-0.014</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.011)</td>
<td>(0.010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>age90</td>
<td>-0.002</td>
<td>-0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age90sq</td>
<td>0.000016</td>
<td>0.000019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000026)</td>
<td>(0.000025)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hsgrad90</td>
<td>-0.034</td>
<td>-0.026</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.022)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bach90</td>
<td>0.047</td>
<td>0.055</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.026)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>totgov90pop</td>
<td></td>
<td>-0.072</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.022)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.036</td>
<td>0.195</td>
<td>0.252</td>
<td>0.271</td>
<td>0.293</td>
</tr>
<tr>
<td></td>
<td>(0.001)**</td>
<td>(0.026)**</td>
<td>(0.026)**</td>
<td>(0.043)**</td>
<td>(0.042)**</td>
</tr>
<tr>
<td>Observations</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

**Notes:** The time period is 1990-2003. The sample is the 254 counties in the state of Texas. All data come from the U.S. Census Bureau, the Bureau of Economic Analysis, and the Bureau of Transportation. gr9003 is the dependent variable in all regressions. See Tables 1 and 2 for detailed variable sources, definitions, and summaries. hispanic90, a variable measuring the proportion of Hispanics in a given county’s population in 1990 relative to the county’s total population in 1990, and distance90, a variable measuring the minimum logged distance from the county to any of the eleven most-traversed points of entry along the Texas-Mexico border, are used as instruments for immigpop9003 in a two-stage least squares estimation procedure. The estimates from the second stage regression are listed above in Table.
THE SENEGALESE EXPERIENCE:
RETHINKING FERTILITY THEORY FOR HIGHLY RELIGIOUS SOCIETIES

Corinne S. Low
Duke University

ABSTRACT
Despite improvements, traditional fertility theory still remains unprepared to cope with developing countries, such as Senegal, where deep religious beliefs dictate a passive acceptance of natural fertility. Because of an unwillingness to use modern contraception, factors that can reduce fertility in these societies will primarily be factors that influence natural fertility. Particularly, my study finds that age at first marriage, cultural taboos against sex while breastfeeding, living with extended families, and extended periods of breastfeeding can all reduce family size. Education is found to increase fertility at low levels because it increases fecundity, but reduces fertility at higher levels. It also acts through a multitude of indirect pathways, clearly modeled for the first time in this paper.

I. Introduction
Sub-Saharan Africa has long been the locus of the debate over population policy. The debate has swung back and forth between questions of the best way to control population growth to whether curbing population growth should be a goal at all. At the close of the last century, most theorists had settled on the conclusion that, while a decrease in population growth often came with development, reduced population growth was not necessarily a precursor or determinant of development. The recent development and refinement of the “convergence model” of demographic effects on growth shows that excess population may have some effect on the rate of productivity gains (see Kelley and Schmidt 2004), but there is little direct correlation between population growth rate overall and GDP growth. Rather, lower population growth can be seen as one component in a system of “virtuous circles” (Birdsall and Sinding 2001) that reinforce one another. Development leads to better education, which leads to lower fertility, which leads to higher productivity, which strengthens gains in development.

But whether one accepts or rejects the idea that population affects growth, it seems
clear that for some families and in some areas, excess fertility constrains resources and has negative effects on welfare. Returning to the macro view, we can see that at the core of the questions over population is really a question about development and well-being. When the policy question of decreasing population growth arises, the real question is, “How can we increase our country’s development and the well-being of our people by curbing population growth?” Countries that enact some kind of population policy are implicitly accepting the idea that development and well-being can be increased by population growth reduction. But if the aim is increased welfare, focusing solely on macro population reduction may not be the most efficient strategy.

To meet its aims of increasing well-being, it seems that the true direction that population research should take is targeted population growth reduction in areas where large family size impedes development or strains resources. Taking this one level further, the true goal is not population reduction, but family-size reduction for families unable to support large numbers of children. If population targets are met because the well-off reduce their family sizes further while the poorest members of society continue to have insupportably large families, such a policy can hardly be considered a success. The policies that enact targeted change in family sizes among disadvantaged populations may be very different from the policies that are expected to curb population overall, and will certainly involve a careful reexamining of many of the assumptions about limiting population growth.

Senegal makes an ideal case study to examine these factors because it is a society in which high fertility has been largely unaffected by the family planning programs implemented over the past quarter-century. The problem is that the policies currently in place to reduce fertility focus on education about and provision of modern contraception, a program that ignores Senegal’s specific religious and cultural factors in favor of a one-size-fits-all solution.

There is already a broad literature on fertility reduction, but when it comes to the developing world, this literature is deficient in a number of ways. Much of the literature relies on the assumption that people in the developing world want fewer children, but are unable to meet these targets due to lack of information about contraception. However, in many developing countries desired fertility is actually higher than natural fertility, and contraception use is not adopted despite widespread knowledge. This has led researchers such as Easterlin (1978) and Kelley (1982) to conclude that these societies are actually supply constrained, and therefore not seeking to limit family size at all. However, even the literature that acknowledges that families in developing countries are not simply unaware of or unable to access contraception, but are rather consciously declining to use it, subscribes to a narrow explanation of this behavior. As Pritchett (1994) asserts, “To a striking extent the answer to why actual fertility differs across countries is that desired fertility differs. In countries where fertility is high, women want more children.” However, this perspective is at odds with the plethora of evidence that in many highly religious developing countries, women express that family size should be left for God to decide. Because so many women in many developing countries do not claim agency over family size, attributing large family sizes to large family aspirations is an incomplete picture of
complex cultural realities.

While the modeling techniques appropriate for a supply-constrained framework may function similarly well for women who do not claim agency over family size—in both situations actual fertility is expected to equal natural fertility—the implications will be markedly different. Women who do not use contraception because of large family-size aspirations could be persuadec to lower their family-size targets and adopt modern contraception if their opportunity cost increased, if they better understood the costs of child-rearing, or if they were made aware of alternative sources of satisfaction. On the other hand, women who steadfastly refuse to “tamper” with natural fertility because of religious reasons will not be affected by such initiatives. Rather, for these women what must be examined are the factors that impact natural fertility, such as marriage duration.

A refusal to adopt family planning methods does not mean that family size and, indeed, population growth cannot be reduced as development progresses. What it means is that researchers must come up with new ideas for how to approach fertility in a framework without consciously demanded family size targets. By focusing on the determinants of natural fertility, such as marriage duration and frequency of intercourse, and in turn their determinants, such as age at first marriage and living situation, it is possible to identify avenues through which fertility can be impacted without conscious agency over family size ever being expressed.

Thus, this paper fills a gap created by the prevailing notion in the literature that women who bear many children must either be uninformed about contraception or seeking large families. For many women in highly religious countries such as Senegal, neither framework is entirely appropriate. Rather, many women do not express conscious agency over family size and therefore experience actual fertility equal to natural fertility. My research, therefore, looks for factors that can reduce family size without agency and factors that may spur women to develop agency over family size. I use modeling techniques similar to earlier research, but focus on interpreting my results for women who do not claim agency over family size.

To examine and test these ideas, I break down the determinants of fertility in Senegal, a highly religious, high-fertility society. Senegal, a former French territory, is 94% Muslim. Despite widespread knowledge of contraception methods, few residents have adopted its use. Instead, 20% of Senegalese women surveyed by the 1997 Demographic and Health Surveys expressed a nonnumeric preference for family size, for example saying that it should be left “up to God.”

Because most women in Senegal do not claim agency over family size, policies that continue to promote the use of contraception will likely have a minimal effect. Moreover, while these programs may lower population overall, in countries where religion is a strong factor, it is unlikely to reduce the family sizes of the households that need it most—poor,
rural families who also tend to have the strongest religious beliefs.

Therefore, my study models fertility in Senegal with the specific goal of breaking down inputs into factors that require agency over family size (such as the use of contraception) and those that do not (such as age at first marriage). I also improve on previous literature by paying special attention to how female education affects women who do and do not claim conscious agency over family size. Few researchers have specifically broken down the effects of education into its components that require agency and those that do not. For women who do have specific fertility goals, education can affect actual fertility by altering opportunity cost and awareness of alternative sources of satisfaction, lowering family size goals and potentially spurring family planning. For women who lack this agency, however, education must be thought of in terms of its effects on cultural norms that impact natural fertility. The four main factors that are expected to impact fertility without conscious agency are: age at first marriage, duration of breastfeeding, cultural taboos regarding sexual practices, and living arrangements. All four affect exposure to pregnancy, and thus natural fertility. Depending on education’s effects on these four variables and other factors that influence natural fertility, education may increase or decrease actual fertility for women who do not express agency. To help sort out the myriad effects of education on women acting within different fertility frameworks, I develop a new model that explicitly sorts the effects of education by those whose impacts require women exercising conscious agency over family size and those whose impacts do not.

My findings are that later age at first marriage, longer duration of breastfeeding, abstaining from sex while breastfeeding (a cultural taboo), and living with extended families all decrease fertility. Policies that can affect these variables may, therefore, be able to curb Senegal’s high fertility without requiring the use of modern contraception, or even conscious family planning. As for education, the net effect is ambiguous: Primary education appears to increase fertility, while secondary and higher education reduces it. This positive effect of primary education is found to be strongest in rural areas. However, when a measure of natural ability to supply children, or fecundity, is controlled for, the effect largely disappears. This leads me to conclude that the positive effect of primary education is largely because it increases health, and thus fecundity. Primary education may also increase fertility in rural areas due to its negative effect on breastfeeding, which naturally delays pregnancy. Contrary to expectations, however, education does not reduce adherence to cultural taboos against sex while breastfeeding, and actually increases instances of living with extended families. Moreover, education is found to increase age at first marriage, which in turn leads to lower family sizes.

To summarize, my paper will improve on earlier work by acknowledging that, in certain societies, women may choose not to interfere with natural fertility for religious reasons. However, my case study of Senegal suggests that fertility can still be reduced in these societies by attempting to influence factors that affect natural fertility, such as age at first marriage. Education may be a key policy to enact some of these changes, although it also carries the unintended consequence of increasing fertility due to increased health. On net, however, education appears to be a positive policy alternative to misdirected and inef-
fective family planning programs.

Section II of my paper will be a brief overview of fertility theory, with special attention to its deficiencies when it comes to highly religious countries. It also explains the evidence that Senegal is operating in a framework where family size decisions are not made consciously, and therefore will not be—and have not been—affected by family planning programs. Section III will review some of the models for fertility that have been developed over the past thirty years, and highlight which features of these models fit the Senegalese framework. It also includes a special focus on the complex ways education is expected to interact with fertility. Section IV reviews the data used in this study, the 1997 Demographic and Health Survey for Senegal, and specifies my dependent and independent variables. Section V reports the results of my analysis in two sections: one that explores the proximate determinants of fertility and one that explores education’s effects on these factors. Section VI concludes my paper by offering suggestions for future research and policy implications of my findings.

II. Literature Review

Fertility theory and its limitations

The literature on fertility and family planning is immense, but incomplete. While the literature has done an excellent job of explaining behavior in countries where family planning programs have been implemented successfully and resulted in large reductions in family size (See Bongaarts 1994, Mauldin and Ross 1991, and Bongaarts et al. 1990), it has been drastically less successful in explaining or even approaching the anomalous cases where family planning programs fail, usually in low-income, highly religious countries. By undertaking a brief review of fertility theory, this section will outline why the current paradigm remains unprepared to cope with fertility behavior in highly religious countries.

The two principal schools of fertility theory are the Easterlin and Becker camps. However, as Sanderson (1976) points out, since the mid-seventies the two once opposed sides have moved closer together, both producing a model for fertility that cannot explain behavior in countries such as Senegal. The history of the dueling and then converging economic literature on fertility began in 1960, when Gary Becker proposed a radical idea: Perhaps people chose the number of children they would have in the same way they set other consumption targets, such as pizza consumption. Unlike with pizza, the supply constraint was internal; a woman could only bear so many children in her lifetime. Yet the underlying tenet of the model was that a household would want to consume less children than it could naturally supply, because each additional child took resources away from other goods. Child consumption, therefore, was expected to be a function of income.

It became clear almost immediately that such a model was drastically out of step with actual fertility behavior. Blake (1968) presented one of the first critiques, pointing out that most empirical evidence did not support Becker’s theory. In fact, as income rises, fertility tends to decrease. Poor families and families in poor areas “consume” many more children than wealthy families in developed countries. Blake suggested several reasons why the Becker model did not fit, arguing among other things that, unlike with cars, as income
rises so does the opportunity cost of consuming children, as they take time away from work, especially for the mother. Several modifications to Becker’s initial model have brought it closer to being able to explain fertility behavior in the developing world, but countries in which there are no clear family size aspirations still present a challenge to models based around explicit demand for children.

Richard Easterlin (1966) made the first refinement by suggesting that fertility behavior was dependent not only on household income, but on the ratio of one generation’s income to that of the previous generation. In other words, a wealthy couple might not consume more children than a poor couple because in comparison to the previous generation, the wealthy couple was not relatively wealthy. This led to the development of the notion in both camps that parents have aspirations not just for the quantity of their children, but for their quality as well (i.e., whether the child can attend school). Becker and Lewis (1973) proposed that parents make tradeoffs between child quantity and child quality, and that wealthy parents choose quality in lieu of quantity. For his part, Easterlin (1976) proposed that parents wished to give their children as much wealth as their parents had given them. Through varying avenues, both authors concluded that the observed income-fertility relationship should not be expected to be positive.

Other authors suggested that people may have children in developing countries for reasons other than consumption, including production value (such as farm labor, as in Espenshade 1977) and investment value (the provision of financial security in old age, as in Nugent 1985). Overall, the literature points to at best a belief on the part of poor families that they will be increasing their consumption opportunities now or at some future point by having an additional child. In reality, according to the generalized literature, a child does not usually start being of positive net value until well into his or her teens, and overall additional children decrease parental consumption possibilities (Espenshade 1977).

A second strand of fertility literature departs from trying to model why households make given consumption decisions regarding family size and begins to develop a model for what factors influence this decision, and then what factors affect the household’s ability to meet this target. This is commonly referred to as the “supply-demand” model of fertility, because it is concerned both with natural supply of children and how many are actually desired by a household. While this new model, developed first by Easterlin in 1978, provides some valuable tools for analyzing fertility, it still leaves unquestioned the primary assumption that fertility is a function of explicit demand and natural supply. Whether people make specific choices over the size of their family is left largely unexamined by the literature. Nonetheless, a review of Easterlin’s model is still instructive for this analysis in several ways.

Easterlin’s model is useful because it allows for the possibility of dividing fertility into two distinct categories: fertility decisions when natural fertility is below desired fertility, and decisions when fertility is above desired fertility. The Easterlin model breaks down fertility into factors determining demand for children and factors determining supply. As development progresses, family size targets are expected to decrease, creating a gap between natural supply and the now-lower demand, which contraception is used to fill. On
the other end of the spectrum are women whose natural fertility is actually lower than their family-size target, resulting in a supply constraint. Under the supply-constrained model, women do not seek to limit family size in any way, and therefore do not use birth control. This new model helps to explain away the fertility behavior in countries where children are tied to social status, have significant production value, and may have investment value—families in these areas seek more children than they can naturally supply, and therefore do not engage in family planning.

The graph below illustrates this model visually, with development progressing along the x-axis. Countries where families are supply-constrained would be operating in section I of the framework below, with modern countries operating in sections III and IV.

**Figure 2.1: Easterlin Supply-Demand Framework**

In Easterlin’s framework, demand for contraception should equal the economic cost of the natural oversupply of children, or the distance between the graph of natural fertility and desired fertility. Stage II demonstrates the lag between when desired fertility crosses below natural fertility and when fertility-reduction measures such as modern contraception are adopted. The key illustration of the above graph is that once the gap between desired and natural fertility becomes large (costly) enough, natural fertility will be reduced. The idea is that anyone can access contraception, because if the cost of an additional child is high enough it would be worth the cost of engaging in abstinence, for example. Therefore, the Easterlin model opened up the possibility for the first time that people don’t fail to use contraception because it is too expensive or not readily available enough, but rather they deem not to use it because the net benefit of an additional child is still positive, or at least not negative enough to justify the initial cost of family planning.

Several researchers have used this framework to explain the slow adoption of modern contraception, or any form of family planning, in certain sub-Saharan African countries, notably Kelley et al. (1982) in their study of rural Egypt. Kelley et al.’s treatment is particularly instructive because they lay out three frameworks under which observed fertility could remain quite high (too high according to Western standards), even as a country develops along other metrics.

Taken directly from their work, they are:

1) Irrational behavior: family size is determined outside a framework of rational choice; the number of children is “up to God,” and the number of surviving children is, thus, the by-product of sexual activity and mortality.
2) Rational behavior with overproduction: family size is determined by rational choice. Parents weigh the benefits and costs (broadly viewed) of children and attempt to attain a family size goal. However, most families exceed that goal due to lack of knowledge or the high cost or the improper use of contraception. Large families are, therefore, explained by the presence of unplanned children.

3) Rational behavior: family size is determined by rational choice, and, while there may be some overproduction of children, large families are explained to a great extent by the relatively high benefits and low costs of children.

Kelley et al. find the third framework to be the most convincing explanation of Egyptian behavior, although all three elements likely contribute, since families are not homogeneous.

What is most instructive about this work is that Kelley uses a different assumption about the fertility framework of Egyptian women than I do for Senegal, yet models fertility using many of the same variables of interest, including age at first marriage and education. Observed behavior under the supply constrained model (the third framework, above) and the “irrational” behavior model (first framework) is expected to be almost identical—in both frameworks fertility becomes a function of factors like health, duration of marriage, and frequency of intercourse. Therefore, my modeling techniques will be very similar to Kelley’s and others (including Bongaarts 1987, Kelley 1988, and Cochrane 1979) who have attempted to model fertility under a different framework. The principal difference between the frameworks, then, and where I will make my main contribution to the literature, is in the interpretation of the results. Women who are supply constrained, but still willing to plan family size, could possibly be persuaded to develop lower targets for family size, and, with sufficient family planning program efforts, be persuaded to reduce fertility. But under the “irrationality” framework, family size must be impacted through indirect avenues, because very few women are willing to exert control over their family sizes. Therefore, the proximate determinants of fertility, to borrow a phrase from Bongaarts (1987), such as age at first marriage, become critical in determining family size.

The next section will review the evidence that demonstrates many Senegalese families to be operating primarily under the “irrationality” framework, which I will henceforth refer to as the innumeracy framework.

**Leaving it “up to God”: Lack of agency over family size**

Evidence from the past thirty years show that while Senegal has invested effort into family planning programs, modern contraception has not been adopted by the broad populace. Nonetheless, Senegal’s fertility has declined over the same time period, indicating that alternative avenues have the potential to impact family size. The evidence shows that Senegal is not operating primarily under either the rational choice or supply-constrained framework, but rather in an environment where innumeracy over family size leads to high fertility by default.

Cultural research in Senegal reveals a society where childbearing is an ultimate value, not only a sign of social status, but a physical tribute to God. Any tampering with natural fertility, therefore, is seen by many as a rejection of gifts from God. Because of this, the
theoretical literature is largely unprepared to explain the determinants of fertility in Senegal or how it might be reduced. The common thread in the theoretical literature is that, as development progresses, women develop a desire to limit family size, and then seek out a method of carrying out this desire. But to even express a desire to limit family size in Senegal is seen as going against God’s will. Without a recognition of agency over fertility, there is no fertility decision to target or reduce. Therefore, policy based on the assumption that households want to have fewer children, or have a specific number of children in mind that can be manipulated somehow, will be ineffective.

A review of the position of women in Senegalese society illustrates how many may come to believe that family size is outside their control. According to research by Kane (1972) and Boye, Hill, Isaacs, and Gordis (1991), women in Senegal experience a high degree of religious constraint and spousal pressure. As a Muslim wife, the woman is expected to both bear children (a symbol of power and wealth) and care for them. The father must provide resources for the family, but it is the wife’s responsibility to account for her children’s health. Thus, if the children are malnourished, the wife is held accountable. Polygamy is legal, and some sources suggest a husband might threaten a wife with a new marriage if the first wife does not bear a satisfactory number of children or care for them properly. The Senegalese marriage code has measures to protect women, but also solidifies the husband’s control over women in many ways. One article allows the husband to oppose his wife’s pursuing a profession. Marriages continue to be established by the male suitor offering a dowry to the bride’s family, which precludes women having full agency over their choice of partner, since their families stand to benefit. Women can legally marry at age 14, though many marry earlier despite laws. Maternal deaths are common due to the young ages of mothers at first births. There is no law preventing the use of contraception, but it is widely disapproved of.

Such an atmosphere makes Senegal ill-fitted for the types of family planning programs that have had large effects elsewhere. Mauldin and Ross (1991) found that in developing countries as a whole, the availability of contraceptives could account for 72% of the variance in fertility decline from 1975 to 1990. The major weakness of their study, however, is that it includes countries such as China (where family planning is widely accepted) alongside those like Senegal (where it has been all but ignored). Although the 1997 Demographic and Health Survey showed that 86% of married women in Senegal knew of some method of contraception and 79% knew of a modern method—reflecting strong programs to spread knowledge of contraception—only 13% reported they were using a method. Careful, country-based analyses reveal that, because sentiments surrounding fertility are so influenced by cultural factors, it is critical to treat countries individually when modeling effects or searching for policies. Lumping countries with drastically different cultural frameworks together results in overly simplified and ill-fitting models.

The limited use of contraception in Senegal despite widespread awareness undermines the long-held tenets of fertility researchers that contraception use is the key (and perhaps the only) determinant of fertility decline. Robey, Rutstein, and Morris epitomize this view in their 1993 article, which holds that “differences in contraception prevalence explain
about 90% of the variation in family planning rates.” Even researchers such as Pritchett (1994), who have admitted that family planning programs may have limited impact in some countries, subscribe to the narrow over-supply or under-supply model. Pritchett’s analysis crystallizes the problem, separating population researchers into the camp that believes a lack of contraception is the main determinant of high fertility and those who attribute it to high family-size targets (Pritchett places himself in the latter group). An examination of cultural values in Senegal, however, reveals that neither framework is appropriate. Many Senegalese women simply do not wish to choose a specific family size.

LeGrand, Koppenhave, Mondain, and Randall (2003) discovered a deep inability and unwillingness to quantify ideal family size in Senegal when they studied whether the “insurance effect” against infant mortality held in Senegal and Zimbabwe. If it held, people would be expected to have more children when infant mortality rates are high, to ensure a minimum family size. If infant mortality fell, family size should, too. LeGrand et al. found that families in Senegal were, for the most part, not making fertility decisions based on a (narrowly defined) rational weighing of costs and benefits, in contrast to families in Zimbabwe.

For the insurance effect to act, households must both have specific ambitions for family size and understand and internalize how changes in infant mortality affect the ability to reach that desired family size. LeGrand et al. found the second part of this requirement to hold generally in Senegal, but not the first. The researchers concluded that the high degree of belief that God, not humans, should determine total fertility would render the insurance effect, or any conscious fertility choice, imperceptible in Senegal.

LeGrand et al. highlight that in the 1997 Demographic and Health Surveys (the dataset used for this analysis), 20% of Senegalese women gave non-numeric answers when asked for their ideal family size. These answers would be statements such as, “It’s up to God,” or “I have no control over that.” While 20% is not a majority, it reflects a deep-seated unwillingness to quantify family size, even when asked by an authority figure. Still more women may have responded with very large numbers that were not necessarily sincere, such as 10 or 15 or, in one case, 23, as a way of satisfying the interviewer while not violating religious principles. Zimbabwe, a non-Muslim country, provides contrast, with only 2% of respondents offering non-numeric replies. Urban Senegalese displayed more willingness to control family size, stating that they desired to leave it up to fate, but recognized that scarce resources required careful planning. Rural Senegalese, however, bristled at the suggestion that births were something to be controlled and planned for. LeGrand et al. reported their findings as follows:

In contrast, in the Senegalese village there was little reproductive agency in terms of the deliberate control over the number of children born to a woman. There was a consensus that God alone should influence family size, with one woman saying that limiting fertility is akin to stealing lives of children God wants you to bear. The only acceptable justification for a married woman to stop childbearing was when an additional pregnancy would jeopardize her life (Randall 2001). To the extent that some women may have sought to limit their fertility, they attempted to do so through longer birth spacing, exag-
gerated claims of health problems, or terminal abstinence. Villagers were aware of modern contraception and knew that it could be obtained from a nearby dispensary at a moderate price; yet no one admitted to using it or to knowing someone who did.

The same kind of non-numeric replies and unwillingness to quantify family size has been found in other papers, including van de Walle’s 1992 study of Mali and Kelley et al.’s Egypt study. However, neither paper fully deals with how family size may be controlled without women expressing agency over it. This question was outside the scope of the LeGrand paper and has not been taken up anywhere else, revealing a large deficiency in the literature.

The Le Grand study provided one additional insight into how people thought about childbearing in Senegal. Polygamy seemed to play a major role in shaping responses, with men often thinking in terms of number of wives rather than number of children. Therefore, it seems reasonable to think of a single wife’s fertility as more independent of her husband’s fertility aspirations than in some other scenarios. Men who desire more children could take on additional wives, reducing the cost to women of bearing a smaller number of children. Since it is largely the investment of the mother’s time that affects child welfare, a reduction in children per woman, even if overall household children remained constant, could still provide gains.

Because of unwillingness to quantify family size, and thus adopt family planning, those gains that have been made in fertility levels cannot be attributed to increased contraception use. However, the 1997 Report from the Demographic and Health Surveys shows that fertility per woman fell from 6.7 children in 1985 to 5.7 children in the 1997 data, a sharp decline. Cohen (1998) found that, while other researchers had dismissed this change as anomalous because of the lack of accompanying contraceptive use, there were reasonable explanations for the decline in fertility.

The decrease in fertility could be almost entirely attributed to increases in age at first marriage. In data from 1992-93, the proportion of women age 20-24 who were married by age 20 was 59.7%, whereas 82.5% of women aged 45-49 had been married by age 20. By comparing women who have recently crossed the 20-year-old threshold with women who would have been in the same age group twenty-five years ago, Cohen essentially compared the percentage of women 20-24 who were married before age 20 in 1967 with the percentage who were married before age 20 in the 1992-3 data. While slightly less reliable (since the comparison relies on 25-year-old memories) than if a true comparison were available, Cohen’s findings are powerful and show a significant trend toward later marriages. The median age at first marriage, however, remained a very young 16.2 (18.2 urban, 15.7 rural). Cohen also hypothesized that this trend had been obscured for a time because most research examined rural and urban Senegal together. While urban age at first marriage had been increasing, and fertility declining, for some time, both variables in rural areas lagged behind.

Cohen also established that age at first marriage increases with education, with those with no education having a median age of 15.8 at first marriage, those with primary education an age of 19.3, and those with secondary education an age of 23. If it holds that
these women had correspondingly later first births, this may provide evidence of education decreasing fertility through later first marriages.

This finding supports the hypothesis that factors impacting natural fertility may be the more effective route to spur family-size reduction in Senegal. While Cohen demonstrated the effect of age at first marriage in a broad, macro framework, my study will examine the effects of this and other natural fertility determinants on a micro level, showing that for Senegalese women the most important determinants of actual fertility are those that affect natural fertility.

One thing is clear: the unique way people think about children in Senegal requires new ideas about fertility control separate from the family planning mindset of Western culture. This broad re-imagining of fertility control will be particularly salient for rural Senegal, where women are most likely to marry young, be illiterate, reject notions of family planning, become engaged in polygamous unions, and suffer from high levels of infant mortality and malnutrition. It is in these areas that women need family planning most, and here that it has been the hardest to implement because conclusions from greater sub-Saharan Africa are incompatible with the Senegalese experience.

In conclusion, fertility theory has evolved from its early beginnings of treating children purely as a consumption good to a more realistic model that better explains high fertility in low-income societies. However, the literature still remains unprepared to cope with societies in which family size is not consciously determined, as LeGrand et al. have shown to be the case in Senegal. In these societies, it is critical to look to the determinants of natural fertility, such as age at first marriage, because these will largely be the factors that determine actual fertility.

III. Modeling Fertility

Micro and Macro Determinants

Although the evidence shows Senegalese households to be operating primarily under the innumeracy framework, the inputs we expect to determine actual fertility are largely unchanged. For this reason, earlier work done under different frameworks will inform my choice of modeling techniques.

Perhaps the most notable work in modeling fertility determinants is the 1978 Bongaarts model. Although his examination is of macro fertility, it is still informative. Bongaarts asserted that fertility is first a function of socioeconomic, cultural, and environmental variables, but that these factors do not act directly on fertility. Rather, they act through fertility’s direct determinants, which on a macro level are the percentage married in a country, contraception use, induced abortion, infertility during breastfeeding, frequency of intercourse, sterility, miscarriages, and the duration of the fertile period. Bongaarts’ model contributes some interesting ideas to the modeling of micro fertility. While we expect factors like education and income to affect fertility, Bongaarts explains that they do not act directly on it. Rather, they either change tastes or change natural supply. However, these variables may still be useful for inclusion in a theoretical model, as many of Bongaarts’ direct determinants, such as very early miscarriages, are unobservable. This
idea of direct (or proximate, as he calls them) and indirect determinants of fertility largely informs this analysis. The fact that social variables do not directly impact fertility is critical to keep in mind when interpreting coefficients. For example, education’s effect should be thought of in terms of its impact on fertility’s direct determinants. An increase in education might, for example, move a woman out of the innumerate framework, encourage her to target a lower family size, and spur the use of contraception (the true direct determinant), or it may actually increase fertility by increasing the woman’s health, and therefore decreasing early miscarriages (again, the direct determinant of children born). Therefore, it is important to remember that the more direct determinants of fertility a model includes, the less significant the education effect should be.

As far as what variables should actually be included in a micro model of fertility, both economic literature and socioeconomic literature are informative. Because Kelley et al.’s work on Egypt is most closely related to my own study, it provides a good jumping off point for creating a theoretical model. Kelley estimates the effects of age, age squared, age at first marriage, wife’s education, husband’s education, personal assets, real assets, electricity, and child deaths on total children ever born. Kelley’s regression is, in effect, a pared down version of the sociological model presented by Cochrane (1979). This paper will fully integrate the sociological model into an economic framework, adding other variables that may have an important effect on family size. Cochrane’s sociological model expects age at marriage, age, lactation, health, sexual activity, sexual taboos, marital status, living arrangements, current number of children, contraceptive use, wife’s occupation, husband’s occupation, and husband-wife communication to all affect children ever born, as shown below:

**Figure 3.1: Socioeconomic Model of Fertility**

Cochrane’s model is, in a way, a fancy version of the Easterlin model, as it, too, explicitly deals with factors that go into supply of children (labeled fecundity) and those that go into demand (represented here as fertility control). The supply and demand factors come together to create births. In this analysis, I am most concerned with the factors that influ-
ence fecundity, although some factors that influence family size desires will also be included to account for women who are functioning in the conscious-choice framework.

In the above diagram, Cochrane acknowledges the endogeneity of some of these factors with arrows showing how one factor may impact several others, all of which in turn impact family size. While this endogeneity is not accounted for in my model, future research can test the robustness of my results by using some of the estimation strategies suggested by Kelley and Schultz (1988), including finding instrumental variables for those factors thought to be endogenous.

My model is ultimately a more complicated version of Kelley’s regression and a simplified version of Cochrane’s socioeconomic model. My principal variables of interest are those that can affect family size without a conscious fertility target on the part of the woman, namely age at first marriage, cultural taboos, living arrangements, and breastfeeding. The effects of each of these variables go back to Bongaarts’ model of the proximate determinants of fertility. Each variable must be interpreted in terms of how it acts on the direct determinants of fertility. For my variables of interest, these effects are: Age at first marriage may decrease fertility by decreasing exposure and therefore shortening the fertile and married time period, cultural taboos against sex while breastfeeding may artificially extend the infertile period following childbirth, alternative living arrangements may reduce frequency of intercourse, and breastfeeding may again extend the infertile period after childbirth. Education is of interest insofar as it affects each of these variables, as an even farther removed determinant of fertility. It may also have effects unrelated to these variables, however, as it can act on women who already express agency over family size by reducing family size targets, and may additionally, at a certain threshold level, remove women from the innumeracy framework and encourage the creation of family size targets.

Because education encompasses so many things, the next section will be devoted to specifically breaking down what its expected effects are and through which avenues it acts.

**Thinking beyond family planning: the uncertain history of education**

The existing literature on family planning has looked to female education as a key variable because it may change demand for children as well as encouraging the modern attitudes that lead to the adoption of family planning methods. Yet by only thinking of female education in terms of its effects on demand for children and contraceptive use, the existing literature glosses over a huge portion of its potential effects. Education doesn’t just change the mindset of a woman when it comes to making fertility choices; it changes the entire framework in which she is making those choices. Education occurring early in life affects the entire timeline of a woman’s existence. The literature in this area has failed to highlight the effects of education on factors that influence total family size outside of conscious fertility decisions. What is needed is a clear sorting of the effects of education into those that act on fertility decisions and require agency (e.g. education increasing opportunity cost, hence decreasing demand for children, hence increasing contraception use, hence decreasing total family size) and those that act on fertility outcomes through indirect pathways (e.g. continued education delaying first marriage, leading to less total childbearing time, leading to lower completed family size). This sort of separation is essential
because the latter avenues are those more likely to have an effect in highly religious countries such as Senegal.

However, the effect of education on fertility is not unambiguously negative. In areas that are supply-constrained—when people would like to have more children than their natural fertility allows—education may actually increase fertility because it increases health and thereby the ability to have children. It may also subtly alter other variables in a way that produces an unpredictable net effect on fertility. As T.W. Schultz (1974) put it:

The education of parents, notably that of the mother, appears to be an omnibus. It affects the choice of mates in marriage. It may affect the parent’s preferences for children. It assuredly affects the earnings of women who enter the labor force. It evidently affects the productivity of mothers in the work they perform in the household, including the rearing of their children. It probably affects the incidence of child mortality, and it undoubtedly affects the ability of parents to control the number of births. The task of specifying and identifying each of these attributes of the parents’ education in the family context is beset with analytical difficulties.

The most extensive work on education’s impacts on fertility, Susan Cochrane’s 1979 book, Fertility and Education, provides useful background for this study, although she fails to acknowledge that education can act through both demand-based and incidental avenues on fertility.

Cochrane first establishes that the relationship between fertility and education is convoluted. The negative relationship, she finds, is strongest for women in the urban sector, which may be why people are so quick to assume education acts primarily through increased contraception use. However, this result may be because of factors that simultaneously increase fertility in the rural sector when education rises, such as increased health. Where the relationship is negative, Cochrane points to multiple paths, both direct and indirect, through which education acts to reduce fertility. In addition to arguing that education both increases contraception knowledge and willingness to use contraception, Cochrane also points to increases in age at first marriage, decreases in the perceived benefits to children, increased sensitivity to cost of children, increases in awareness of alternative sources of satisfaction, and improved husband-wife communication.

The important caveat of her analysis is that in some areas education’s effects that increase fertility may be stronger than those that decrease it. She notes that in some low income areas there may be an initial increase in fertility from small amounts of education, perhaps because it increases the perception of being able to afford children and perhaps because it decreases child mortality, but principally due to decreases in length of breastfeeding, decreases in adherence to sexual taboos, and increases in natural fertility from better health. On balance, she expects female education to decrease fertility, with some possible exceptions.

Kelley et al. (1982) take a more skeptical view of education. Their analysis is useful, however, because it is one of the few sources to sort the effects of education into direct and indirect effects. Kelley et al. assert that low levels of education can be expected to increase age at first marriage and therefore decrease family size indirectly, while high levels may
decrease family size directly by decreasing demand for children. The Kelley book is the first to suggest that education may have different effects for different women, depending on whether the schooling is sufficient to create conscious family size goals.

However, they anticipate that other effects such as a reduction in child deaths and fewer incidences of prolonged breastfeeding will increase fertility. Additionally, Kelley et al. suspect that the effect of female education increasing fertility because of health effects is likely to be the strongest, especially for rural areas where the baseline education is low. They postulate that this effect will more than likely outweigh the small reductions.

In their empirical study, Kelley et al. find the effect of education to be positive on family size when age at first marriage is controlled for. However, because education is found to have a large effect on age at first marriage, they suggest that the overall effect of education in rural Egypt may be to decrease family size. They also suggest that education may have a more significant negative impact on family size if there were more roles for women in the workforce, which would increase women’s opportunity cost. Without these roles, some of the expected effect of education may be lost.

While the literature on education and fertility has identified all the variables that will be needed in this analysis and explained their likely effects, nowhere has there been a clear description of how education works through both demand-based avenues (increasing opportunity cost and hence decreasing demand for children, for example) and indirect avenues (such as increasing age at first marriage). Without breaking education down into these separate mechanisms, it is difficult to analyze its effects. For different women, even within the same country, fertility will be determined in different frameworks. To truly understand the effect a policy like education will have on a country, one must understand how it affects each of these frameworks. So while in Senegal, where many women do not express conscious agency over family size, education is expected to act primarily through the unconscious effects, this framework will coexist with a narrowly designed rational framework. I hypothesize that there may be some threshold level of education that increases women’s agency over family size, moving them into the conscious decision-making framework. Education, then, would both act on women in the innumerate and rational framework and act to move women from one framework to the other.

Cochrane (1979), Cohen (1998), and Kelley (1982) all expect education to increase age at first marriage, which in turn decreases fertility. This yields the arrow to the right of the variable, indicating that as education increases, children ever born should decrease through this pathway. Cochrane and Kelley expect breastfeeding to decrease with education, and it is widely accepted as a traditional means of fertility reduction, so through this pathway children ever born should increase as education increases. Cochrane expects education to decrease adherence to cultural taboos, which again are expected to reduce fertility, so through this pathway children ever born increases as education increases. There is no literature on whether education should increase or decrease living with extended families, so it is unclear how education will affect fertility through this channel. Education is widely expected to increase health (in Cochrane 1979 and Kelley 1982, among others), which increases fertility resulting in more children ever born, so this effect yields a posi-
tive education-fertility relationship. Therefore, the effect of education in the innumerate framework remains ambiguous, as proposed by previous literature. However, age at first marriage is expected to have accounted for significant fertility declines in Senegal historically (Cohen 1998) and so may outweigh the other effects. My analysis empirically tests each of these effects to determine whether the net effect is expected to be positive or negative on fertility.

After passing the agency threshold, women are willing to aspire to specific family size targets, thereby altering the effects of education. Here, factors effecting demand are expected to have more of an effect than factors affecting supply, because as shown in Easterlin’s 1978 graphical representation (Figure 2.1), actual fertility will eventually converge with desired fertility. The marginal cost of fertility control is expected to be much smaller than the initial cost (Easterlin 1978), so once the threshold has been crossed it is not expected to be a high cost proposition to bring realized fertility down to desired fertility. The main effect of interest here, then, is that education is expected to exert a negative influence through all the factors influencing actual fertility. The effect of education on fertility in the rational choice framework is expected to be unambiguously negative, because women form and aspire to specific family size goals.

This analysis updates Cochrane and Kelley’s works by using a similar framework under the assumption of innumeracy. By carefully sorting the effects of education into avenues that create and lower family size targets and those which affect the natural supply of children, we can move toward a better understanding of how education affects fertility overall. Under this framework, the demand-side effects of education can only appear in the cases of women who have specific family targets or who are spurred by their education level to formulate them. We can therefore expect education to act on all four of our indirect variables of interest (age at first marriage, breastfeeding, cultural taboos, and alternative living arrangements) as well as by removing women from the innumerate framework and causing them to target specific family sizes. It will also act on women who are already expressing agency by encouraging lower family size targets, and helping them better meet these targets with contraception use. Because education is expected to have both increasing and decreasing effects on the factors influencing fertility, I pay special attention to which factors carry more weight.

IV. Data

My analysis uses data from the Demographic and Health Surveys (DHS) to examine the determinants of Senegalese fertility. The DHS data is widely established as the best data for this type of analysis (notably, Cochrane 1979). In particular, the Senegal data have been used by Cohen (1998) and Garenne and Joseph (2002). However, few researchers have used the Senegal data for an analysis specific to Senegal. Instead, the Senegal data have been used in concert with the broader sub-Saharan Africa data in order to prove macro theories about greater sub-Saharan Africa. Because part of my thesis is that Senegal differs from other sub-Saharan nations in important cultural factors, I reverse this trend in my paper, and focus solely on the Senegal data.
The Demographic and Health Surveys are a global survey initiative sponsored by the U.S. Agency for International Development. The surveys include a wide range of family and health-related questions. Although the DHS are a worldwide initiative, questions are tailored specifically to each country included, and many questions are country-specific. The questions are then translated into the native language of the country, and interviews are conducted in person at the subject’s residence. Women are the units of analysis, and the survey is administered to a representative sample of women age 15 to 49 in the country.

Volunteers are trained by the DHS to pose the questions, and they may return several times in order to complete the survey. However, despite this vigilance, the Senegal survey still contains many missing values. Values are most frequently missing for sensitive or uncomfortable topics or when the answer would rely on the subject’s memory. Because the survey covers the entire childbearing history of the subject, answers that recall greater recollection are more likely to be missing. Additionally, for some questions information is simply unavailable. For example, the weight of a child at birth might be unavailable because no health professional was on hand at the birth to weigh the child, or the level of household income might be unavailable because many people are subsistence farmers or part of the barter economy. For this analysis, variables with large numbers of missing values were not used. This required some creativity in specifying certain variables, but the dataset’s many positive features more than made up for the drawbacks.

For my analysis, I use the DHS performed in Senegal in 1997. The dataset includes information in the following areas: background characteristics, reproductive behavior and intentions, contraception exposure and use, prenatal and postpartum care of all children in household, breastfeeding and nutrition of children, children’s health, the relative status of the woman in the household, the husband’s background, and AIDS and other sexually transmitted infections. My analysis draws variables from all sections of the survey except the AIDS section. The dataset contains 8,593 observations, which comprise women from both urban and rural areas. Some of these observations are dropped because of missing values, but every effort is made to preserve the size of the dataset.

My principal dependent variable is “Total Children Ever Born,” which I will regress on a number of demographic inputs, suggested by the theoretical literature as important factors in determining family size. However, I have also added some features to the theoretical model that set my analysis apart. My analysis features a specific variable for rural living, since my research shows the rural sector to be different from the urban sector in significant ways. Other researchers have tried to account for the rural effect with income and education proxies or by focusing only on rural areas, as in Kelley et al. (1982). I am adding a specific variable for a non-numeric response to a question asking ideal family size to see if a lack of numeracy over child choice directly impacts total family size.

My analysis explicitly treats women as the unit of analysis, instead of looking at household fertility. This is important because a traditional “household” model assumes a husband and wife, and in Senegal living arrangements frequently do not match this paradigm: many women lived with their extended families, were part of polygamous unions, or had
absent husbands. Making the women the units of analysis eliminates any confusion over trying to estimate household factors such as family size or income or total children, and restricts the analysis to factors relating to the specific woman.

From the theory and my own research, it seems the following factors should largely determine a given woman’s total family size: age, age at first marriage, household income, education, occupation, desired family size, numeracy over family size, sexual taboos, living arrangement, length of breastfeeding, child deaths, contraceptive use, urban/rural sector, frequency of intercourse, and health. Because of data availability, some of these variables have been specified in alternate ways.

Total Children Ever Born: This is my main dependent variable. It refers to the total number of live births, as reported by the woman. Because it does not account for still births, it may not capture a full measure of fertility. However, still births and miscarriages are a supply constraint, and therefore treated in the literature as no different from being unable to conceive. Some researchers have noted that total children may be underreported because women could be reluctant to report live births where the child died soon after childbirth. This omission could bias the variable slightly, but those who do underreport children born are unlikely to be different in significant ways from those that do not. While this paper is focusing on large family sizes, the mean for total children ever born is only 3.19. This may be artificially low because 29.05% of respondents had never had any children. This is largely accounted for by the 25.15% of women in the sample who have never been married. Mean children born for married women is 4.61. For women 40 to 45 who are expected to have completed their childbearing, mean children born is 7.63, giving us a better idea of how many children a Senegalese woman generally bears in her lifetime.

Age: Age is available in the dataset, measured in years, and so is used as a simple continuous variable. There are no missing values for this variable, and responses are expected to be fairly accurate. To account for the non-linear effects of age, an age-squared variable is also included.

Rural/Urban: 64.4% of respondents were from rural areas while 35.6% were from urban areas. While I expect the effect of this variable to be strong as-is, it is important to note that the rural sector may be so different from the urban sector that it needs to be accounted for in other ways. For example 87.83% of the rural population cannot read at all, compared to only 50.83% of the urban population. These sort of huge disparities mean the effects of the two sectors are unlikely to be captured fully by a simple change in intercept. Additionally, Cochrane (1979) found education to impact rural and urban areas differently. For this reason, an interaction term of rural times level of education will be added.

Marital Status: It is obviously expected for unmarried women to have fewer children than married ones, especially because the country is so religious. Women can be married, divorced, widowed, separated, or never married. 74.85% of the sample has been married at least once. Marital status is first accounted for with a dummy variable for “never married,” and later the regression is rerun for only women who have been married, since 87% of never married women have no children.
Education: Education is measured using a survey question for highest level of education attained. I use three dummy variables for “primary,” “secondary,” and “higher,” with no education being the baseline. Overall, 19.84% of women had primary education, 9.25% had secondary, and .85% had higher. The rest had no education. This is interesting because it may be difficult to sort out the effects of education since the sample in each category will be so small. Additionally, even if education is found to have a large effect, it may be difficult to implement because it is currently so uncommon.

Child Deaths: This variable is a measure of total children who were born live and have since died. It includes children who have died at any stage in their lives, so could theoretically include a 25-year-old dying if the mother is within the sampled age range. I initially broke down this measure into sons who have died and daughters who have died, since theoretically families may be more eager to replace sons in pursuit of additional labor, but the effects were found to be near identical. The final model, therefore, treats all child deaths the same.

Cultural Taboos: The only relevant cultural taboo that would affect childbearing in Senegal is a taboo against sex while breastfeeding because it is thought to spoil the mother’s milk. This is a zero/one dummy referred to in tables as “No sex.” This cultural taboo may provide an unconscious mechanism of birth spacing. 21.46% of women acknowledged they avoid sex while breastfeeding.

Age at First Marriage (AFM): This variable is self-reported by women who have been married. It is calculated in years. Average age at first marriage is around 17.

Ideal family size: This variable represents the woman’s childbearing ambitions. Responses ranged from 0 to 23. Additionally, about 20% of respondents gave a non-numeric answer such as “That’s up to God.” The effect of ideal family size will be captured in the continuous “family size” variable, with non-numeric response being captured in a zero/one dummy. The mean for people who did give an ideal number was slightly more than five. It is also important to note that some numbers such 10 or 15, which were much more common than, say, nine and 14, display a lower level of numeracy in regards to family size, and may therefore have a stronger positive effect on total children than just a high ideal number. In rural areas 24.11% of respondents gave non-numeric responses, in urban areas this was only 14.14%, again displaying the differences between the two sectors.

Labor Force participation: The data does contain a question about the woman’s occupation, but it is sorted into many different categories that would be impractical to include. For this reason, I condensed the variable into a “Formal Sector” zero/one dummy. All jobs that involve significant time outside the house were considered formal sector employment, with “no job,” “subsistence farming,” and “sales” considered to be non formal-sector employment. Subsistence farming was included because this response is indistinguishable from the “not working” response, since both would be applicable to a woman who serves as housekeeper, childrearer, and occasional farmhand. Sales was included because this may include occasional street vending, and may therefore not entail significant time outside the house.
Contraceptive Use: This is captured in a variable for whether the respondent has ever used any method of contraception, including modern, traditional, and folkloric types. I included all types of contraception because to engage in any type of family planning, even less effective types like folkloric methods, is to cross a barrier from passive family size determination to active reduction. 10.5% of respondents were currently using some contraceptive method. 3.64% were currently using a modern method. There were still stark differences between urban and rural, but neither sector showed high patterns of use.

It is important to note that the rate of contraceptive use does not include extended periods of breastfeeding, which has widely been accepted as a traditional method of birth control. Some women might be using breastfeeding to delay pregnancy (control birth spacing), but might not be comfortable expressing it as a means of contraception. In this data, the use of “traditional methods” is so low (3.40%) because it only includes those methods consciously undertaken as “birth control,” such as withdrawal or periodic abstinence. Because there is a cultural taboo against controlling family size, traditional birth control may be more prevalent than expressed here, but viewed as a method of birth spacing or undertaken for other benefits apart from limiting family size.

Income: This variable cannot be captured directly because so many families do not have conventional incomes to speak of, but rather trade in goods and services. The literature suggests using proxies such as electricity, material possessions, and distance from water source, but none of these were found to be significant. Furthermore, several measures of material possessions included at once were not jointly significant. I therefore constructed a proxy measure of income by dividing number of rooms for sleeping by total household size, creating a ratio of rooms to people. This should proxy for income because a wealthy family would have a bedroom for each person, while very poor families often live as large extended families in single-room structures.

Breastfeeding: Length of breastfeeding is measured by using the length of breastfeeding in months of the second to most recent child. This is because many of the most recent children were still breastfeeding at the time of the survey, which lowers the variable’s predictive power. However, there is only data on breastfeeding for women who had a child in the past three to five years, meaning almost 75% of the sample was excluded, leaving 2,284 observations. This information also relies on the woman’s memory, resulting in a large number of focal point answers such as 18 or 24 months.

Some Missing Factors: I do not include a variable for relative persuasive power, since this is difficult to quantify, and also because my analysis hopes to remove the assumption that family size is a negotiation where both parties have a desired outcome. Health was left out of this analysis because the Senegal survey contains no data on the health of the woman, including measures of height and weight. While this variable may have considerable predictive power, omitting it should not drastically affect my results, as health in terms of ability to bear children, or fecundity, will be picked up by other variables such as income and education. I also do not include a variable for frequency of intercourse, because the data on this was found to be unreliable. This effect, as well as the effect of health, will be accounted for later in the analysis by the addition of a “birth interval” vari-
able that normalizes fertility by the number of days between marriage and first birth. Summary statistics for my variables are listed below.

### Table 4.1: Summary Statistics for Included Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Expected Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total children ever born</td>
<td>8593</td>
<td>3.194228</td>
<td>3.155993</td>
<td>0</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>8593</td>
<td>28.18434</td>
<td>9.329213</td>
<td>15</td>
<td>49</td>
<td>+</td>
</tr>
<tr>
<td>Age-squared</td>
<td>8593</td>
<td>881.3809</td>
<td>567.6174</td>
<td>225</td>
<td>2401</td>
<td>-</td>
</tr>
<tr>
<td>Rural</td>
<td>8593</td>
<td>0.644478</td>
<td>0.478699</td>
<td>0</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>Never married</td>
<td>8593</td>
<td>0.251484</td>
<td>0.433891</td>
<td>0</td>
<td>1</td>
<td>- strongly</td>
</tr>
<tr>
<td>Education Dummies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>8593</td>
<td>0.198417</td>
<td>0.398831</td>
<td>0</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>Secondary</td>
<td>8593</td>
<td>0.092517</td>
<td>0.289772</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Higher</td>
<td>8593</td>
<td>0.008495</td>
<td>0.091783</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Not wife</td>
<td>8593</td>
<td>0.63319</td>
<td>0.481962</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Child deaths</td>
<td>8593</td>
<td>0.583149</td>
<td>1.0864</td>
<td>0</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>No sex</td>
<td>8593</td>
<td>0.214593</td>
<td>0.410564</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>AFMXmarried</td>
<td>6432</td>
<td>17.13169</td>
<td>3.497683</td>
<td>8</td>
<td>42</td>
<td>-</td>
</tr>
<tr>
<td>Family size</td>
<td>6814</td>
<td>5.414147</td>
<td>2.357809</td>
<td>0</td>
<td>23</td>
<td>+</td>
</tr>
<tr>
<td>Nonnumeric</td>
<td>8593</td>
<td>0.205633</td>
<td>0.404187</td>
<td>0</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>Absent husband</td>
<td>6021</td>
<td>0.282511</td>
<td>0.450258</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Formal sector</td>
<td>8507</td>
<td>0.120372</td>
<td>0.325415</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Contraceptive</td>
<td>8593</td>
<td>0.104969</td>
<td>0.306532</td>
<td>0</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>Rooms to ppl</td>
<td>8176</td>
<td>0.371951</td>
<td>0.1383</td>
<td>0.071429</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Birth interval</td>
<td>5312</td>
<td>27.95105</td>
<td>30.40223</td>
<td>0</td>
<td>378</td>
<td>-</td>
</tr>
<tr>
<td>Breastfeeding</td>
<td>2284</td>
<td>16.6979</td>
<td>7.048159</td>
<td>0</td>
<td>36</td>
<td>-</td>
</tr>
</tbody>
</table>

While there are small concerns with certain variables, most of the theoretical variables of interest have readily available counterparts in the DHS dataset. Of these included variables, I am most interested in how education, age at first marriage, cultural taboos, alternative living arrangements, and breastfeeding affect children ever born.

### V. Findings

The main goal of my empirical analysis is to show that age at first marriage, length of breastfeeding, cultural taboos, and living arrangements all have large effects on total children ever born. Education is also expected to affect children ever born, both through these pathways and by impacting the fertility decisions of those women who do make conscious choices and by removing some women from the innumerate framework. Through these avenues, family size and infant mortality for low-income women can be reduced without necessarily requiring conscious agency over family size on the part of the woman. This section will be divided into two parts: the first breaks fertility down into its various deter-
ominants and highlights factors that have large effects, and the second shows how education may impact each of these important factors.

**Modeling Fertility**

I regressed “total children ever born” on a number of factors that have been shown to be important determinants in previous literature, with the goal of identifying factors that act without conscious agency over family size. The structure of my model, ordinary least squares, allows me to look for things that do reduce family size, rather than factors that make people want to reduce family size or make people able to meet their desired family size. My model removes the assumption of conscious choice.

I expect the coefficients on age at first marriage, no sex while breastfeeding (a cultural taboo), not wife of household head, breastfeeding, and education to be large and negative. The following table presents results from my first regression.

To confirm that this regression yields reasonable results, let’s first examine the coefficients of some basic control variables. The first two variables account for the major determinants of children born, marital status and age. Both yield statistically significant coefficients. We expect the effect of never being married on total children ever born to be strongly negative, which it is. Women who have never been married have almost three fewer children than women who have. We expect the effect of age to be strongly positive, which it is. The coefficient is not as large as that as marital status because the range of age is so large, from 15 to 49. The coefficient can be loosely interpreted as saying that for each

| Total children ever born | Coef.  | Std. Err. | t     | P>|t|  | [95% Conf. Interval] |
|--------------------------|--------|-----------|-------|------|----------------------|
| Never Married            | -2.89963 | 0.1182    | -24.53| 0    | -3.13134 to -2.66793 |
| Current age              | 0.234148 | 0.014006  | 16.72 | 0    | 0.206693 to 0.261602 |
| Age-squared              | -0.00084 | 0.000221  | -3.81 | 0    | -0.00127 to -0.00041 |
| Rural                    | 0.191281 | 0.041821  | 4.57  | 0    | 0.1093 to 0.273261   |
| Education Dummies        |         |           |       |      |                      |
| Primary                  | 0.150596 | 0.048221  | 3.12  | 0.002| 0.056071 to 0.245121 |
| Secondary                | -0.12344 | 0.067399  | -1.83 | 0.067| -0.25556 to 0.008676 |
| Higher                   | -0.7855  | 0.188218  | -4.17 | 0    | -1.15446 to -0.41654 |
| Not wife                 | -0.5915  | 0.041598  | -14.22| 0    | -0.67304 to -0.50995 |
| Child Deaths             | 0.841647 | 0.018726  | 44.94 | 0    | 0.804938 to 0.878355 |
| No sex                   | -0.14088 | 0.041303  | -3.41 | 0.001| -0.22184 to -0.05991 |
| AFMXmarried              | -0.12608 | 0.005955  | -21.17| 0    | -0.13775 to -0.11441 |
| Family size              | 0.042528 | 0.00847   | 5.02  | 0    | 0.025924 to 0.05913 |
| Nonnumeric               | 0.171027 | 0.0643    | 2.66  | 0.008| 0.044983 to 0.297071 |
| Formal sector            | -0.1536  | 0.054097  | -2.84 | 0.005| -0.25964 to -0.04756 |
| Contraceptive            | 0.654028 | 0.057383  | 11.4  | 0    | 0.541541 to 0.766514 |
| Rooms to ppl             | -2.78676 | 0.122142  | -22.82| 0    | -3.02619 to -2.54733 |
| Constant                 | 0.230378 | 0.245953  | 0.94  | 0.349| -0.25175 to 0.71251  |
additional year of age a woman will have .23 of an additional child, or about 1 child every five years. The Age-squared term then reduces this, meaning that this effect decreases as age continues to increase.

The next variable is one of our principal variables of interest, education. The baseline here is no education, so in accordance with the theory we expect primary education to have a positive effect, while higher education will have a negative effect. Cochrane (1979) attributes this initial positive effect to education first increasing the ability to have children, through such avenues as better nutrition and better healthcare. She also asserts that small amounts of education may cause women to abandon traditional patterns of breastfeeding and postpartum abstinence. If this is responsible for the increase in fertility, including information on the health benefits of breastfeeding in any planned female education policy might eliminate this unintended effect.

My results for education mirror Cochrane’s 1979 results. The coefficient on primary education is positive, and statistically significant. The coefficient for secondary education is negative, as expected, and statistically significant. Women with secondary education tended to have about .12 fewer child births than women without. The coefficient for higher education is extremely large, saying these women tended to have almost one full less child than women without any education.

The effect of abstaining from intercourse during periods of breastfeeding has a statistically significant effect on total children ever born. Literature suggests that women do not abstain from sex while breastfeeding to consciously space their children farther apart, but rather to avoid “spoiling” their milk, according to cultural tradition. While the effect of this taboo is not overwhelming, a reduction of .14 children for women who abstained during breastfeeding, it is still another avenue through which family size can be reduced without agency. This finding matches a 1973 report by the United Nations, cited in Cochrane (1979), which found that cultural taboos on sex should act negatively on total fertility. The next section will explore how education affects this factor.

The effect of age at first marriage is also statistically significantly negative, although it is not overwhelmingly large. For each additional year before first marriage, women tended to have .13 less children. This finding is in line with the Senegal-specific findings of Cohen (1998), and Cochrane’s 1979 survey of the work of previous scholars, including McGreevy and Birdsall (1968), Kim et al. (1974), Davidson (1973), Yaukey (1972), and Palmore and Ariffin (1969), all of whom found the effect of wife’s age at marriage on fertility to be negative. However, the size of this effect means it would require a change in age at first marriage of eight years to get a one-child reduction in family size. Compared to the standard deviation of age at first marriage of 3.5 years, this hardly seems feasible. However, it is impossible to know how the size of this effect breaks down across individuals—for some families the effect from just a single year increase in first marriage might have a significant impact on family size. And, as in Cohen’s findings, the effect is expected to have a significant impact on macro fertility levels.

Recall “rooms to people” was constructed as a proxy for income. It is the ratio of the number of rooms for sleeping listed in the survey to the number of household members.
While there is little precedent in the literature for this construction, none of the literature-supported income proxies, such as whether a household has electricity, were found to be statistically significant. One reason for the slight effects of such variables could be because income is thought to act negatively on fertility through channels such as the increased opportunity cost of women. Because other variables such as education and occupation may account for this effect, the income effect may be weakened. However, the “rooms to people” ratio gives us a large, statistically significant, negative effect of about three fewer children for each unit increase in the ratio. However, when compared to the standard deviation of the ratio, .1383, to get this effect would require a change of eight standard deviations. More work should be done to see if this coefficient represents the income effect or other factors.

One interesting feature of this regression is the large negative effect of living in a household headed by someone other than the respondent’s husband. Women who were not the wife of the household head had .6 fewer children than those women who were. This is consistent with earlier findings that separate location of spouse (which may be implied by the household head variable), Williams (1976), and joint family living, Williams (1976) and United Nations (1973) both have negative effects on fertility.

The effect of a job in the formal sector is negative, as expected, but not very large. It may also not be as robust as some other factors, as specifying this variable as “working” in general versus no employment does not yield statistically significant results, perhaps because this includes occupations like subsistence agriculture and street vending/panhandling, in which the woman’s opportunity cost may not be large. Additionally, this variable loses its statistical significance once unmarried women are excluded.

Child deaths have a positive effect on fertility, with a statistically significant coefficient of .84. Schultz (1997) points out that the causality for this variable may run both ways, explaining the large effect. Families may have more children in response to child deaths, perhaps acting to replace the child. Families may also have more children in anticipation of child deaths, the so-called insurance effect tested by Le Grand et. al (2003). There is also the possibility, however, that the causality runs backward, with rates of infant mortality being higher in large households because resources are more likely to be spread thin.

Contraceptive use is another such mystery. Intuitively, the use of contraception should reduce family size, by helping to bring family size down to desired levels. However, this effect may be confounded by the fact that use of contraceptives is seen as extreme in Senegalese society. Only families severely overburdened with children may consider the use of contraception. This might explain the positive coefficient of contraceptive use in the above results. However, ideally the model would account for all factors influencing natural fertility, and therefore produce a negative coefficient on contraceptive use. The fact that contraceptive use has a positive coefficient because we believe it to be correlated with fertility means that some factors affecting fertility are still unaccounted for. These are the “unobservable” factors in Bongaart’s 1978 model. Adding a variable for the interval from marriage to first birth should normalize the results by overall fertility, and help
solve this problem (see table 5.3).

This analysis supports my predictions about how three of my variables of interest—
age at first marriage, cultural taboos, and living with extended families—affect children
ever born. My fourth variable of interest, breastfeeding, will need to be analyzed separately.
The effect of education matches that predicted in the literature, with primary education
increasing fertility and secondary and higher education decreasing it, but more analysis
needs to be done to sort out how this variable influences other factors. The second stage
of my findings section addresses this very question, but first I engage in some sensitivity
analysis to test how my variables are affected by a few refinements of the model.

Testing Differences Between Married and Unmarried Women

Because Kelley (1982) chose to look only at married women, it is worth examining
whether my findings are robust under this specification. Results for these separate groups
are shown below.

<table>
<thead>
<tr>
<th>Table 5.2: Breakdown by Marital Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total children ever born</td>
</tr>
<tr>
<td>R-squared</td>
</tr>
<tr>
<td>Never Married</td>
</tr>
<tr>
<td>Current age</td>
</tr>
<tr>
<td>Age-squared</td>
</tr>
<tr>
<td>Rural</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education Dummies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
</tr>
<tr>
<td>Secondary</td>
</tr>
<tr>
<td>Higher</td>
</tr>
<tr>
<td>Not wife</td>
</tr>
<tr>
<td>Child deaths</td>
</tr>
<tr>
<td>No sex</td>
</tr>
<tr>
<td>AFMXmarried</td>
</tr>
<tr>
<td>Family size</td>
</tr>
<tr>
<td>Nonnumeric</td>
</tr>
<tr>
<td>Formal sector</td>
</tr>
<tr>
<td>Contraceptive</td>
</tr>
<tr>
<td>Rooms to ppl</td>
</tr>
<tr>
<td>_constant</td>
</tr>
</tbody>
</table>

***significant at 1% level, **significant at 5% level, *significant at 10% level

Breaking down my results by married and unmarried women, we can see that all find-
ings except the impact of formal sector employment hold up for married women.
However, there is a troubling lack of good explanatory variables for unmarried women.
This is probably because women who have children out of wedlock are exceptions to the
rule in Senegal, as doing so goes against religious practices. The determinants of having
children out of wedlock are unlikely to be the same that explain higher fertility after mar-
riage. For this reason, my next refinement includes only married women in the sample.
Controlling for Birth Interval

The next refinement to the model is normalizing some of the omitted fertility effects by adding a “birth interval” term defined as the interval between marriage and first birth. Logically, this regression was run for only married women.

The principal difference in this regression is that the effect of primary education is no longer statistically significant. This is expected, however, as primary education is hypothesized to increase fertility by increasing natural ability to supply children, or fecundity. Recall that primary education is expected to have positive health effects, due to a better understanding of basic health practices and nutrition. If this is the principal avenue through which primary education increases fertility, controlling for this fecundity effect in birth interval should yield an insignificant coefficient on primary education.

Primary education may also reduce adherence to cultural taboos that delay childbirth as well as decrease length of breastfeeding, but these factors are not accounted for in birth interval since they act after the first birth. Because accounting for birth interval alone is enough to make the coefficient on primary education statistically insignificant, we can infer that these effects of education alone are not strong enough alone to create a statistically significant positive effect.

Therefore, these results suggest that primary education’s main impact on fertility occurs through the positive change of increasing health. This lends credence to the idea that while education may initially increase fertility, it does so through effects that are generally considered positive, such as increased health. Since this increased knowledge of health and nutrition can be expected to transfer to the mother’s care of her children, this
might still create a more supportable family size even if it is actually larger in number, because the woman may be more able to care for her children.

Contraceptive use continues to have a positive correlation with total children ever born, however, indicating that some fertility factors may still be unaccounted for.

**Rural/Urban Interactions with Education**

To test whether education has different effects in rural versus urban areas, I now add an interaction term for rural combined with education. Recall that Cohen (1998) found that the urban versus rural populations had undergone a demographic transition at completely different time periods. Cochrane’s 1979 work also found differing effects for rural versus urban areas, and she suspected the effect of education might be stronger in urban areas. Additionally, Kelley (1982) looked at only rural women, so this makes the results more comparable. Because here we are interested in the fecundity effects of education, such as how it impacts health in the different sectors, this regression omits birth interval.

**Table 5.4: Results with Rural Interactions—R-squared=.7035**

| Total Children Ever Born | Coef.  | Std. Err. | t     | P>|t| | [95% Conf. Interval] |
|--------------------------|--------|-----------|-------|-----|---------------------|
| Age                      | 0.421634 | 0.01866 | 22.6  | 0.385054 | 0.458213 |
| Age-squared              | -0.00355 | 0.000287 | -12.37 | 0.00411 | -0.00299 |
| Education interactions   |        |           |       |       |                     |
| Ruralx1                  | 0.253179 | 0.115913 | 2.18  | 0.029 | 0.025947 | 0.480411 |
| Ruralx2                  | 0.460764 | 0.221792 | 2.08  | 0.038 | 0.025972 | 0.895557 |
| Ruralx3                  | 0.378338 | 1.211515 | 0.31  | 0.755 | -1.99667 | 2.753343 |
| Education Dummies        |        |           |       |       |                     |
| Primary                  | 0.038344 | 0.085097 | 0.45  | 0.652 | -0.12848 | 0.205165 |
| Secondary                | -0.43235 | 0.113643 | -3.8  | 0.06513 | -0.20957 |
| Higher                   | -0.73346 | 0.306341 | -2.39 | 0.017 | -1.33399 | -0.13292 |
| Not wife                 | -0.47415 | 0.046427 | -10.21 | 0 | -0.56516 | -0.38314 |
| Child deaths             | 0.815391 | 0.020745 | 39.31 | 0 | 0.774724 | 0.856059 |
| No sex                   | -0.22871 | 0.054425 | -4.2  | 0 | -0.33541 | -0.12202 |
| AFMxmarried              | -0.14001 | 0.006769 | -20.68 | 0 | -0.15328 | -0.12674 |
| Family size              | 0.060975 | 0.010466 | 5.83  | 0 | 0.040458 | 0.081491 |
| Nonnumeric               | 0.27233 | 0.08086 | 3.37  | 0.001 | 0.113814 | 0.430845 |
| Formal sector            | -0.12311 | 0.078524 | -1.57 | 0.117 | -0.27704 | 0.030827 |
| Contraceptive            | 0.671447 | 0.068658 | 9.78  | 0 | 0.536853 | 0.80604 |
| Rooms to ppl             | -3.31237 | 0.156567 | -21.16 | 0 | -3.6193 | -3.00544 |
| Constant                 | -2.32802 | 0.314747 | -7.4  | 0 | -2.94504 | -1.711 |

This new model is interesting because the coefficients for both primary and secondary education interacted with the rural term are positive, while the general education effect is now insignificant. This means the effect of primary education in urban areas may not be positive at all, while its effect in rural areas is positive, significant, and robust. One possible explanation for this is that health in urban areas might already be at a higher level, making the effect on fecundity of better health insignificant. The lower incidence of
breastfeeding in urban areas may also play a role, as primary education is also expected to increase fertility through shorter periods of breastfeeding.

Perhaps more surprisingly, the positive interaction term on secondary education is enough to outweigh the overall negative effect of secondary education, indicating that even secondary education may increase fertility in rural areas. This may be because health is so low initially in rural areas that it is a large enough effect to outweigh any of the fertility-reducing effects of education. It may also be because breastfeeding is an important method of fertility control in rural areas, and is lessened with education. It may also be that women in rural areas are more “stuck” in the innumerate framework than in urban areas, and therefore the effects of education will barely include effects from creating or lowering family size targets and adopting contraceptive use. This would be consistent with LeGrand’s 2003 finding that people in urban areas were much more willing to express agency over family size and were much more similar to the Zimbabweans surveyed than were rural Senegalese. For this reason, policies aimed at reducing rural fertility in Senegal may need to focus entirely on the innumerate side of the framework, since even at quite high (for Senegalese society) levels of education, rural women do not exhibit characteristics consistent with having crossed the “agency threshold.”

Overall, these findings are consistent with Cochrane’s results that education can be expected to have the largest impact in urban areas. The small and insignificant general primary education term implies that the initial boost in fertility from primary education occurs mostly in rural areas. It also implies that the fecundity effects of education—or other undiscovered effects increasing fertility—are quite large in rural areas. The fact that Kelley et al. (1982) limited their examination to rural areas might explain why their general conclusions about education are more pessimistic than my own.

Breastfeeding

The next question I try to answer is whether length of breastfeeding after childbirth decreases overall fertility. Because data on this variable only includes women who had a child in the last 3-5 years, the sample size is limited. For this reason, this variable was not included in the main regression.

From these results we can see that breastfeeding an additional month is expected to decrease fertility by .01 children. Because the maximum length of breastfeeding in the sample was 36 months, and the standard deviation 7 months, this is not a very economically significant finding. It would take 14 standard deviations, or a change of 100 months in length of breastfeeding, to reduce children born by one child. Alternate specifications of this regression omitting certain variables yielded either insignificant results or a positive coefficient, showing that this finding is not robust. Better data availability would help to explore this effect, but, for now, the effect of breastfeeding is unclear. However, I will defer to the literature and accept the hypothesis that breastfeeding decreases children ever born.

The regression also eliminates the statistical significance of all levels of education. While this is primarily due to sample size, it has an intuitive payoff: the more variables we add that are proximate determinants of fertility the less the indirect determinants should
matter. Theoretically, if we could account for every proximate determinant the coefficient of education would be zero.

Thus, the previous sections have yielded expected coefficients on age at first marriage, cultural taboos, and living with extended family, and an expected but not very robust coefficient on breastfeeding. Increased age at first marriage, increased adherence to cultural taboos, increased instances of living with extended families, and increased length of breastfeeding are all expected to decrease total children ever born. The next task will be to see how education impacts each of these effects, as well as effects outside the illiterate framework, to see if it is worth pursuing as a policy for fertility reduction.

| Total children ever born | Coef.   | Std. Err. | t      | P>|t|  | [95% Conf. Interval] |
|-------------------------|---------|-----------|--------|------|----------------------|
| Current age             | 0.397822| 0.025414  | 15.65  | 0    | 0.34798              |
| Age-squared             | -0.00091| 0.0004    | -2.28  | 0.023| -0.0017              |
| Rural                   | 0.040968| 0.056749  | 0.72   | 0.47 | -0.07033             |
| Education Dummies       |         |           |        |      |                      |
| Primary                 | 0.088004| 0.066537  | 1.32   | 0.186| -0.04249             |
| Secondary               | 0.035715| 0.131858  | 0.27   | 0.787| -0.22289             |
| Higher                  | -0.49692| 0.409738  | -1.21  | 0.225| -1.30051             |
| Not wife                | -0.12546| 0.043767  | -2.87  | 0.004| -0.21129             |
| Child deaths            | 0.328601| 0.021178  | 15.52  | 0    | 0.287065             |
| No sex                  | -0.04861| 0.05444   | -0.89  | 0.372| -0.15538             |
| AFMXmarried             | -0.32569| 0.008208  | -39.68 | 0    | -0.34178             |
| Family size             | 0.010719| 0.010456  | 1.03   | 0.305| -0.00979             |
| Nonnumeric              | 0.061812| 0.081805  | 0.76   | 0.45 | -0.09862             |
| Formal sector           | 0.149608| 0.084509  | 1.77   | 0.077| -0.01613             |
| Contraceptive           | 0.236448| 0.064573  | 3.66   | 0    | 0.109807             |
| Rooms to ppl            | -1.07032| 0.179081  | -5.98  | 0    | -1.42154             |
| Breastfeeding           | -0.01123| 0.003037  | -3.7   | 0    | -0.01719             |
| Birth interval          | -0.02837| 0.000845  | -33.59 | 0    | -0.03003             |
| Constant                | 0.364554| 0.409967  | 0.89   | 0.374| -0.43948             |

Thus far, the results for education have been expected but inconclusive. Primary education was found to increase fertility, while secondary and higher education reduced it. When birth interval was accounted for, the effect of primary education became insignificant, indicating that this effect may be primarily due to increased fecundity from better health. Splitting the effect of education into its general and rural-specific components yielded positive effects on fertility for both primary and secondary education, indicating that fertility-increasing effects must be strong enough in the rural sector to outweigh many of the fertility-reducing effects of education. However, to see whether education is, on bal-

Table 5.5: Results with breastfeeding—R-squared=.8863
ance, likely to produce a positive or negative effect on fertility, we must examine how it affects the four variables of interest that were controlled for, and thus not entered into the education term, in earlier regressions: namely age at first marriage, adherence to cultural taboos, living with extended families, and length of breastfeeding. We should additionally explore how breastfeeding affects women outside of the innumeracy framework by reducing family size goals and increasing contraceptive use, and also how it may remove people from the innumeracy framework demonstrated by a change in nonnumeric answers to the ideal family size question. The next section addresses these issues in detail.

**The Effects of Education**

This section explores how education affects each of the variables of interest. This will help us examine whether education overall is expected to have a negative or positive effect on fertility. I find that education is expected to increase age at first marriage, decrease length of breastfeeding, have no statistically significant effect on adherence to cultural taboos, increase the chance of living with extended family, and increase contraception use as well as increase numeracy over family size and reduce family size goals. Therefore, education is expected to decrease fertility through multiple indirect avenues that do not require agency while at the same time encouraging women to express agency over family size and lower their family size targets. The only avenue through which education may increase fertility (besides health, as has already been shown) is by decreasing length of breastfeeding.

I first examine the effects of education on the main determinants of fertility in the innumerate framework, which are the determinants of natural fertility explored earlier. Each dependent variable is regressed on the basic controls of age, age-squared, rural/urban, and formal sector employment, in addition to the three levels of education.

<table>
<thead>
<tr>
<th>Dependent</th>
<th>AFM</th>
<th>Breastfeeding</th>
<th>No Sex</th>
<th>Not Wife</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td>0.1616</td>
<td>0.0246</td>
<td>0.008</td>
<td>0.2714</td>
</tr>
<tr>
<td>Education Dummies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>1.442173***</td>
<td>-0.90851**</td>
<td>0.012203</td>
<td>0.090856***</td>
</tr>
<tr>
<td>Secondary</td>
<td>3.438614***</td>
<td>-1.21485</td>
<td>-0.01798</td>
<td>0.134741***</td>
</tr>
<tr>
<td>Higher</td>
<td>5.885442***</td>
<td>-6.73271**</td>
<td>-0.05536</td>
<td>0.069154</td>
</tr>
<tr>
<td>Age</td>
<td>0.348948***</td>
<td>0.658469***</td>
<td>-0.01454***</td>
<td>-0.05576***</td>
</tr>
<tr>
<td>Age-squared</td>
<td>-0.00507***</td>
<td>-0.00987***</td>
<td>0.000191***</td>
<td>0.000535***</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.95463***</td>
<td>1.156927***</td>
<td>0.018268*</td>
<td>-0.08151***</td>
</tr>
<tr>
<td>Formal Sector</td>
<td>0.730718***</td>
<td>-1.45834**</td>
<td>-0.02465*</td>
<td>0.088524***</td>
</tr>
<tr>
<td>Constant</td>
<td>11.73927***</td>
<td>5.694764***</td>
<td>0.447846***</td>
<td>1.74423***</td>
</tr>
</tbody>
</table>

***significant at 1% level, **significant at 5% level, *significant at 10% level

These tables provide some valuable insights into the various pathways of education that were, in effect, hidden in earlier regressions. Primary education alone results in a 1.44-year increase in age at first marriage. Secondary education yields a 3.43-year increase, and higher education results in a nearly six-year increase. Because age at first marriage decreases fertility, education will also decrease fertility through this avenue. In
a sense, then, controlling for age at first marriage in earlier regressions camouflaged some of the fertility-decreasing effects of education. Indeed, rerunning earlier regressions without including age at first marriage as a control results in negative, statistically significant, coefficients for all levels of education, including primary.

As for breastfeeding, this regression shows that all levels of education are expected to decrease duration of breastfeeding, although secondary education is not statistically significant. This is consistent with the conclusions of Cochrane (1979). Because breastfeeding is expected to decrease fertility, this is one avenue through which education will likely increase fertility, and thus an area policymakers should carefully consider when implementing an education program.

Contrary to expectations, and the findings of Cochrane (1979), here education has no statistically significant effect on adherence to the cultural taboo against sex while breastfeeding. Intuitively we expect education to decrease this adherence, as education results in more modern attitudes that are less likely to incorporate folklore into decision-making. While the true coefficient on abstaining from sex while breastfeeding may be negative, we can safely conclude it is not so negative that it is a major concern for policymakers looking to decrease fertility through education. This belief may be so deeply rooted in Senegalese society that it is not affected by education, or it may be tied up in the respondent’s understanding of health, which is expected to increase with education.

Additionally, living with extended families in which the woman’s husband is not the household head is actually increased by education. Because there was no literature on this variable, we did not have an expected direction for this effect. Again, this coefficient is small so the finding may not have much impact. However, it may be reassuring to policymakers to know that this is another natural check on fertility that will at least not be decreased by broader education.

I now examine how education impacts factors that affect demand for fertility and fertility control, which are expected to function outside the innumerate framework.

With all levels of education, contraceptive use increases, indicating that more educated women are more willing to engage in family planning, although the overall predictive power of this model indicated by the R-squared is low. There is a significant but economically slight positive effect on contraceptive use for primary education, but a larger effect for secondary and higher education.

All levels of education also decrease the likelihood of offering a nonnumeric response when asked about ideal family size, offering more evidence that education can potentially remove people from the innumerate framework. Because this effect appears at all education levels, it is evident there is not a single agency threshold, but rather that the level of education needed to embrace agency over family size differs for each woman. Some women would likely continue to adhere at any education level to religious beliefs dictating family size should be left up to God.
Additionally, for those women who already respond numerically, education lowers ideal family sizes. This indicates that education may be raising the opportunity cost of women or enhancing awareness of alternative sources of satisfaction, as Cochrane (1979) suggests.

These findings lead to the conclusion that while education does not unambiguously reduce family size, it can have powerful effects both on the natural determinants of fertility and on the factors that determine fertility outside the innumerate framework. First and foremost, education is expected to decrease natural fertility by significantly increasing age at first marriage. Education also increases instances of living with extended families and has no effect on adherence to cultural taboos, leaving both of these natural fertility checks in place. At the same time, education removes some women from the innumerate framework, evidenced by lower instances of offering nonnumeric ideal family sizes at higher education levels, and acts on women already making conscious fertility decisions by lowering their family size targets and increasing their use of contraceptives. Therefore, it seems that for these factors, education affects fertility in more ways that decrease family size than increase it. I have therefore adjusted my theoretical model for education’s effects to make it consistent with my findings.

### VI. Conclusions

This study aimed to discover whether there are ways to reduce family sizes in rural Senegal within a framework where people are reluctant to admit control over childbirth. By carefully examining both conscious and unconscious ways family size can be impacted, my research improved on earlier literature that did not pay close attention to women who did not admit agency over family size. I first identified factors in the literature expected to impact family size without conscious agency, then tested how these factors affected total children ever born for a representative sample of Senegalese women. I then used a new model for how education acts on fertility in both the innumerate and conscious-agency framework to untangle the myriad impacts of education and try to determine which

---

**Table 5.7: The Effects of Education on Fertility Control and Fertility Demand**

<table>
<thead>
<tr>
<th>Dependent</th>
<th>Contraceptive</th>
<th>Nonnumeric</th>
<th>Ideal family</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td>0.0702</td>
<td>0.0366</td>
<td>0.1278</td>
</tr>
<tr>
<td>Education Dummies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>0.084978***</td>
<td>-0.11383***</td>
<td>-0.738***</td>
</tr>
<tr>
<td>Secondary</td>
<td>0.13353****</td>
<td>-0.16948***</td>
<td>-1.10509***</td>
</tr>
<tr>
<td>Higher</td>
<td>0.147538****</td>
<td>-0.17469***</td>
<td>-1.37152***</td>
</tr>
<tr>
<td>Age</td>
<td>0.030168***</td>
<td>-0.01436***</td>
<td>0.017462</td>
</tr>
<tr>
<td>Age-squared</td>
<td>-0.00042***</td>
<td>0.00025***</td>
<td>0.00014</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.05244***</td>
<td>0.039052***</td>
<td>0.942133***</td>
</tr>
<tr>
<td>Formal sector</td>
<td>0.021447**</td>
<td>-0.0108</td>
<td>-0.19807**</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.37359***</td>
<td>0.406478***</td>
<td>4.548597***</td>
</tr>
</tbody>
</table>

***significant at 1% level, **significant at 5% level, *significant at 10% level
effects were likely to be strongest.

My research drew largely on Kelley’s 1982 framework, and my results are mostly in line with his findings. However, Kelley’s conclusions about education were more pessimistic than my own. Interestingly, Kelley et al. were studying rural Egypt in 1976, which was remarkably similar to modern-day Senegal in demographic characteristics. These societies, which lag behind their regions in demographic indicators of development, require new ideas and fresh models to try to improve standard of living on their own cultural terms. It is my hope that my analysis will provide a basis for this sort of new thinking about fertility in Senegal.

The major results of my study are threefold: 1) Age at first marriage is negatively correlated with fertility. By raising age at first marriage, total children ever born can be reduced. This will also result in reductions in infant mortality, since older mothers are shown in the literature to be better able to care for infants, and smaller family sizes are less likely to have child mortality. The direct effect on children born is not overwhelmingly large, but this small gain coupled with other gains could make a significant impact. This result is robust for different alternate specifications of the model. 2) Cultural taboos against sex while breastfeeding, living arrangements in which the husband is not the household head, and increased breastfeeding all reduce total children ever born, although my findings on breastfeeding were less robust than other findings. 3) Education acts on fertility in many ways, and it is unclear whether education will increase or decrease fertility in rural areas. Primary education increases fertility, especially in rural areas, when age at first marriage is controlled for. However, this is mainly due to increased health, and thus fecundity, and decreased breastfeeding. Secondary and higher education reduce fertility, but both are rare. Because education is positively correlated with age at first marriage, however, and because increased health increases welfare overall, I am optimistic for the prospects of education increasing welfare in Senegal, even if it does not decrease family size in all cases.

Implications

The implications of these findings are that education may not be the panacea that was hoped for rural women, yet is still worth pursuing because of its multiple positive side effects. Additionally, the effect of education on increasing age at first marriage alone may be enough to make it smart policy. Although small levels of education may increase fertility, it does so only by either increasing the health of the mother or decreasing rates of breastfeeding. The latter could be accounted for by including education on the health benefits of breastfeeding in any education program. Moreover, since education also is expected to increase both the health of the mother and her children, its increases of fertility through health can hardly be seen as a negative impact. As a result, strengthening the education system for women in Senegal would most likely benefit the country, whether or not it directly impacts fertility.

My other findings provide other possibilities for policies to reduce fertility. The promotion of formal sector jobs for women is expected to decrease fertility and, Kelley (1982) suspects, strengthen the effect of education. Cultural taboos about having sex while
breastfeeding, once considered antithetical to a modern society, need to be reconsidered in the wake of evidence that they serve a practical purpose. The government may also want to encourage families to live together as extended units and share resources, since this was shown to reduce family size for each woman. This could be done easily through tax breaks for extended families.

And, perhaps most clearly, later ages at first marriage should be encouraged, first because they reduce total family size, but also because they reduce infant mortality for all children born. This could be done by strongly enforcing, or raising, the legal marriage age of 14 already in place, or through education programs. Support for later marriages is the strongest conclusion of my analysis. While the impact may be initially small, as the culture changes toward later marriages, the effects on well-being for individual families may be immense.

Limitations and Avenues for Future Research

My study was limited by the data available to me. As such, my first recommendation is that future surveys by DHS be carried out with a mind to the direction of the research the surveys hope to foster. Information on the health of the woman and her children is critical in conducting analyses of well-being in sub-Saharan Africa. Moreover, it is one of the few things that can be objectively measured by the researcher by calculating height and weight percentiles so that reliable data is more likely. Gathering information on health for mothers and children should be a foremost priority of the DHS.

My study is also limited in that it does not reflect the newest developments in the economic theory of fertility. While this was intentional, because I hoped a simpler model would offer clearer implications, future research may want to re-complicate the model to account for endogeneity of some of the right-hand variables, all the while bearing in mind that the goal of this research should not be an ideal model, but a usable answer. Researchers should look to Kelley and Schmidt (1988) for more robust estimation techniques and ideas for reliable instruments.

Future studies may also want to update the Senegal study with the latest data available, compare Senegal to other highly religious African countries, and incorporate a longitudinal element. Each of these improvements on my work may help to answer the ultimate question of how sub-Saharan African countries can be made better off, specifically by reducing family sizes to levels supportable by individual households.

While making these improvements, however, future research should not lose track of the necessity of tailoring models to specific cultural and religious factors in a given country rather than seeking increasingly complex models that elegantly explain fertility in some situations but fail to capture its mechanisms in others. That an innumerate, or passive, framework for fertility can coexist with more conventional cost-benefit decision-making, and that both must be addressed if fertility theory hopes to remain relevant to developing countries, is the single most important conclusion of my analysis.
References


EFFECTS OF MINIMUM WAGE LEGISLATION ON EMPLOYMENT AND INEQUALITY:
EVIDENCE FROM EASTERN AND CENTRAL EUROPE

Dubravka Colic
Wellesley College

1. Introduction

Among the more established labor market mechanisms for social protection is the minimum wage -- the legislation has existed in most countries in the Northern Hemisphere for the majority of the twentieth century. Although on any given day only about 5 percent of workers earn the minimum wage in the U.S., researchers estimate that up to 60 percent of workers have earned the minimum wage at some point in their lives (Card and Krueger, 1995), and the number is likely higher in countries with more generous provisions. The legislation disproportionately hits low income families and therefore attracts the attention of social activists and politicians. But it is the clear and unambiguous prediction of the simple Economics 101 labor supply and demand model that an earnings floor reduces the number of jobs that has always attracted economists to study the impact of legislated minimum wage programs.

Over the years, a significant amount of research has been conducted in the United States and Western Europe on the effects of minimum wage laws and levels on labor market and economic outcomes. Each change in the level of minimum wage -- in the United States especially -- is accompanied by an avalanche of papers analyzing the effects on variables such as employment, poverty, labor force participation rate, women’s participation in the workforce, etc. The resulting conclusions, especially about the effects on employment, are varied and highly controversial; some researchers find significant negative effects associated even with modest increases in minimum wage levels, and others detect relatively small to no employment effects, or even job growth.

Empirical analyses of changes in minimum wage levels in the Western world have proven challenging for several reasons. Mostly, it is difficult to conduct a reliable economic analysis of such minor changes in levels as the most recent increase of the federal minimum wage in the United States from $4.75 to $5.15, the nominal level which has been maintained since 1997. Even when the nominal minimum wage remains the same, the real level (around which the predictions of economic theory revolve) does change, but the rel-
atively low inflation rate in the United States causes even the real levels of minimum wage to vary only slightly. Similarly to the United States, in Western Europe, Canada, and Japan, the general idea of minimum wage and its proper level has been relatively stable for many years, and changes to minimum wages are minor and very gradual, as are differences in levels across states within the US.

Whereas minimum wage levels in Western Europe have remained fairly stable, Eastern and Central Europe have experienced large fluctuations in minimum wages both within any given country and across the region. After the fall of the Iron Curtain in 1989, many countries either introduced minimum wages in their usual form for the very first time, or decided to increase the level of the wage in order to be more on par with the developed world. Other countries almost eliminated some employment protections altogether, or cut the minimum wage levels by so much that the extremely low monthly minimum wage no longer provided any kind of protection against poverty; in some countries, the minimum wage decreased to the point that the monthly amount would not even pay for a pair of shoes. In countries of the former Soviet Union, for example, abnormally high inflation rates in the 1990s decreased the ratio of the minimum wage to average wage below 10 percent in some years (compared to past values of up to 50 percent). While it cannot be said that large changes in minimum wage levels should also inevitably have a significant impact on workers (particularly when the fraction of workers earning the minimum wage was small to begin with and increased only slightly as the minimum wage rose), it is still more likely that a wider variation in the minimum wage will eventually affect more workers than a smaller one, particularly when already high levels of minimum wage are further increased.

Eastern and Central Europe have therefore become a laboratory of sorts for policy makers and politicians, since few labor law changes are a result of carefully planned and coordinated efforts. Instead, the countries of the region have sometimes almost randomly implemented changes to their labor market policies – after regime changes, at the advice of foreign agencies, or as a response to changes in ideology. Minimum wages and other provisions were, for the most part, generous in ECE countries during socialism. Many of these countries began the 1990s with very similar labor market institutions, which then evolved and developed in markedly different ways. Such different policy evolutions could allow for an interesting analysis of how the minimum wage affects labor market outcomes, particularly employment and inequality. Such an analysis should produce reliable results not only because of the wealth of variation in the data, but also because unplanned policy changes in the region greatly diminish the likelihood of endogeneity of the legislation.

Considering how Eastern and Central Europe provide a fruitful ground for empirical analyses of the labor market, it is surprising that more research has not been conducted on minimum wages in the region. The lack of research may be due to language barriers. Aside from the more developed Central European countries and Russia, who provide literature and data in English, data in many countries is only available in the several distinct, native languages. While some country-specific event studies and narratives do exist, the empirical research has yet to take full advantage of the variations in minimum wages across time.
and across countries. The goal of this honors thesis is to take advantage of available data from this region and evaluate the impact of minimum wage increases on employment and inequality.

A basic economic model of the labor market predicts unambiguously that employment will drop in the presence of a minimum wage and continue to fall as the level of protection is increased. However, research from the United States, Canada, and Western Europe finds both positive and negative changes in employment in response to increases in the minimum wage. I first review this evidence. Then I use new information on minimum wage legislation from Eastern and Central Europe to perform an original econometric analysis designed to estimate the impact of the minimum wage on employment and inequality. This analysis provides a contribution to the literature not through innovative economic modeling, but by using a relatively simple yet effective fixed effects framework to take advantage of a newly compiled panel data set with information never before used to analyze the effect of minimum wage on employment and inequality.

I find that a 1 percent increase in the ratio between minimum and average wages decreases the employment rate by 0.01 percent. Alternatively, doubling the minimum to average wage ratio at the outset of reform in 1995, from 27.3% to 54.6%, would decrease the employment rate from 42.7% to 42.27%. This small and negative elasticity is consistent with recent prominent research from the United States, and robust to different specifications, as well as to performing the analysis on only the subset of countries which belonged to the former Soviet Union. An equal increase in the ratio between minimum and average wages does not, however, have a statistically significant effect on the Gini coefficient. These results suggest that minimum wages have a very small negative effect on the employment rate, but cannot be said to improve economic equality in turn.

The format of my thesis is as follows: Section II provides background information on recent history of the region and its minimum wage legislation, as well as descriptive statistics for minimum wage trends and labor market data; Section III provides the economic theory of minimum wage effects on employment; Section IV reviews existing literature on case study, time-series, and panel data evidence from the US and abroad; Section V proposes the econometric model used in this thesis, as well as describes the data sources and construction; Section VI presents the major findings and proposes improvements to the econometric analysis; Section VII concludes.

II. Background

Since the fall of the Iron Curtain in 1989, Eastern and Central Europe have experienced their most tumultuous period since World War II. Beginning with the dissolution of the Soviet Union and the creation of the Commonwealth of Independent States in 1991, and culminating with many countries of the region becoming full members of the European Union in 2004, the region has undergone dramatic economic, political, and social change. Even more pronounced have been the economic reforms within particular countries - especially those of Central Europe and the Baltic states - which were able to move away from the command economy model quickly, having avoided armed conflicts.
and major political obstacles. Many (if not most) countries have retained their commitment to generous social and welfare policies, while others have moved farther toward a less regulated, more capitalist society.

Minimum wages were expected to play a significant role in easing the lives of the poor during the tumult of the transition period of the early 1990s. However, the period saw clear erosion in the levels of minimum wage with respect to the average wage, although wages were, in general, falling dramatically during this time. Following the advice of the IMF and the World Bank, policymakers in most countries were more inclined to see minimum wage laws as distortions to the new market economies than as protections for low wage workers. As a result, the regional average of the ratio between minimum and average wages fell somewhat.

Of the countries in my sample, only two – Poland and Estonia – increased minimum wages faster than average wages in these early years, while all other countries experienced downward pressure in minimum wages. In Bulgaria, Romania, Ukraine, and Hungary, initial increases in the minimum were followed by a steady fall until the middle of the decade. Russia and other countries of the former Soviet Union generally had lower minimum wage levels even before the reforms, yet these countries saw the sharpest falls in social protections, including the minimum wage. High rates of inflation along with the inactivity of legislative bodies allowed the deterioration of minimum wages from a social protection to a merely symbolic number in official records.

The trend in the employment rate in individual countries and the region was similar to that of minimum wages; nearly all countries departed from an era of full employment. As Figure 2.1 below shows, the regional average of the employment rate fell steadily in the 1990s and has rebounded only slightly in the current decade. Since layoffs were mostly centered around privatization, lack of investment in the tumultuous region, and other factors which had little to do with minimum wage, the regional trend – when compared to that of minimum to average wages – tells us very little in the absence of control variables which would account for a large part of the dramatic drop in employment.

The extent of variation in minimum wage levels in countries of Eastern and Central Europe is most easily noticed by observing the time-series trends within countries. Figure 2.2 presents several representative time trends of the ratio of minimum to average wages. In Belarus and most other non-Baltic countries of the former Soviet Union, the ratio fell steadily until only recently. Other countries, such as Lithuania, experienced the same decreasing trend early in the 1990s and then a sharp reform around 1995, maintaining the ratio relatively steady afterward. Bulgaria experimented early on and raised minimum wages to as high as 64 percent of the average wage, the largest percentage of all counties in the time period covered by my data set. After 1992, there was a fall in the minimum to average wage ratio in Bulgaria until the mid 1990s, when it recovered (and stabilized) to the present value of roughly 40 percent. Romania represents a country struggling to keep minimum wage levels at approximately one third of the average wage, with high inflation (against which minimum wages have traditionally not been protected) quickly dragging the ratio down.
The extent of variation in minimum wage levels in countries of Eastern and Central Europe is most easily noticed by observing the time-series trends within countries. Figure 2.2 presents several representative time trends of the ratio of minimum to average wages. In Belarus and most other non-Baltic countries of the former Soviet Union, the ratio fell steadily until only recently. Other countries, such as Lithuania, experienced the same decreasing trend early in the 1990s and then a sharp reform around 1995, maintaining the ratio relatively steady afterward. Bulgaria experimented early on and raised minimum wages to as high as 64 percent of the average wage, the largest percentage of all counties in the time period covered by my data set. After 1992, there was a fall in the minimum to average wage ratio in Bulgaria until the mid 1990s, when it recovered (and stabilized) to the present value of roughly 40 percent. Romania represents a country struggling to keep minimum wage levels at approximately one third of the average wage, with high inflation (against which minimum wages have traditionally not been protected) quickly dragging the ratio down.

Figure 2.3 shows a similar situation in countries of the former Soviet Union. Although the trends are more similar to each other, there is still a fair amount of variation, with the Baltic states taking a more active role in minimum wage reform and in maintaining the ratio of minimum to average wages substantially higher than in Russia, Ukraine, Belarus and Moldova. The first reforms occurred in Latvia and Lithuania, with Estonia’s initial dramatic drop in minimum wage followed by a very moderate increase over the next 10 or so years. More recently, Moldova followed in the footsteps of Lithuania and increased its minimum wage in 2001 from 7 percent to approximately 38 percent of the average wage.
It is useful to compare the extent of the variation in minimum wage in Eastern and Central Europe to a similar snapshot of the ratios across states in the United States of America. To that effect, Figure 2.4 below shows the range of ratios of minimum to average wages in countries of the region in the year 1992, which extends from only 12 percent in Belarus to 64 percent in Bulgaria.
Aside from minimum wages and employment, the change in inequality (as measured by the Gini coefficient) over the past decade and a half is equally disappointing; this region which once saw extremely high levels of equality has now seen its Gini coefficient steadily rise since the mid 1990s. After an initial drop in inequality in the early part of the 1990s, inequality soared to nearly .35, one of the higher regional averages (second to those of Africa and Latin America). Even though the Gini coefficient (particularly that of earnings, which is used in my analysis) is an imperfect measure of inequality in a region where a large portion of income is earned in the informal economy and derived from remittances from foreign countries, the calculations from earnings alone indicate an alarming trend: the region has seen a substantial increase in inequality.

Even though little can be said based on trends alone about the effects of minimum wages on employment and inequality in Eastern and Central Europe, it can be said that minimum wages no longer provide any substantial form of protection for workers in most countries in the region. In Russia and other CIS countries, the value of the minimum wage as a percentage of average wage has fallen well below 10 percent, and similar (though perhaps less dramatic) developments can be observed in other countries as well. Most Central European countries have maintained the ratio between 30 percent and 50 percent (like the US and Western Europe), and only in Poland and Estonia has the ratio increased substantially and remained high. The debilitating effect of weak and disorganized labor market institutions in Eastern and Central Europe provides additional incentive for an informative panel study of minimum wage effects on employment and on income inequality.

**III. The Economics of the Minimum Wage**

One of the most commonly discussed models in introductory economics textbooks is the simple model of the labor market in the presence of the minimum wage:
Using the conventional supply and demand framework to analyze the labor market, we notice that in the absence of wage regulations, the equilibrium real wage and the number of employees hired are determined by the intersection of the labor supplied by workers and labor demanded by firms, represented on the graph above by W1 and E1, respectively. The model assumes that the minimum wage covers all sectors of the economy, that labor is homogenous, and the labor market competitive. A newly introduced minimum wage acts as a price floor for labor, forcing employers to hire E2 workers. The existence of minimum wages thus creates a surplus of workers who can no longer be employed because the legislated wage is too high. The model predicts this drop in employment in the presence of a minimum wage in all cases other than the unlikely scenario where the demand for labor is completely inelastic. The same decrease in employment would result from an increase in the level of the minimum wage (as depicted on the graph), which would further negatively influence employment and reduce it to E3.

This classical model of minimum wage is widely adopted by economists despite its simplistic nature and the fact that much of the empirical evidence does not support its predictions. Precisely because the model is so simple and powerful, and because it does predict the behavior of some firms as well as employment drops at high levels of the minimum wage, alternative models are yet to become widely used (and evaluated) in labor economics. These alternative models depart from the basic framework only slightly, yet yield strangely similar results. They incorporate into the model of the labor market such “imperfections” as incomplete information, search costs in hiring, uncovered sectors, heterogeneous labor, long term vs. short term effects, etc. Regardless of what complications are incorporated into the model, the prediction that some workers will inevitably lose jobs from higher minimum wages remains. Only some dynamic search models (such as a dynamic monopsony model) which have surfaced in recent years even begin to allow for the flexibility of employment gains in the face of minimum wage increases; such models
are beyond the scope of this project, whose main goal is to evaluate new data using the standard framework and not pioneer economic modeling of minimum wage.

As discussed, the classical framework is widely accepted by economists, but is source of much of the controversy (both political and economic) surrounding the minimum wage because economists often interpret its implications rather differently. Some economists argue that the minimum wage hurts both the economy as a whole and low-wage earners – it decreases employment in the economy (with mostly low-wage workers losing their jobs, presumably), and also decreases the share of a society’s wealth that the least fortunate receive, assuming that the employment cuts occur among workers who previously earned the minimum wage and whom employers cannot retain because the minimum wage increased. Other economists conclude that while some low-wage workers receive an increase in earnings as a result of minimum wage, other (similar) workers are forced out of the labor market due to the minimum wage, which is a questionable tradeoff. Those who support minimum wage increases are most often non-economists who attempt to downplay the predictions of the classical model and argue that employment losses are not as dramatic as theory predicts. Interestingly, a large portion of empirical analyses supports the assumption of small to no employment losses, as outlined subsequently in the literature review.

The existing level of the minimum wage and the extent of its coverage of the labor market sometimes determine the changes in the distribution of earnings which could result from increases in the minimum wage. Studies have shown that a typical (US) increase in the minimum wage generates a mere 10 to 15 percent increase in wages of affected workers, who are fewer than 10 percent of the workers in the economy (Card and Krueger, 1995), so the effect on inequality in the entire economy will inevitably be small. The direction and magnitude of the effect depend on several factors, but most notably on our assumptions regarding the minimum wage effect on employment. Although the standard model of minimum wage predicts an unambiguous negative effect on employment, the empirical evidence is certainly mixed.

Predicting the impact of changes in minimum wages on the distribution of earnings depends on what assumptions we make about how minimum wages affect employment. If we accept the prevalent empirical results of small decreases in employment in response to minimum wage increases, it follows that some of those who would have been employed under the old minimum wage lose jobs once the level is increased. The elimination of some of these low-wage jobs should effectively remove the very lowest end of the wage distribution, therefore making the distribution more equal – although policy makers probably did not have such an equalization in mind when they were hoping for an improvement in inequality. Additionally, workers who earned less than the increased minimum wage before and who have maintained employment now earn more – namely, the new minimum wage – which should also improve outcomes at the lower end of the wage distribution and help inequality.

Finally, the increase in the wages of workers directly affected by changes in the minimum wage may raise demand for substitutes for such workers, like those earning just
above the minimum (and perhaps also lead to increases in their wages), resulting in more workers with higher wages in the labor market; this possibility is less likely if a low-wage worker is replaced by a machine or his/her job is outsourced. There may also be some repercussions for wage-earners higher up in the wage distribution, although such effects would likely be small. In summary, higher minimum wages are generally expected to somewhat improve inequality, although some of that improvement may come from the decrease in employment of lowest-wage workers and their subsequent removal from the wage distribution.

IV. Literature Review

Evidence from the US - Employment Effects

The model of the labor market in the presence of minimum wage presented in the previous section predicts a negative effect on employment when the minimum wage is increased. Evidence presented in this section is far from consistent in evaluating this prediction; the data, methodology, and findings for each study are outlined in Table 1.

Most cross-sectional studies in the United States exploit the variation in minimum wages across states which arise from the states’ imposition of a minimum wage higher than the federal minimum wage. Neumark and Wascher (1992) were one of the first to combine federal and state minimum wages into a single variable, and use differences in those “combined” minimum wages across states to estimate the effect on the employment of teenagers, who are disproportionately affected by minimum wage legislation. Their estimates use the Kaitz index (a coverage-weighted ratio of minimum and average wages) and are essentially zero if school enrollment rates are not taken into account. When enrollment is controlled for, Neumark and Wascher show that a 10 percent increase in the minimum wage decreases employment of 16-24 year olds by 0.7 percent. The result is slightly higher when both the minimum wage and its lagged value are included in the regression.

The conclusions of the above study were challenged in Card et al. (1994) in a paper that generated a series of responses. The main issues that arise have to do with the specification of the model and the construction of the main independent variable. Card et al. argued that the minimum wage variable should simply be the higher of the federal and state minimum wage, and not the coverage-adjusted measure that Neumark andWascher (1992) construct. Neumark and Wascher also do not account for changes in the state average wage, which may lead to more negative employment effects. Card et al. (1994) find that this methodology produces a positive minimum wage coefficient, although it is not significant once enrollment is included in control variables. In one of their responses to Card et al., Neumark and Wascher (1994) report estimates that do include the state minimum wage; while the coefficients are still negative, they are no longer significant. Finally, the results are highly sensitive to enrollment figures, and it can be argued that the relationship between minimum wages and enrollment might also exist – something Neumark and Wascher (1992) did not carefully consider.

Similar to previous studies, but taking the recommendation of Card et al. to pick the higher of state and federal minimum wage levels, Burkhauser et al. (1997) also used cross-
EFFECTS OF MINIMUM WAGE LEGISLATION ON EMPLOYMENT

state differences to estimate an employment effect. The study took advantage of the 1989, 1990, and 1991 increases in federal minimum wage levels by calculating the fraction of teenagers in each state that would be affected by the change in minimum wage and regressing this fraction against the mean log(wage) of teenagers and their employment to population ratio. The study finds that higher minimum wages significantly reduce teenage employment, with coefficients ranging from 4% to 9% in response to a 10% increase in the minimum wage. The results are highly sensitive to the sample used, and are particularly critiqued both by Burkhauser and Card-Katz-Krueger (in a later paper) for not distinguishing between January of one year and the next (i.e. including year dummies). The authors find that augmenting the model to account for time variation and not only cross sectional variation is the culprit for the large discrepancy between this and other studies.

Some of the more convincing and interesting studies use low-wage sectors, such as fast food and retail, to conduct their analyses. Workers who earn minimum wages are disproportionately found in the low-pay sector, so estimates are usually higher in either direction when data from this industry is used. Contrary to the large negative effect of Burkhauser et al., Card (1992) finds that employment of California teenagers in retail grew faster than in other states after the 1988 increase in state minimum wage. The study finds that wages rose 5 percent faster in California, and employment either rose or fell by one percentage point, depending on the choice of base year and comparison group, further demonstrating the sensitivity of employment effects to methods and data used.

Using a similar approach, Deere et al. (1995) obtain drastically different results for the change in the teenage employment/population ration following the 1990 and 1991 changes in the federal minimum wage. The results show that had the minimum wage level not been increased, the employment of teenagers would be 7 percent higher for males and 11 percent for females, which are some of the highest negative results found recently. Compared to the results in Card (1992), these elasticities are consistently large and negative even after augmentations to the model.

Kim and Taylor (1995) use variation in minimum wages across states to reevaluate Card’s conclusions from the study of California’s retail industry. They study the trends in wages and employment in sub-industries of retail and find no relationship between wage and employment growth before the minimum wage increase in 1988, but find a strong negative elasticity (-0.9 to -0.7) following the increase. The authors argue that the negative effects of the minimum wage were difficult to detect because of high demand in the years following the minimum wage increase, but that there indeed exists a strong negative relationship. Kim and Taylor (1995) repeat their analysis using inter-county variation within California and find the same strong negative effect on employment, thus rejecting the null hypothesis of no impact.

The fast food industry has also been subject to several studies due to the extremely low wages of its workers. In anticipation of the April 1991 rise in the federal minimum wage, Katz and Krueger (1992) surveyed fast food restaurants in Texas before and after the law change took effect. The paper takes a look at the difference between high- and low-wage restaurants in the state of Texas to find employment responses to the minimum wage hike.
The restaurants that were forced to comply with the new federal standard had faster employment growth than restaurants that were paying above the minimum already, which again stands against the predictions of the classical supply and demand mechanism.

Card and Krueger (1994) take advantage of the “natural experiment” from the April 1992 increase of the New Jersey state minimum wage level by comparing fast food restaurants in New Jersey to similar ones in an adjacent county in Pennsylvania and to higher-paying restaurants in New Jersey itself. Like Katz and Krueger (1992), and regardless of the comparison group used, they find that the new minimum wage increased employment in restaurants that were forced to comply with the new law. Welch (1995) suggests that the employment increases in response to higher minimum wage levels found in the NJ study came at the expense of employment in non-chain restaurants. Card and Krueger replied in a later paper that this argument is unlikely to be true because the high-wage NJ group that was used for comparison would be affected by the closing of non-chain restaurants just as much as the treatment group.

In an attempt to challenge the results of Card and Krueger (1994), Neumark and Wascher (1998) collected a similar data set for New Jersey and Pennsylvania fast-food restaurants, mostly because they suspected measurement error in the 1994 data set. The authors go to great lengths to justify the differences between the two data sets, and use primarily data from payroll reports to conduct their analysis, and not survey data. They present different results from Card and Krueger, including smaller standard deviations and a small but negative coefficient, contrary to the positive effect Card and Krueger detected. Both Neumark and Wascher as well as Card and Krueger attempted to refine the methods and the data, with surprisingly similar results. In a 2000 paper, in another reply to Neumark and Wascher, Card and Krueger use a new data set provided by the Bureau of Labor Statistics and perform an updated analysis of the same counties and years, finding exactly the same results as before.

Even though the general trend has been toward cross-sectional and panel data studies, time series studies still provide important evidence in the minimum wage debate. Time series data on the impact of Federal minimum wage changes has often been a controversial source of evidence for employment effects of minimum wage because federal legislation affects both high-wage and low-wage states. The vast majority of time series evidence (unlike cross-sectional or panel) finds negative employment effects. In Myth and Measurement: The New Economics of the Minimum Wage, Card and Krueger summarize the time series evidence available; the employment loss due to a 10 percent increase in the minimum wage was between 0.5 percent and 3 percent in all studies, with more recent evidence citing the smallest estimates. The authors also update some of the data sets to include the 1980’s; this decreases most estimates showing a negative employment effect and even turn the results around to indicate minor positive effects of minimum wage increases on employment. Even the effect on teenagers and young adults turn out ambiguous using updated data – increasingly, it is difficult to reject the hypothesis that the true effect of minimum wages on employment is zero.
**Effects of Minimum Wage Legislation on Employment**

**Effects on Inequality**

Estimates of the effect of minimum wages on the wage distribution and inequality are far more scarce. DiNardo et al. (1996), for example, study the period between 1979 and 1988, when inequality increased for both men and women. They found that as the ratio of the minimum wage to average wage fell by 0.27 for men, the standard deviation of log(wage) increased by 0.072, which implies that the fall in minimum wage can account for 0.018 of the increase in inequality. The impact is higher for women. Overall, the study finds that minimum wages play a significant role in determining inequality.

Card and Krueger (1995a) pick up where DiNardo left off and consider the changes in inequality between 1989 and 1992 by looking at the fraction of a state’s workers who are affected by minimum wage increases in 1990 and 1991. Their study confirms DiNardo’s results, noting that higher minimum wages significantly reduce inequality as measured by the number of workers in each earnings decile.

**International Evidence**

Most of the evidence from countries other than the United States comes from Canada, the UK, and Puerto Rico. Castillo-Freeman and Freeman’s (1992) detailed analysis of the impact of minimum wages on employment in Puerto Rico included both an aggregate time-series analysis and an industry level analysis. Their time-series analysis uses the Kaitz index, which measures the impact of minimum wage by dividing it by average wage in the industry and multiplying by the fraction of workers who are covered by the minimum wage in the industry. The analysis shows a strong negative effect of minimum wage increases and was presented as decisive evidence of strong negative effects of minimum wage changes when the existing level is particularly high, as it was in Puerto Rico. Card and Krueger challenge these results in Myth and Measurement. Given the nature of Puerto Rico’s minimum wage setting, the direction of causality might go the other way; in other words, minimum wages are often increased in those industries where employment is expected to grow. Card and Krueger included a minimum wage index and two lags of said index; using this approach, they find that the interpretation of the impact of minimum wage in the original research no longer holds up.

Baker and Benjamin (1999) analyze a panel data set of minimum wage and employment data in Canada encompassing the period 1975-1993 and initially find that a 10 percent increase in minimum wage leads to a 2.5 percent reduction in employment. However, this result is attributable mostly to low variation in the data; at high frequencies, the employment effect is positive and insignificant. This has implications both for the interpretation of empirical research on minimum wage and on the design of experiments. In a more recent Canadian study, Yuen (2003) addresses the usual complaint that most panel data studies are designed as comparisons between low-wage workers (those affected by minimum wage increases) and high-wage workers. The author considers low-wage workers only and forms different control groups from the low-wage population. The results confirm the suspicions – when comparison groups are more carefully selected, estimates are virtually zero. The evidence from Canada in general confirms the ambiguity over the true effects of minimum wage on employment that dominates the debate over minimum
DUBRAVKA COLIC

wages in the United States.

Rama’s (2001) study of the Indonesian labor force takes advantage of a large change of the minimum wage (which tripled in nominal terms and doubled in real terms) in Indonesia. Using data from a series of labor force surveys, the author finds modest impacts on the Indonesian labor market, with increases in average wages of 5-15 percent and a decrease in urban employment of 0-5 percent. Rama also found that employment effects varied substantially by firm size, decreasing employment most severely at small firms and even increasing slightly for larger firms, indicating that workers at those firms gain the most from the minimum wage increase both in terms of wages and employment.

In the United Kingdom, several studies by Machin and Manning (1994, 1996) took advantage of the wage-setting policies of the Wage Councils, which set different minimum wages in different industries pre-1986 and introduced a single minimum wage in 1986. The authors find that increases in the minimum wage are not connected with drops in employment. Indeed, they find that large minimum wage increases lead to faster employment growth, although this result is not quite significant at the conventional levels.

Compelling empirical evidence on effects of minimum wage on the labor markets in Eastern and Central European countries is scarce at best. Several important papers have been published explaining in detail the wage-setting practices in the countries of the region, as well as descriptions of general trends in the ability of minimum wages to provide a sufficient income. In Minimum Wages in Central and Eastern Europe: From Protection to Destitution (Standing and Whitehead, 1995), for example, the authors tie together trends from Eastern Europe and support their recommendations for increases in the generosity of labor market provisions by reviewing evidence from Western Europe which indicates a lack of clear evidence on the negative effects of minimum wage. One study, from the Czech Republic, even uses time series evidence to isolate the effect of the 1992 increase in the minimum wage in this country. Like so many other attempts do, the author finds that the modest increase had no appreciable effect. Although the minimum wage covered a small fraction of workers nationwide, the numbers in select industries were more substantial.

V. Econometric Model and Data Description

The Model

The purpose of this thesis is to estimate the potential impact of minimum wages on employment and inequality by taking advantage of differences in minimum wage levels across countries and over time. It is therefore crucial to construct measures of the minimum wage and employment/inequality which can be used to compare data both between countries with different currencies, standards of living, inflation, etc. and across time. To normalize the variables of interest, I use the ratio of employment to the size of the population (the employment rate) and the ratio of the minimum wage to the average wage to allow for easier comparison. For regressions measuring the impact on inequality, I simply use the Gini coefficient, which was constructed precisely to allow for cross-country comparisons and ranges from 0 to 1, with 1 reflecting a perfectly unequal society. In order to
isolate the respective elasticities and more easily compare my results to those of other economists, I use the log of both the dependent and the key independent variable, the ratio of minimum to average wage, in my regressions.

In evaluating the effect of changes in the minimum wage on employment or inequality in Eastern and Central Europe, we need a data set consisting of information on the minimum and average wage levels, as well as employment and population figures in countries of interest in a particular year. We could include any other explanatory variables such as the inflation rate or the size of the economy. The model would be as follows:

\[
\text{emp/pop} = \beta_0 + \beta_1 \times \text{min/avg} + \beta_2 \times X + \epsilon
\]

where emp/pop is the ratio of employment to population size, min/avg is the ratio of minimum to average wage, Xi is a vector of other explanatory variables, and \( \epsilon \) is the error term, all in a particular country i.

This cross-sectional analysis is problematic because excluding one or more explanatory variables could introduce bias into our estimates. In other words, if instead of the above, the correct model should take the form of:

\[
\text{emp/pop} = \alpha_0 + \alpha_1 \times \text{min/avg} + \alpha_2 \times X + \alpha_3 \times Z + u
\]

where Z is the excluded variable related both to min/avg and emp/pop, the expected value of \( \alpha_2 \) in the model we estimate will be \( \alpha_2 + \alpha_3 \times b_z \), where \( b_z \) indicates the impact of variable Z on min/avg. Sometimes the excluded variables are simply not observable, such as a population’s work ethic or compliance with laws, all of which could influence both employment figures and minimum wage legislation. Unobservable heterogeneity among the countries of Eastern and Central Europe would thus introduce bias into our regression results. Although the countries of Eastern and Central Europe are similar in many ways, certainly there exist some inherent differences in employment rates and inequality which would complicate our analyses and bias the results from a simple cross-sectional regression.

Fortunately, a panel data set allows us to account for the existence of unobservable heterogeneity in a model and use a framework with country and time-specific effects to obtain unbiased coefficient estimates. In a panel data set, each observation has a country index (i) and a time index (t); if we specify the model to take the form

\[
\text{emp/pop}_{it} = \beta_0 + \gamma_i + \gamma_t + \beta_1 \times \text{min/avg}_{it} + \beta_2 \times X_{it} + \epsilon_{it}
\]

where min/avg is the value of the ratio of minimum to average wage in country i and year t, Xit is a vector of explanatory variables, the term \( \gamma_i \) represents the unobservable country-specific component, and \( \epsilon_{it} \) represents unobserved time specific components that affect all countries in the sample uniformly. In the fixed effects model, the country-specific term shifts the intercept for each country in the data set, acting like a dummy variable. For a two-period data set, we can eliminate the unobservable heterogeneity by subtracting the variables one period from the next and using the method of first differencing. For 15 years of data, the process is somewhat different. The unobservable components are eliminated by including dummy variables for each country and year.

The model can be augmented easily for an estimate of the effect of minimum wage lev-
els on the Gini coefficient. It takes the form of
\[ \text{gini}_{it} = \delta_0 + \delta_1 \times \frac{\text{min/avg}_{it}}{\text{avg}_{it}} + \delta_2 \times X_{it} + \varphi_i + \varphi_t + \psi_{it} \]

where \( \text{gini}_{it} \) is the Gini coefficient in country \( i \) and year \( t \), and other notation is analogous. The procedure for estimating this fixed effects model is also appropriate here, resulting in unbiased estimates of the impact of the log of ratio of minimum to average wage on inequality. From a theoretical standpoint, this model may have the problem of endogenous variables if we accept that one of the motivations behind the minimum wage is to increase the income of low-wage workers and thus improve wage inequality. It may be the case, then, that countries with high inequality are likely to pass more generous minimum wage legislation in order to combat inequality, but this potential problem seems less likely to occur in my analysis precisely based on my earlier discussion regarding the likely exogeneity of minimum wage legislation in ECE countries.

**The Data**

Without a doubt, one of the major obstacles in conducting research in Eastern and Central Europe (ECE) is the difficulty with which reliable data can be obtained. Most reputable databases, such as the World Development Indicators or UN Social Indicators, contain information on ECE countries for only the last few years. Data for the pivotal early 1990s, when most labor laws were reshaped, are almost never available for labor market indicators such as employment or labor force size. Even when the information is available, it is often at odds with countries’ self-reported figures. For example, the range of fluctuations in population and labor force sizes reported in comprehensive international databases is often systematically different from the country data for any given variable. Because of this phenomenon and because of a need for consistency in related groups of data, statistical yearbooks for each country turned out to be the most reliable and complete source. Other major sources include the World Bank, the United Nations, and the European Union (for more recent data on candidate and member countries) among others.

The data set covers the years 1989 through 2004 for the following countries: Belarus, Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Moldova, Poland, Romania, Russia, Slovakia, Slovenia, and Ukraine. Countries of the Balkans, as well as Georgia were not included mainly because of the lack of data, which is linked to the disruptions caused by civil wars in these countries. Data for the selected countries will therefore provide evidence on how minimum wage levels can affect labor market outcomes in transition economies with relatively minor political or economic disruptions, allowing for an easier comparison with existing evidence from North America and Western Europe.

Information on the population, the labor force, and employment in each country for the time period covered are obtained from country statistical offices, either through the statistical yearbook (which the offices issue), or through their website when data is available online. In rare instances, direct inquiries with the statistical offices resulted in a response; for example, a large portion of the data for Poland was obtained in this way.

For the most part, recording the labor market data itself was straightforward. Most countries have the labor force figures broken down into the number of employed and unemployed by gender or region. In some instances, only labor force and employment fig-
ures were displayed clearly, and the fractions had to be constructed either by simple division of the number of women by the total number, or from labor force participation rates and employment rates for women. For example, the fraction of the employed that are women in Romania was constructed by finding the number of women in the country from population data, multiplying that number of women by their employment rate (thus finding the number of women who are employed) and dividing the product by the total number of the employed.

Average monthly wages and the size of the population were easily available, and often came with revised figures provided in the form of a convenient time series. One common problem was that employment and labor force figures would sometimes not only differ across statistical yearbooks (for the same year), but would not add up even within a single statistical yearbook. This appears to be a result of exclusions of certain employees from employment figures – anyone from women on maternity leave to very young workers or retirees would be excluded from certain tables and not others, with no clear pattern. The issue was addressed by consistently picking the same table from which to record data for a given country, since the country-specific effects in the model will be able to account for lower employment or labor force figures in a particular country due to the data specification in the yearbooks, as long as the specification is consistent.

The minimum wage level, the crucial explanatory variable, was generally not available directly through the statistical offices, except in the case of Lithuania. Since the minimum wage is a legislated and clearly specified figure, a single consistent source was not a necessary prerequisite for accuracy. The information came from a variety of sources, most notably from studies by the European Industrial Relations Observatory (EIRO), the World Bank, and the State Department through the Country Reports on Human Rights Practices. Other sources include the European Union, International Labor Organization, Database Central Europe, and the UN’s TransMONEE databases for 2004 and 2005. For the most part, the minimum wage levels were available in the local currency, but it had to be converted from euro or dollar amounts in a few instances. Most countries amend the laws on minimum wage once or twice a year, but the majority of minimum wage figures are either unweighted averages of all values in a given year or represent a level in effect for the largest part of the year.

The data on average and minimum wages had to be converted for nearly every country because of redenomination and currency changes that occurred during the 1990s. The information on factors by which the currency changed (or replaced an old currency) was available through most of the statistical yearbooks to explain which currency was used for the given year in cases where the statistical office did not place already converted values in the yearbooks. The conversion factors ranged from 1000 in Bulgaria and Romania to 100,000 in Poland, so detecting which figures were already converted and which were not was relatively straightforward.

The United Nations TransMONEE database provided data on the Gini coefficient of earnings in countries of Eastern and Central Europe for the analysis of minimum wage effects on inequality. Data for all years are not present for all countries, but the informa-
tion from this database was by far the most complete of all cross-country data sources. Examining changes in inequality and poverty in ECE countries is particularly challenging because most countries change poverty lines through history, which complicates comparisons across time. Other measures of poverty include the number of people living on less than $1 (in some cases, $2) per day, but since a dollar has drastically different purchasing power in different countries of this region, the Gini coefficient provides the most objective tool for comparisons across both countries and time.

Apart from labor market data, the data set also includes a series of dummy variables for political and economic changes that have occurred in the chosen countries. The variables include political turmoil (such as the collapse of the former Soviet Union, the separation of Czech Republic and Slovakia in 1993, or significant regime shifts which could result in changes in minimum wage), whether or not a country is a member of the NATO pact, and whether a country is a candidate for membership in the European Union. Since the first ECE countries to join the EU did not officially become members until 2004, being a candidate country is a better measure of any economic advantages to cooperation with the European Union. The TransMONEE database for 2005 provided several control variables, such as the inflation rate, GDP per capita, GDP growth, etc.

VI. Results

Effect on the Employment Rate

The bare bones model for estimating the effect of minimum wages on the employment rate contains only the ratio of minimum to average wages on the right hand side, along with the country and time-specific fixed effects. In other words, no other explanatory variables are included at first. Results show that the ratio of minimum to average wage along with the fixed effects accounts for approximately 76 percent of the variation in the employment rate. The estimated elasticity is -.017 (significant at the 1 percent level) indicating that a one percent increase in the ratio of minimum to average wage decreases the employment rate by .017 percent, consistent with compelling evidence from case studies by Card and Krueger (New Jersey fast food industry study), as well as Katz and Krueger (California retail industry study).

Further variations on the model used in my estimations are as follows. I first augmented the basic model by adding several variables which control for major political and economic shifts that often dominate the politics of minimum wage or have a significant impact on employment in Eastern and Central Europe. The variables include candidacy to join the European Union (since the first countries formally became part of the EU only in 2004, the last year in my sample, and candidacy itself does imply economic and political ties with the European Union), membership in the NATO pact, as well as a variable constructed from historical, political, and economic histories of my countries of interest. These variables should help distill the effects of changes in the minimum wage on the employment rate from other influences on the dependent variable.

I then further expand the model to include two macroeconomic variables: GDP growth and the inflation rate. Economies which are growing fast will likely also experience more
dramatic changes in the labor force, in the form of increases or restructuring in the employed population. The inflation rate, on the other hand, significantly changes how the minimum wage affects the employed and, subsequently, those who may lose their jobs due to the increases in the minimum wage. The notion of “sticky wages” is particularly important here; nominal monthly wages in the region usually adjust much faster than the legislated level of the minimum wage, which means that the ratio of minimum to average wage deteriorates faster than intended due to high inflation. This phenomenon is particularly pronounced in the years of hyperinflation, which reached thousands of percent in countries like Belarus, Ukraine, and Moldova. The inflation rate, along with the growth of the economy, therefore most likely belongs in the model as a control variable.

The main results of the basic regression, as well as the two expanded models, are presented in the table below:

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum wage</td>
<td>-.015</td>
<td>-.015</td>
<td>-.010</td>
</tr>
<tr>
<td>(.003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Political</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Macroeconomic</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Country fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R2</td>
<td>0.780</td>
<td>0.782</td>
<td>0.824</td>
</tr>
<tr>
<td># observations</td>
<td>157</td>
<td>157</td>
<td>155</td>
</tr>
</tbody>
</table>

The coefficient estimate on the minimum wage variable is significant at the 1 percent level for each specification. The variables included in each model are noted in the table. Results change only negligibly when variables for political and economic change are included, but the magnitude of the coefficient falls by one third when inflation and GDP growth are added as explanatory variables, indicating that some of the variation in the employment rate could be explained by those variables. The R2 increase from .782 to .824 confirms this conclusion. The regression results therefore indicate that a 1 percent increase in the ratio of minimum to average wage decreases the employment rate by .01 percent. Perhaps more intuitively, a 10 percent increase of the ratio of minimum to average wage in 1997 (so, from 26.62 percent to 29.28) would lead to a decrease in the employment rate of .1 percent, or from 42.08 percent to 42.04. Alternatively, doubling the minimum to average wage ratio at the outset of reform in 1995, from 27.3 percent to 54.6, would decrease the employment rate from 42.7 percent to 42.27.

In order to gauge whether the results would be consistent for a different set of countries in Eastern and Central Europe, I perform the same estimation using data for only the countries of the former Soviet Union; the results are practically identical. The bare bones model produces an elasticity of -.017, with a decrease in the coefficient to -.012 once polit-
ical and macroeconomic variables are included in the regression. The size of the R^2 is comparable to that in the previous table, with the number of observations consistent across models. The results are presented in Table 6.2 below.

For a further robustness check, I run these same regressions (both for all countries and the subset of former Soviet republics) without using the population as a weight. The use of this weight is called into question in some literature, and could potentially be an issue in my analysis because Russia is by far the largest country in the region and that with the largest drop in the minimum wage. The concern, then, is that the negative elasticities are driven by the particular experience of the Russian economy. The estimates show consistent results when no weight is employed in the regressions, indicating that the experience of falling employment rates in response to increasing minimum wage levels is not specific to Russia or even to former Soviet countries alone. The main regression results using data from all counties and employing no weight is presented below:

**Table 6.2: Elasticity of the Employment Rate (former Soviet Union only)**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum wage</td>
<td>-.017</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.003)</td>
<td>-.012</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Political</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Macroeconomic</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Country fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R^2</td>
<td>0.763</td>
<td>0.770</td>
<td>0.790</td>
</tr>
<tr>
<td># observations</td>
<td>70</td>
<td>70</td>
<td>70</td>
</tr>
</tbody>
</table>

**Table 6.3: Elasticity of the Employment Rate (no weight employed)**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum wage</td>
<td>-.024</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.009)</td>
<td>-.023</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.008)</td>
<td>-.018</td>
<td></td>
</tr>
<tr>
<td>Political</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Macroeconomic</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Country fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R^2</td>
<td>0.828</td>
<td>0.836</td>
<td>0.850</td>
</tr>
<tr>
<td># observations</td>
<td>157</td>
<td>157</td>
<td>155</td>
</tr>
</tbody>
</table>

**Effect on the Gini Coefficient**

Results for estimates of minimum wage effects on inequality are not as clear as those
for the employment rate. When population is used as a weight, the regression results indicate an increase in inequality as measured by the Gini (of .002 with all control variables included) in response to a one percent increase in the ratio between minimum and average wage. This estimate, however, is not statistically significant. The results are presented in Table 6.4 below.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum wage</td>
<td>.001</td>
<td>.005</td>
<td>.002</td>
</tr>
<tr>
<td>Political</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Macroeconomic</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Country fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R²</td>
<td>0.961</td>
<td>0.966</td>
<td>0.976</td>
</tr>
<tr>
<td># observations</td>
<td>110</td>
<td>110</td>
<td>108</td>
</tr>
</tbody>
</table>

When no weight is employed, the elasticity of the Gini coefficient becomes negative, indicating that higher minimum wages do decrease inequality. This change makes sense in light of the Russian trend in inequality – namely, Russia’s increase in the Gini coefficient from .27 in 1989 to .52 in 2001 – more pronounced than in any other country of the region. Since Russia’s population is by far the largest in Eastern and Central Europe, the positive elasticity of the Gini in previous regressions was likely driven mostly by Russia’s unique experience, which is certainly not reflective of the entire region. The control variables (along with country and year fixed effects) included in my regressions cannot account for the entirety of the variation in Russia’s Gini, so the minimum wage coefficient was likely picking up other effects on inequality. Although the coefficient estimate is still not statistically significant at the conventional levels when no weight is employed in the regressions, at least the elasticities are consistent with theoretical predictions outlined in Section III. Results from these regressions are presented in Table 6.5

**Limitations and Extensions**

Although the above results are consistent and robust to alternate specifications, the model is still imperfect. The culprit, largely, is the difficulty of obtaining data for the countries of Eastern and Central Europe. To account for as many effects on the employment rate as possible, I would have liked to include several variables for which data was either not available at all, or so scarce that including these data would significantly reduce the number of observations and introduce doubt in the results – in particular, several labor market variables such as the extent and duration of the unemployment insurance benefit. The generosity of the unemployment insurance system affects similar workers as the minimum wage, and could influence both firms’ hiring practices and individuals’ decisions to
enter (or remain in) the labor market while earning the minimum wage. Another variable of interest is the extent of government involvement in the economy - if a government has a large stake in the employment of its population, it may be more inclined to keep minimum wage levels low, so this information might be important as well.

Inclusion of the inflation rate in the last model helped control for the presence of high inflation rates which would deteriorate the social protection aspect of the minimum wage. However, the inflation rate alone cannot account for all channels of high inflation’s impact on wages and employment. Another crucial variable which goes hand in hand with inflation is the extent of unionization of workers in countries of Eastern and Central Europe. In most countries, labor unions were strong before the collapse of communism and have remained influential in some countries, but become almost entirely obsolete in others. In the context of my models, their influence is felt most strongly through the extent to which protection from inflation is included in wage legislation. Stronger unions will refuse to negotiate on nominal wages alone and will demand that wages increase together with inflation. Since it is not always the case that countries with stronger unions extend protection from high inflation to minimum wages as well, there exists a possibility of effects on employment through minimum wage which should really be attributed to other protection from inflation, seeing as the ratio of minimum to average wages will evolve differently under the influence of strong unions. Including a variable which measures the extent of unionization in my regressions would better account for this influence.

**Table 6.5: Elasticity of the Gini Coefficient (All Countries, No Weight)**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Minimum wage</strong></td>
<td>-.004</td>
<td>-.003</td>
<td>-.007</td>
</tr>
<tr>
<td>(.015)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.015)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.016)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Political</strong></td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Macroeconomic</strong></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Country fixed effects</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Year fixed effects</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>R2</strong></td>
<td>0.879</td>
<td>0.883</td>
<td>0.894</td>
</tr>
<tr>
<td><strong># observations</strong></td>
<td>117</td>
<td>117</td>
<td>115</td>
</tr>
</tbody>
</table>

When it comes to the effect of minimum wage on inequality, it is puzzling that despite the clear drop in employment when the ratio of minimum to average wage increases, an equally clear improvement in inequality is also not detected – lowest-paid workers are presumably removed from the wage distribution, which should, theoretically, improve inequality. A potential problem is certainly the source of data for the Gini coefficient – while all labor market data comes from statistical offices, the information on the Gini is from World Bank’s TransMONEE database. As mentioned earlier, data from sources like the UN or World Bank are sometimes systematically different from the countries’ self-reported values, which could introduce measurement error. While errors in the measure-
EFFECTS OF MINIMUM WAGE LEGISLATION ON EMPLOYMENT

ment of the dependent variable do not generally introduce bias into our results, it is possible that they create difficulty in detecting an effect on inequality in light of the extremely small drop in employment when minimum wages increase. Additionally, the availability of the data on the Gini coefficient in this database essentially claims that information on the wage distribution is available every year for each country, which is unlikely since this would require yearly labor force surveys.

Finally, the analyses in this thesis rely heavily on the assumption that minimum wage policy was largely exogenous in countries of Eastern and Central Europe during the transition period. It is certainly possible that some of the increases in minimum wage were not entirely dissociated from similar policies, particularly before elections. If politicians are attempting to draw in votes, for example, they may do so by increasing the minimum wage in order to attract those already in the labor force, but also expand the number of government jobs in the economy at the same time to please the unemployed – both policies would affect employment. While it is possible that my data set encompasses such events, it still remains that the majority of the changes in minimum wage were exogenous to the labor market, so Eastern and Central Europe remain a fruitful ground for evaluation of labor market policies. The next step in the study of minimum wage in ECE countries, however, should certainly include a careful consideration of the exogeneity of labor market legislation alongside the search for better data in order to take advantage of the events in this tumultuous region to bring new evidence of government policy on the labor market.

VII. Conclusion

Eastern and Central Europe have seen tremendous changes since 1989 in the areas of politics, economics, and societal organization. The institution of the minimum wage has been the source of much experimentation during the transition period of the early 1990s. The deterioration of the minimum wage relative to the average wage in some countries - to as low as 10 percent in Russia - and its increase dramatically in others - to as high as 64 percent of the average wage in the case of Bulgaria - provide the extensive variation across countries that allows for an interesting study of the impact of minimum wages on employment and on the Gini coefficient.

The classical model of the labor market effects of the minimum wage predicts unambiguously that an increase in the minimum wage will reduce employment. The empirical evidence, however, does not always support this prediction. Results of case studies using increases in state levels of minimum wage in the US cannot support the predicted drop in employment, with the most recent panel data evidence even finding employment increases in response to a hike in minimum wage. The evidence on the employment effects of minimum wages thus remains highly controversial. A few studies on the impact of minimum wage on inequality, however, do detect a reduction in inequality in the presence of a minimum wage. In this thesis, I apply the techniques developed in these studies to the case of Eastern and Central Europe.

Estimating a fixed effects model using a panel data set of Eastern and Central European countries, I find that a one percent increase in the ratio of minimum to average
wage leads to a decrease in the employment rate on the order of .01 percent. The coefficient is statistically significant at conventional levels for every specification, and remains small, negative, and statistically significant when the regressions are restricted to the sub-sample of former Soviet republics. Minimum wage as a percentage of average wage cannot be said to have a statistically significant effect on inequality measured by the Gini coefficient; the estimate is also positive in some cases, indicating that higher minimum wages may worsen inequality, which is contrary both to economic theory and existing empirical evidence.

Although there is much room for improvement of the econometric model used in this paper through the data available and the selection of control variables, the consistent and robust results of the analysis so far provide convincing evidence that the impact of minimum wages on the employment rate is not particularly strong even in response to large swings in the minimum wage. Even doubling the ratio of the minimum wage to the average wage (as in the 1995 regional average example from Section VI) would decrease the employment rate by only .043 of a percentage point. However, within the group of workers who earn minimum wage in Eastern and Central Europe, the employment impact on the targeted group may not be acceptable by social standards. For a labor force where roughly 5 percent of workers earn the minimum wage, it would indicate that doubling the minimum wage ratio would result in four out of every hundred workers in this wage group losing their job, assuming low wage workers the ones getting laid off when minimum wage rises.

When a positive impact of minimum wages on inequality cannot be shown as statistically significant, the countries of the region should certainly carefully weigh the benefits of higher social protection against the efficiency costs and job loss among those lowest-paid workers. Most countries rarely raise minimum wages as dramatically as by 100 percent of the minimum to average ratio, so a modest increase in the minimum wage should have a negligible effect on the employment rate, and perhaps a more modest effect on the employment of low-wage workers.

References


Over the past decade, the popular press has perpetuated a depiction of an African continent ravaged far and wide by the HIV epidemic. In fact, many parts of Africa still have HIV prevalence rates lower than 1 percent. However, some parts of Africa are also growing very quickly. Economic growth often brings disruption and disease along with it, but it can also lift people out of poverty and allow greater investment in health. Thus, theory provides no clear answer as to how economic growth should affect HIV prevalence, but that answer will have serious implications for development in Africa and elsewhere.

In this paper, I attempt to determine the net effect of economic growth rates on HIV prevalence rates in the developing world. I then deconstruct this effect to identify its major components. Although there have been a number of studies on the determinants of HIV prevalence rates (Bates 2004; Garnett 1998; Poundstone 2004), the specific relationship between economic development and HIV rates remains unexplored. In an unpublished draft paper examining the structural determinants of HIV prevalence rates in Sub-Saharan Africa, Tsafack Temah (2004) finds that income level is an insignificant determinant of HIV prevalence rates. However, the study does not examine the reduced-form relationship between national income and HIV prevalence.

The data, which come from the United Nations, discuss the relationship between HIV prevalence rates and economic growth and seem to suggest a non-linear relationship. There are a number of ways to model non-linearity. Following Coleman, Rayner, and Bates (1997), I attempt to model the data as a Kuznets Curve. Since the data are not obviously U-shaped, however, I introduce another form of nonlinearity using a power function, a model that lends itself to transformation for linear regression and provides a great deal of flexibility in shape.

The primary results are reported in Tables 3 and 4. Of three specifications tested, only the power functional form produces a significant coefficient on national income. A random effects regression that applies a regional dummy for Sub-Saharan Africa mimics the fixed-effects regression while maintaining the sufficient degrees of freedom. When ln(Y_{it}) increases by 1, the HIV prevalence rate increases by 15 percent. That is, when a country’s GDP per capita roughly triples, its HIV prevalence rate will be 1.15 times as large.
I explore the channels by which national income acts on HIV prevalence using simultaneous equations. Limitations in the data do not permit panel analysis. Under a cross-sectional regression using HIV and GDP data from 2001, the relationship between HIV prevalence rates and GDP per capita becomes slightly more significant, but the coefficient changes sign. HIV seems to have been in most countries of the world for roughly the same amount of time. Countries that were at high income levels during the advent of the disease, and still are today, are likely to have healthier populations, and in turn experience lower transmission rates from a given sexual contact (Stillwagon, 2000). If this is the case on a large scale, GDP per capita will be negatively correlated with HIV prevalence in a cross-section, but positively correlated with HIV prevalence over time. Nevertheless, the way the cross-sectional regression reacts to changes in the specification can give information about the importance of the different channels.

National income has a significant effect on each of the channel variables studied. My regression of HIV prevalence rates on the channel variables, controlling once again for refugee populations and democracy, yields two significant results. Migration rates and malnutrition rates have a positive impact on HIV prevalence rates. The coefficient on women in the labor force, a variable shown previously to be an important determinant of HIV prevalence rates (see Tsafack Temah 2004), is insignificant.

Finally, I regress the channel variables, the control variables, and GDP per capita on HIV. The results, reported in Table 7, suggest that migration and malnutrition do, in fact, act as important channels for the effect of GDP per capita on HIV prevalence rates.

Limitations in the data engender concerns about omitted variables. Official development assistance (ODA) is likely to be correlated with economic growth, and countries receiving a great deal of ODA are also likely to receive funds for the prevention and treatment of HIV. Ethnolinguistic fractionalization (ELF), which is highly correlated with ethnic fractionalization, has been shown to be correlated with corruption, which then has an impact on economic growth (Mauro, 1995). The results, shown in tables 8 and 9, indicate that these omitted variables do not have a significant impact on HIV prevalence rates. Their inclusion renders the coefficient on national income insignificant simply because of the reduction in sample size.

Reverse causality is also a concern in the study. Bonnel (2000) found that Africa’s average HIV prevalence rate of 8 percent reduced its GDP per capita growth rate by approximately 1.1 percent. However, if there is a problem of reverse causality, my positive result should only be an underestimate. More robust analysis should point to an even greater impact of economic growth rates on HIV prevalence rates.

Nevertheless, I attempt to address the problem of reverse causality in two ways. I estimate the coefficient on 5-year-lagged growth rates (Table 10). This result does not contradict the primary finding that economic growth has a weakly positive impact on HIV prevalence, but it does indicate that the result is very sensitive to changes in the specification. Secondly, I introduce distance from the equator as an instrumental variable for economic growth. I use latitude data gathered from the Central Intelligence Agency’s annual World Factbook. Under the preferred specification, which includes the Sub-Saharan Africa (SSA)
dummy, the coefficient on GDP per capita becomes insignificant, but it remains positive and becomes slightly larger. When the SSA dummy is left out, however, the coefficient on national income becomes highly significant and, more importantly, changes sign. The negative effect suggested here is smaller than the positive coefficient seen in the preferred specification, but the negative result is slightly more significant.

These results are self-contradictory, however, for the same reason that the panel and cross-sectional results were different. The random effects estimation that includes the regional dummy mimics a fixed effects regression in controlling for socio-cultural effects, so that the main source of variation is time. The random effects specification that excludes the regional dummy, on the other hand, allows for a great deal of cross-country variation. Because the panel is extremely wide, it is likely that the cross-sectional effect dominates the vertical effect over time.

These results suggest that while economic growth may have some overall impact on HIV prevalence, its effect is negligible. Using three different specifications and two estimation methods for each specification, I was only able to achieve one significant result. The result, moreover, implies that a near tripling in the level of GDP per capita would only cause the HIV prevalence rate to grow by 15 percent. In low prevalence countries, like those in West Africa, this difference is barely noticeable. Even in higher prevalence countries, the time frame is so lengthy that prediction is almost irrelevant.

The channel analysis is more enlightening. In particular, it reveals why economic growth might have such a small effect on HIV prevalence: migration and improved nutrition work in opposite directions to moderate the effect of economic growth on national HIV prevalence rates. However, even if nutrition improvement is moving some segments of the population away from HIV infection, those involved in migrant labor or transport industries are still at a very high risk of contracting the disease. Therefore, it will be important for policymakers to continue to target prevention efforts at these sectors of society.

The insignificance of urbanization is surprising, but it is likely that the effects of urbanization are captured by the international migration regressor. In previous studies that have found urbanization to be highly significant, it is the only indicator of spatial mobility studied (Dyson 2003, Tsefack Temah 2004). The policy implication of my result is that HIV prevention programs must be implemented along migration routes, not just in cities, since migrants seems to be bringing the disease along with them.

Finally, the strong significance of the regional dummy for Sub-Saharan Africa in both the panel analysis and the cross-sectional analysis has important implications for how policymakers understand the epidemiology of HIV/AIDS. It is likely that the regional dummy is capturing the importance of STI coinfection. Oster (2005) identified STI coinfection as the largest source of variation in transmission rates between the United States and Sub-Saharan Africa. Africans have the highest rate of untreated curable STIs in the world (Poku 2000). The consistent significance of SSA supports the conclusion that fighting other STIs, particularly ones that cause genital sores, could be one of the most effective HIV prevention measures.

The greatest limitation to this study is the lack of data. Future research on this topic
should use a third year of UNAIDS HIV prevalence data, which alone could vastly improve the robustness of these regressions. The dichotomy between migrant stocks and urbanization also merits further study. Both of these data sets are imperfect, and research using more precise measures of urbanization or migration might better explain why, or if, one matters more than the other. That result could have important implications for HIV prevention policies, growth-promoting strategies, and development in general.

<table>
<thead>
<tr>
<th>Table 3: Quadratic Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation type</td>
</tr>
<tr>
<td>Dep. Variable</td>
</tr>
<tr>
<td>cgdp</td>
</tr>
<tr>
<td>cgdp2</td>
</tr>
<tr>
<td>refpop</td>
</tr>
<tr>
<td>freedom</td>
</tr>
<tr>
<td>ln cgdp</td>
</tr>
<tr>
<td>ln cgdp2</td>
</tr>
<tr>
<td>ln ref pop</td>
</tr>
<tr>
<td>ln freedom</td>
</tr>
<tr>
<td>ssa</td>
</tr>
<tr>
<td>_cons</td>
</tr>
<tr>
<td>N</td>
</tr>
</tbody>
</table>

Legend: * p<.1; ** p<.05; *** p<.01

<table>
<thead>
<tr>
<th>Table 4: Power Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation type</td>
</tr>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Lncgdp</td>
</tr>
<tr>
<td>Lncgdp2</td>
</tr>
<tr>
<td>Ln ref pop</td>
</tr>
<tr>
<td>Ln freedom</td>
</tr>
<tr>
<td>Ssa</td>
</tr>
<tr>
<td>_cons</td>
</tr>
<tr>
<td>N</td>
</tr>
</tbody>
</table>

Legend: * p<.1; ** p<.05; *** p<.01
**Table 5: HIV and GDP per Capita Cross-Section**

<table>
<thead>
<tr>
<th>Dep. Variable</th>
<th>lnhiv</th>
<th>lnhiv</th>
</tr>
</thead>
<tbody>
<tr>
<td>lncgdp</td>
<td>-.54623175***</td>
<td>.19184034</td>
</tr>
<tr>
<td>lnrefpop</td>
<td>.11868401</td>
<td>.03433238</td>
</tr>
<tr>
<td>lnfree</td>
<td>.49620819</td>
<td>.49645453*</td>
</tr>
<tr>
<td>Ssa</td>
<td>2.946207***</td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>3.1883362***</td>
<td>-3.1329091***</td>
</tr>
<tr>
<td>N</td>
<td>68</td>
<td>68</td>
</tr>
</tbody>
</table>

Legend: * p<.1; **p<.05; *** p<.01

**Table 6: GDP per Capita and Channel Variables**

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>Migr</th>
<th>urb</th>
<th>health</th>
<th>malnutr</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>cgdp</td>
<td>.00115172***</td>
<td>.00328107***</td>
<td>.00012975***</td>
<td>-.0042716***</td>
<td>-.00052372***</td>
</tr>
<tr>
<td>_cons</td>
<td>2.6853095***</td>
<td>42.626386***</td>
<td>2.4862744***</td>
<td>27.675487***</td>
<td>41.479085***</td>
</tr>
<tr>
<td>N</td>
<td>101</td>
<td>101</td>
<td>101</td>
<td>93</td>
<td>100</td>
</tr>
</tbody>
</table>

Legend: * p<.1; **p<.05; *** p<.01

**Table 7: Channel Variables and HIV**

<table>
<thead>
<tr>
<th>Dep.Variable</th>
<th>Inhiv</th>
<th>Inhiv</th>
<th>Inhiv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lnlnmigr</td>
<td>.29468023**</td>
<td>.01676815</td>
<td>.29378617**</td>
</tr>
<tr>
<td>Lnlnurb</td>
<td>-.28221182</td>
<td>-.25811114</td>
<td>-.21029089</td>
</tr>
<tr>
<td>Lnlnhealth</td>
<td>.28892178</td>
<td>.28606323</td>
<td>.32558285</td>
</tr>
<tr>
<td>Lnlnmalnutr</td>
<td>.98328954***</td>
<td>.16164292</td>
<td>.90336581***</td>
</tr>
<tr>
<td>Lnlnfem</td>
<td>.72527831</td>
<td>1.3942107</td>
<td>.72946983</td>
</tr>
<tr>
<td>Lnlnrefpop</td>
<td>.08481551</td>
<td>.06024943</td>
<td>.08296404</td>
</tr>
<tr>
<td>Lnlnfree</td>
<td>.13950077</td>
<td>.5473593*</td>
<td>.1585257</td>
</tr>
<tr>
<td>LnSsa</td>
<td>2.3989664***</td>
<td>2.8859715***</td>
<td></td>
</tr>
<tr>
<td>Lnlnfree</td>
<td>-.11202372</td>
<td></td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>-4.8192287</td>
<td>-6.6786658*</td>
<td>-4.1849571</td>
</tr>
<tr>
<td>N</td>
<td>64</td>
<td>64</td>
<td>64</td>
</tr>
</tbody>
</table>

Legend: * p<.1; **p<.05; *** p<.01

**Table 8: Impact of Omitted Variables**

<table>
<thead>
<tr>
<th>Estimation type</th>
<th>random</th>
<th>random</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dep. Variable</td>
<td>lnhiv</td>
<td>lnhiv</td>
</tr>
<tr>
<td>lncgdp</td>
<td>.08818336</td>
<td>.05486723</td>
</tr>
<tr>
<td>lnrefpop</td>
<td>.00924172</td>
<td>.00237458</td>
</tr>
<tr>
<td>lnfree</td>
<td>.08558543***</td>
<td>.03549994***</td>
</tr>
<tr>
<td>Ssa</td>
<td>2.6985599***</td>
<td>2.7103021***</td>
</tr>
<tr>
<td>lnoda</td>
<td>.00138087</td>
<td></td>
</tr>
<tr>
<td>lnethnic</td>
<td></td>
<td>-.35543729</td>
</tr>
<tr>
<td>_cons</td>
<td>-1.6921039***</td>
<td>-1.6229863***</td>
</tr>
<tr>
<td>N</td>
<td>116</td>
<td>127</td>
</tr>
</tbody>
</table>

Legend: * p<.1; **p<.05; *** p<.01
## CAN DEVELOPMENT CAUSE AIDS?

### Table 9: Significance of omitted variables

<table>
<thead>
<tr>
<th>Dep. Variable</th>
<th>lnhiv</th>
<th>lnhiv</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnretpop</td>
<td>.00178756</td>
<td>.007822</td>
</tr>
<tr>
<td>lnfree</td>
<td>.03556136***</td>
<td>.0872453***</td>
</tr>
<tr>
<td>ssa</td>
<td>2.6313062***</td>
<td>2.5697277***</td>
</tr>
<tr>
<td>lnethnic</td>
<td>-.36743679*</td>
<td></td>
</tr>
<tr>
<td>lnoda</td>
<td></td>
<td>.00664113</td>
</tr>
<tr>
<td>_cons</td>
<td>-1.2192031***</td>
<td>-1.0582792***</td>
</tr>
<tr>
<td>N</td>
<td>127</td>
<td>116</td>
</tr>
</tbody>
</table>

Legend: * p<.1; **p<.05; *** p<.01

### Table 10: 5-Year-Lagged GDP per Capita

<table>
<thead>
<tr>
<th>Estimation type</th>
<th>lnhiv</th>
<th>lnhiv</th>
</tr>
</thead>
<tbody>
<tr>
<td>lngdp5</td>
<td>.13812205</td>
<td>.10706031</td>
</tr>
<tr>
<td>lnretpop</td>
<td>.00219625</td>
<td>.00196501</td>
</tr>
<tr>
<td>lnfree</td>
<td>.05224449***</td>
<td>.0515838***</td>
</tr>
<tr>
<td>ssa</td>
<td>2.7053734***</td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>-1.9208619***</td>
<td>-.53559597</td>
</tr>
<tr>
<td>N</td>
<td>144</td>
<td>144</td>
</tr>
</tbody>
</table>

Legend: * p<.1; **p<.05; *** p<.01

### Table 11: Distance to Equator Instrument

<table>
<thead>
<tr>
<th>Dep. Variable</th>
<th>lnhiv</th>
<th>lnhiv</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnegdp*</td>
<td>.17180396</td>
<td>-.49412587***</td>
</tr>
<tr>
<td>lnretpop</td>
<td>.00186453</td>
<td>.00193335</td>
</tr>
<tr>
<td>lnfree</td>
<td>.05243719***</td>
<td>.0515838***</td>
</tr>
<tr>
<td>ssa</td>
<td>2.7577196***</td>
<td>.05247559***</td>
</tr>
<tr>
<td>_cons</td>
<td>-2.1834808**</td>
<td>3.556163***</td>
</tr>
<tr>
<td>N</td>
<td>144</td>
<td>144</td>
</tr>
</tbody>
</table>

Legend: * p<.1; **p<.05; *** p<.01

### References


DISCRIMINATION AS A FACTOR OF DEVELOPMENT IN ECUADOR

Natasha Nguyen

University of California—Berkeley

My original goal was to gather survey data in Ecuador similar to the World Bank’s Living Standards Measurement Study and compare it against previous years’ data. During my first few weeks there however, I unexpectedly found the prejudices between classes (and, subsequently, races) to be blatant and omnipresent. The question of discrimination as an active factor in development became a much more compelling question to investigate. Questioning presence and power of discrimination is relevant when discussing tools for development. Trade agreements, in particular, and equalization mechanisms between developed and developing trade partners can be rendered ineffective by internal issues such as discrimination.

By means of empirical evidence, I found that economic development in Ecuador is immobilized by racial discrimination from Spanish descendents against the Indigenous and Afro-Ecuadorian populations. Discrimination was introduced during the early days of the Spanish Conquest and created an earnings gap that is still present today. Lower income levels in the discriminated populations impact healthcare and education quality that restrict access to living wage jobs, resulting in a cyclical poverty trap.

While Afro-Ecuadorians and Indigenous people are legally equal to any other Ecuadorian citizen, they are behind economically. Currently, the average indigenous worker’s wage is 55% of the average non-indigenous worker (World Bank, 2004). Discrimination is the driving force behind the gap. Directly, discrimination in the workplace limits earning opportunities. In fact, the World Bank attributes 74% of the earnings gap to endowment differences, such as educational and occupational resources, and 26% to discrimination (World Bank, 2004). In other words, every one in four qualified but unhired Indigenous person is not hired due to taste-based discrimination. The last three are not hired, at least partially, due to discrimination that compromises schooling quality, social networking, and workforce entry.

How Discrimination Creates the Earnings Gap

The historical social discrimination of the Indigenous and Afro-Ecuadorians from days of the Spanish Conquest immediately put them at an educational disadvantage. During Spanish rule, Indigenous and Afro-Ecuadorians were not considered as equal citizens who deserved equal educational opportunities. Today however, they are required to attend formal schooling until the age of 16 with the majority of the population. Discrimination that
continues in schools today compromises the educational opportunities that these populations do have. Ridicule and ostracism discourages students from actively participating or becoming enthusiastic about school and school-related activities. So while educational opportunities might be improving (World Bank, 2004), students are not able to benefit from their time in school.

Similarly, social networking also suffers long-term consequences from school-place discrimination. When ostracized from social groups, Indigenous and Afro-Ecuadorians are prevented from entering specific social networks. Many jobs are found through connections made during high school and university years. But not being a part of such social networks, Indigenous and Afro-Ecuadorian job opportunities are severely limited. Especially considering that their populations have not had a historical presence in the formal sector, social networking becomes a huge advantage, if not a necessity, in breaking into the formal sector.

Education quality is also negatively impacted by poor healthcare. The Indigenous have considerably lower marks in healthcare quality. According to the World Bank, the amount of professionally-assisted birth deliveries was significantly less for indigenous mothers at 33% compared to 82% for non-indigenous mothers. Similarly, about 36% of indigenous mothers did not have a single prenatal checkup during their last pregnancy compared to 12% of non-indigenous mothers. It is not surprising then to learn that the infant mortality rate is higher for indigenous mothers (10.5%) and is more than double the infant mortality rate for non-indigenous mothers (5.1%). For the children that are not victim to infant mortality, those born to indigenous mothers have higher rates of stunted height (59%) than those born to non-indigenous families (29%) (World Bank, 2004). The majority of the Indigenous population lives in rural villages or shanty towns outside the urban center, which geographically limits access to proper healthcare. General lack of resources also limits healthcare access since families would rather buy food than send their 10-year-old to the doctor for a mild but chronic cough. Poorer health quality, however, directly correlates to poorer education quality. If an illness is serious enough to prevent a child from attending school, he obviously is not learning. If the child is still able to attend school, mild symptoms can be distracting and affect concentration and memory, lowering the overall education quality.

When families lack basic resources, schooling, like healthcare, also becomes a secondary option. If families cannot generate enough income, it is common to send children to work instead. Even though schooling is mandated by law for children under the age of 16, one out of three children do not complete the first six years of primary education. Moreover, just over two out of every ten adults have completed high school (Sanchez, 2004).

**The Poverty Trap**

All of the factors named above (social networking, education quality, healthcare quality, and lack of basic resources) contribute to the cyclic phenomenon known as the Poverty Trap. All factors negatively impact education quality, which in turn affects the ability to obtain jobs that pay living wages. If unable to obtain living wages, families lack the
resources needed to save to come out of poverty. Moreover, they cannot to send their children to school. Instead, their children are needed to work as an additional source of income. Hence, populations become trap in a never-ending lifestyle of poverty.

The “Pre-Sorting Mechanism” is another form of the poverty trap that can arguably be even more powerful. Its origin can be traced to the Spanish Conquest when the Indigenous and Afro-Ecuadorians were typically not as qualified for formal sector jobs for the factors named earlier. As a result, employers have statistically discriminated, possibly to the degree of screening resumes by ethnicity of names. Today, these statistics are not the norm. But even if an Indigenous or Afro-Ecuadorian is able to receive a good education and become qualified for a specific job, he might be so influenced by discrimination that he anticipates the discrimination and either prevents himself from getting the job he deserves. It becomes “pre-sorting” when the applicant anticipates the discrimination and asks for wages below his qualifications or chooses to not apply altogether. Hence, the application process becomes a “self-fulfilling prophecy” (Morena, 2004). This can be worse than any other form of discrimination since the applicant has overcome other forms of discrimination in order to become qualified. But when it comes to the actual application process, the applicant self-discriminates, so that is it he himself that keeps him in the poverty trap.

Employment rates are also a factor in creating the pre-sorting mechanism. Currently, the Afro-Ecuadorian unemployment rate (3.9%) is higher than the national unemployment rate (2.7%). The Indigenous unemployment rate (1.1%) is curiously lower than the national rate. However, this can be explained by the disproportionate amount of Indigenous that are employed in low-paying and unskilled sectors while under 1% of the population are employed in upper management, administrative, or professional technician professions (see Figure 3, Appendix). This history adds to their lack of confidence in the workplace and application process. Similar pre-sorting reasoning can keep parents from sending their children to school or encouraging them to pursue more economically sustainable careers. Since parents pessimistically anticipate their children cannot compete for jobs in the formal sector, they do not see the benefit in pursuing a formal education.

The poverty trap continues to fuel the earnings gap between classes. According to one study, the “classism” gap is twice the size of the earnings gaps based on gender, three times the size of the earnings gaps based on beauty, and three times the size of the earnings gaps based on race (Nuñez and Gutierrez, 2004). The poverty trap becomes a serious dilemma for development by limiting large populations from becoming significantly active economic players in the market. With less consumption power, there is less economic activity. In short, by limiting the economic development of Indigenous and Afro-Ecuadorians, the economic development of Ecuador is effectively immobilized.
**Figure 1:** Minimum and Actual Ecuador Banana Prices January 2001 - July 2003

Source: SONICONTI S.A.

**Figure 2:** Results of 5th Graders on a Modified Version of Aprendo

<table>
<thead>
<tr>
<th>Language</th>
<th>Non-Indigenous</th>
<th>Indigenous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>7.58</td>
<td>6.21</td>
</tr>
<tr>
<td>Women</td>
<td>7.90</td>
<td>6.00</td>
</tr>
<tr>
<td>All</td>
<td>7.74</td>
<td>6.11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mathematics</th>
<th>Quecha Schools</th>
<th>Shuar Schools</th>
<th>Afro-Ecuadorian Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>6.35</td>
<td>6.07</td>
<td>5.50</td>
</tr>
<tr>
<td>Women</td>
<td>6.07</td>
<td>5.50</td>
<td>6.07</td>
</tr>
<tr>
<td>All</td>
<td>6.21</td>
<td>6.18</td>
<td>5.05</td>
</tr>
</tbody>
</table>


**Figure 3:** Indigenous Employment by Sector

<table>
<thead>
<tr>
<th>Occupational Sector</th>
<th>Percent of Population Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural and forest horticultural workers</td>
<td>36.6</td>
</tr>
<tr>
<td>Drivers, transportation workers, and similar</td>
<td>36.4</td>
</tr>
<tr>
<td>Miners, textile workers, and tobacco workers</td>
<td>12.7</td>
</tr>
<tr>
<td>Upper management, administrative personnel, professional technicians</td>
<td>&lt; 1.0</td>
</tr>
</tbody>
</table>


**References**


It Right, or Not at All.” Carnegie Endowment for International Peace, Washington D.C.


Sistema Integrado de Indicadores Sociales del Ecuador (SIISE) 3.5 (2001). Ministerios del Frente Social, INEC, INFA, Consejo Nacional de las Mujeres CONAMU.


World Bank (2005): “World Development Indicators.”


ECONOMIC DETERMINANTS OF DROPOUT RATES IN MEXICO: A CROSS-STATE ANALYSIS

Héber M. Delgado-Medrano
Georgetown University

ABSTRACT
Why do the wealthier northern Mexican federal entities perform worse in terms of dropout rates than their poorer counterparts? The existent theories on dropout rates that focus on the internal and external income-contingent determinants of dropout rates are unable to explain this anomaly. Using panel data and a random effects model, this study attempts to answer this question by analyzing the impact of macroeconomic labor-market factors such as international emigration, unemployment, maquiladora (foreign-owned factory) employment, and wages on secondary dropout rates. The results at the secondary level of schooling suggest international emigration, maquiladora factory employment, and higher wages have a positive and significant effect on secondary school dropout rates. Unemployment has a negative effect on dropout rates, except when wages were included. A supplementary analysis of primary school dropout rates suggests that whereas at the secondary level some macroeconomic variables exert a positive substitution effect on dropout rates, at the primary level these same exert a negative income effect on dropout rates. Ultimately, the results of this study support all of our hypotheses, suggesting that the addition of economic variables to the traditional theories on dropout rates results in a suitable explanatory model for dropout rates in Mexico.

I. Introduction
With the advent of theories on human capital and social capital attainment, education became a central theme in the literature of economic growth, often cited as a stronger determinant of increased worker productivity than physical capital gains (Denison 1974, 1985). Dropping out of school, therefore, came to be considered an alarming social problem with adverse effects for individuals and for society as a whole. Since the 1970s, many researchers have found that dropouts are subject to lower wages; higher employment
volatility; alcoholism; and a host of other economic and social problems (see for instance King 1978 and McCaul et al 1992). Given the unfavorable consequences of dropping out, much research has also been done with respect to the determinants thereof. Most of the research in this area focuses on two main categories of determinants that can be classified as either internal or external factors. The internal factors focus on school-related issues such as school quality, teacher preparedness, supplemental programs, academic performance and retention, and teaching styles (Barnes 1999; Ehrenberg and Brewer 1984). External factors take into account issues such as family composition and background, household income, socioeconomic status, race, gender, and involvement in alcohol, drugs, gangs, and crime (CEPAL 1993; Psacharopoulos and Arriagada 1989; Barnes 1999; Farkas 1996).

The vast majority of the cited determinants of dropout rates in both the internal and external categories are closely associated with wealth and a great number of them arguably are direct functions of income level. In a sense, therefore, the existent literature on the causal factors of desertion rates is highly centered on income as a determinant of dropping out. This focus on income is generally justified given the much higher performance of wealthy countries in terms of dropout rates vis-à-vis their poorer counterparts. However, zooming in at the regional level within Mexico, this pattern simply does not hold. At both the primary and secondary level, many of the wealthiest federal entities in Mexico actually exhibit the highest dropout rates, while the single best performer at both levels, the state of Tlaxcala, ranked 29th out of 32 federal entities in terms of per capita income. Moreover, there seems to be a general regional pattern: southern and poorer federal entities tend to outperform their richer northern counterparts at both levels of schooling. These findings are inconsistent with the literature on determinants of dropping out, which expects wealthier entities to perform better.

Using cross-state panel data for the years 1990 to 2002, this paper attempts to explain these anomalous findings by reexamining the validity of the specification of previous theoretical models and developing an alternative and better-specified model to interpret dropout rates in Mexico. Unlike other models that precede it, ours will take into account several macroeconomic variables that are generally neglected in the literature on dropout rates, but that are nonetheless important determinants of dropout rates in Mexico: international emigration rate, unemployment, maquiladora (factory) workers, and wages. Our hypothesis is that the combined effect of these macroeconomic variables explains the anomalous situation in Mexico. Including these variables should therefore give us a better specification than that provided by the literature on dropout rates.

II. Theoretical Foundations

The literature on the determinants of dropout rates altogether ignores international emigration as a factor. Most of the research on dropout rates focuses on high-school dropouts in the United States and because the United States experiences negative rates of international emigration, most U.S. studies do not seriously consider emigration as a factor. In the case of Mexico, however, given the high rates of international and national emi-
We hypothesize that this force is one of the largest determinants of dropping out:

H1a: secondary school dropout rates will be higher in federal entities where a high percentage of the population emigrates internationally, ceteris paribus

H1b: secondary school dropout rates in a given federal entity will not change significantly due to national migration, ceteris paribus

Hypothesis H1a hinges on the fact that the relevant measure of secondary dropout rates, SDROP, includes people that dropout of school to move to another location. Because most migrants within Mexico emigrate from Mexico to the United States, international emigration is therefore a potentially highly relevant determinant of SDROP. Most migrants to the United States from Mexico are males between the ages of 15 and 44 (MPI 2006; Durand et al 2001). Therefore, because secondary school takes place between the ages of 12 and 16 years, many secondary school-age individuals drop out explicitly to immigrate to the United States. Hypothesis H1b is founded on the assumption that secondary school-age individuals are much less likely to migrate internally within Mexico given the lower magnitude of national migration vis-à-vis international migration (CONAPO 2006). If internal emigration was in fact a significant determinant of secondary school dropping out, then southern federal entities would perform much worse than their northern counterparts since most internal migration within Mexico is south to north (CONAPO 2006).

In terms of the effect of unemployment on dropout rates, the literature does exist, but it is very narrow. In her seminal work, Duncan (1962) showed that the national unemployment rate was negatively related to school dropout rates for certain age cohorts. The rationale behind this finding is that students are reluctant to drop out of school if jobs are scarce. The results of Rumberger (1983) for the United States and Peraita and Pastor (2000) for Spain support Duncan’s original findings, suggesting a strong negative relationship between dropout rates and unemployment. Ehrenberg and Brewer (1994), however, actually found a positive relationship between unemployment and the dropout rate of White low-income students. Ultimately, they were unable to establish the same relationship between unemployment and the dropout rates of Whites of other socioeconomic levels, or of African Americans and Hispanics, generally. Given the consensus that there exists a negative relationship between unemployment and dropout rates, we hypothesize that unemployment is also a strong determinant of dropout rates in Mexico:

H2a: secondary school dropout rates will be higher in federal entities where there is a low level of unemployment (high availability of jobs), ceteris paribus

H2b: wages may potentially explain out some of the effect of unemployment on secondary school dropout rates

Hypothesis H2a rests on two assumptions. First, we assume that the main purpose of
going to secondary school is to become eligible for the pecuniary benefits that arise from employment. Secondly, if, however, a job is readily available to those who do not finish secondary school then the returns to schooling decrease while the opportunity cost of finishing secondary school become high, in terms of forgone earnings. Given the opportunity, therefore, secondary school students who do not place an inherent value on education will choose to drop out. Hypothesis H2b may be relevant since the presence of maquiladoras and the level of wages could at least partially determine the level of unemployment, although not entirely.

Because maquiladoras are an almost exclusively Mexican phenomenon, it is not surprising that the literature on dropout rates does not consider them. Nonetheless, given the importance of maquiladoras in Mexican labor markets, this analysis opts to analyze them as a determinant factor. The theoretical arguments for hypothesis H3, however, are identical to those of hypothesis H2a, since the presence of maquiladora factories substantially increases the probability of finding a job in the regions where they are present (Feenstra and Hanson 1995). In addition, given that many of these factories only require basic levels of education, secondary school students might be reasonably tempted to dropout of school to get a job in a maquiladora. Moreover, earnings at a maquiladora factory can be up to four times the average minimum wage in many of these federal entities (ILO 1999), therefore, the opportunity cost of going to school may be quite high, even if the returns to schooling are also high. Because maquiladora factories are not present in all federal entities, however, and because unemployment in other sectors may not directly be affected by maquiladora production (an export industry), we expect that the effect of maquiladoras on dropout rates is independent of that of unemployment levels generally. This leads to the following hypothesis:

H3: secondary school dropout rates will be higher in federal entities where the presence of maquiladora factories is large (more maquiladora workers), ceteris paribus (independently of unemployment in other sectors)

Finally, the only well-documented relationship between dropout rates and labor market conditions in the literature is that between (minimum) wages and dropout rates. In their various works Neumark (1995), Neumark and Wascher (1995), Card (1992), Cunningham (1981), Ehrenberg and Marcus (1982), and Gustman and Steinmeier (1981) all find evidence that increases in wages lead to dropping out and decreased levels of enrollment in high schools in the United States. The logic behind these findings hinges on the assumption that the higher are wages in a given region, the more will students be tempted to dropout, for this increases the opportunity costs of going to school in terms of forgone wages. We therefore hypothesize that:

H4: secondary school dropout rates will be higher in federal entities where salaries are higher, ceteris paribus
III. Data and Methodology

This analysis employs panel data that includes annual observations on dropout rates at the secondary school level from all thirty-two federal entities in Mexico (thirty-one states and a Federal District), for twelve years, from 1990 to 2001. Because the observations of cross-sectional units have different intercepts in this dataset, we used a random effects generalized least squares (GLS) model in order to avoid potential biases that may arise from employing an ordinary least squares (OLS) model. As the first step of this analysis, we regressed secondary dropout rates on the theoretical variables (traditional internal/external model) and then on the economic variables of interests in conjunction with the theoretical control variables (expanded model). We ran several regressions until we found suitable specifications and then compared the performance of the traditional model vis-à-vis our expanded macroeconomic model. Secondly, in order to enrich our findings, we similarly regressed dropout rates at the primary (elementary) school level using both the traditional and expanded models. We then compared this set of findings with those at the secondary level.

As the key dependent variable for all secondary school level regressions, we used SDROP, the percentage of students that dropout of secondary school; for primary school level regressions, we used PDROP, the dropout rate at the primary school level. For our key independent variables we used EMG1 (net rate of international emigration); UNPL (rate of unemployment for selected urban cities—used as a proxy for unemployment at the federal entity level); MAQ (number of workers per 100,000 people in a given federal entity that work at maquiladora factories); SLR, (a proxy of wages measured as the median wage—in tens of pesos per day—for which social security contributions are required); and EMG2 (net rate of internal emigration—within Mexico). The variable GDPC—GDP per capita (constant 1993 prices)—controlled for the effect of income and served as a proxy for all income-contingent factors. As an alternate measure, we used GDPV—annual rate of growth of GDP per capita. We controlled factors arising from the urban-rural divide cited in the literature by using URBP—the percentage of the population that lives in an urban setting. For the demand for education, we controlled using POPG—rate of growth of the population—as a proxy. To control for rates of return to schooling we used RTN—an estimate of the regional rate of return to schooling in Mexico. Additionally, in some regressions we used PDROP as an independent variable to control for all additional external factors that contribute to dropout rates at both levels of schooling. Finally, we used SFAIL (the rate of academic failure at the secondary level) and STDN (the average number of students per teacher at the secondary level) to control for internal school quality and performance factors. All data was taken from official national statistics, except RTN (Rodríguez-Oreggia 2005).

The baseline regression (excluding key economic variables) is the following:

\[
SDROP_{it} = \mu + \beta_1 GDPC_{it} + \beta_2 URBP_{it} + \beta_3 RTN_{it} + \beta_4 POPG_{it} + \beta_5 PDROP_{it} + \beta_6 SFAIL_{it} + \beta_7 STDN_{it} + (u_i + \varepsilon_{it})
\]
We used equation (1) as a starting point for finding a more significant and parsimonious specification, which gave us:

\[
\text{SDROP}_{it} = \mu + \beta_1 \text{GDPV}_{it} + \beta_2 \text{RTN}_{it} + \beta_3 \text{POPG}_{it} + \beta_4 \text{PDROP}_{it} + (u_i + \varepsilon_{it})
\]

To equation (4) we then added our economic variables of interest and modified it accordingly until we obtained our final specification:

\[
\text{SDROP}_{it} = \mu + \beta_1 \text{GDPV}_{it} + \beta_2 \text{RTN}_{it} + \beta_3 \text{POPG}_{it} + \beta_4 \text{PDROP}_{it} + \beta_5 \text{SFAIL}_{it} + \beta_6 \text{MAQ}_{it} + \beta_7 \text{SLR}_{it} + (u_i + \varepsilon_{it})
\]

**IV. Findings**

**Secondary School Analysis**

In regression (13), our final and best specification (TABLE 1), we can see that all included variables are significant determinants of dropout rates, and in particular, that the economic variables seem to be the driving force behind the anomalous findings. This suggests that we have arrived at a model that more accurately explains the situation in Mexico by taking into account the traditional factors cited in the literature as well as the heretofore-neglected macroeconomic variables introduced in this analysis.

The results on TABLE 1 consistently confirm the predicted positive effect of emigration on dropout rates. In the regression (13), we can observe that for every percentage point increase in the number of emigrants from Mexico to the United States as a percentage of total population, the secondary school dropout rate increases by 2.49 percentage points (range: 2.1-11.9%). To put it differently, one standard deviation increase in international emigration leads to an increase of 0.57 standard deviations in secondary dropout rates (99 percent confidence level). The magnitude of this effect places EMG1, consistently and significantly, as the most important determinant of dropout rates at the secondary leveling Mexico across all regressions. This confirms hypothesis H1a. It is therefore not surprising that the federal entities that exhibit the highest rates of net international emigration are also often the worst performers in terms of dropout rates—e.g. Aguascalientes, Chihuahua, Durango, Jalisco, Michoacán, and Zacatecas. EMG1 resulted significant at the 99 percent confidence level in all regressions.

Moreover, as predicted in our hypothesis H2a, all regressions where UNPL was included show that it has a negative effect on dropout rates, even when we considered MAQ as a variable. This suggests that independently of the effect of low levels of unemployment caused by the presence of maquiladora factories, low levels of unemployment, generally, also contribute to higher dropout rates. Nonetheless, when the variable SLR was included UNPL became not significant (as predicted in hypothesis H2b). We dropped UNPL in the final specification because the effect of MAQ and SLR on dropout rates overtook its effect. However, it is important to note that in all regressions in which UNPL is significant its effect on dropout rates was always among the highest, especially when compared to the
control variables that are cited in the literature as the main determinants of dropout rates. Highly commercial and industrialized federal entities with low levels of unemployment like Chihuahua, Baja California, Jalisco, and Quintana Roo are all among the worst performers in dropout rates also.

All regressions that included MAQ confirmed that it has a positive effect on dropout rates as was also predicted in our hypothesis. In particular, the results in equation (13) suggest that one standard deviation in the number of maquiladora workers per 100,000 inhabitants leads to an increase of 0.2 standard deviations in the dropout rates (95 confidence level). This finding supports the logic behind hypothesis H3. As far as the regional anomaly is concerned, MAQ is an important explanatory factor. Maquiladoras are heavily concentrated in the Mexican Northwest in states like Baja California, Sonora, and Chihuahua and in large southwestern states like Jalisco, all four of which fall among the worst performers. One reservation for including MAQ, however, is that it substantially reduces our number of observations. Nonetheless, the primary school analysis corroborates many of the findings in this section, despite the exclusion of MAQ.

The secondary analysis also confirms the predicted positive effect of wages (SLR) on SDROP. For every 10-peso increase in the SLR dropout rates increase by 0.113 percentage points; one standard deviation in wages leads to 0.25 standard deviations in SDROP (99 percent confidence level). SLR was significant at least at the 90 percent confidence level in all but one specification (where UNPL was also included). As in the case of the other hypotheses, the regression results confirm hypothesis H4.

Out of all the control variables, only RTN and PDROP were consistently significant across all regressions, while GDPC, GDPV, URBP, and STDN where either never or rarely significant. POPG was significant under most specifications; however, in the final specifications where we introduced SLR in conjunction with all other controls its significance disappeared. Conversely, SFAIL resulted not significant until the final specifications, perhaps owing to the addition of all the relevant macroeconomic variables. In the final specification, the most important controls are PDROP, RTN, and SFAIL, respectively. It is particularly interesting that GDPC and GDPV were not among the most important determinants of dropout rates in Mexico, since the theories on dropout rates emphasize income-contingent factors, both internal and external. Moreover, it is also interesting that where GDPV was significant, it had a positive effect on dropout rates. This suggests that the higher is the rate of growth of per capita GDP the higher the dropout rate will be. This makes sense if we taken into account that variation of per capita GDP in Mexico is correlated with availability of jobs, which we have already established as having a positive effect on dropout rates.
<table>
<thead>
<tr>
<th>Independent Variable</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>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
<th>(13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>4.295&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.996&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.102&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.843&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.594&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.819</td>
<td>1.356&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.943&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.584&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.530</td>
<td>-0.788</td>
<td>-1.291</td>
<td>3.123&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>GDPC</td>
<td>3.94&lt;sup&gt;-05&lt;/sup&gt; (5.14e-05)</td>
<td>0.031&lt;sup&gt;a&lt;/sup&gt; (0.007)</td>
<td>0.031&lt;sup&gt;a&lt;/sup&gt; (0.007)</td>
<td>0.031&lt;sup&gt;a&lt;/sup&gt; (0.007)</td>
<td>0.015&lt;sup&gt;c&lt;/sup&gt; (0.008)</td>
<td>0.014</td>
<td>0.015&lt;sup&gt;b&lt;/sup&gt; (0.008)</td>
<td>0.015 (0.008)</td>
<td>0.017</td>
<td>0.017</td>
<td>0.018&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.019&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.018&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>GDPV</td>
<td>-0.002 (0.009)</td>
<td>0.009 (0.013)</td>
<td>0.011 (0.013)</td>
<td>0.002 (0.007)</td>
<td>0.006 (0.011)</td>
<td>0.008</td>
<td>0.007 (0.011)</td>
<td>0.010 (0.010)</td>
<td>0.266&lt;sup&gt;a&lt;/sup&gt; (0.086)</td>
<td>0.368&lt;sup&gt;b&lt;/sup&gt; (0.148)</td>
<td>0.376&lt;sup&gt;b&lt;/sup&gt; (0.153)</td>
<td>0.394&lt;sup&gt;a&lt;/sup&gt; (0.138)</td>
<td>0.419&lt;sup&gt;a&lt;/sup&gt; (0.148)</td>
</tr>
<tr>
<td>URBPI</td>
<td>0.193&lt;sup&gt;a&lt;/sup&gt; (0.072)</td>
<td>0.146&lt;sup&gt;b&lt;/sup&gt; (0.063)</td>
<td>0.144&lt;sup&gt;b&lt;/sup&gt; (0.061)</td>
<td>0.132&lt;sup&gt;a&lt;/sup&gt; (0.051)</td>
<td>0.193&lt;sup&gt;a&lt;/sup&gt; (0.060)</td>
<td>0.219&lt;sup&gt;a&lt;/sup&gt; (0.070)</td>
<td>0.228&lt;sup&gt;a&lt;/sup&gt; (0.069)</td>
<td>0.213&lt;sup&gt;a&lt;/sup&gt; (0.060)</td>
<td>0.266&lt;sup&gt;a&lt;/sup&gt; (0.086)</td>
<td>0.368&lt;sup&gt;b&lt;/sup&gt; (0.148)</td>
<td>0.376&lt;sup&gt;b&lt;/sup&gt; (0.153)</td>
<td>0.394&lt;sup&gt;a&lt;/sup&gt; (0.138)</td>
<td>0.419&lt;sup&gt;a&lt;/sup&gt; (0.148)</td>
</tr>
<tr>
<td>RTN</td>
<td>-0.741&lt;sup&gt;a&lt;/sup&gt; (0.157)</td>
<td>-0.812&lt;sup&gt;a&lt;/sup&gt; (0.179)</td>
<td>-0.786&lt;sup&gt;a&lt;/sup&gt; (0.175)</td>
<td>-0.774&lt;sup&gt;a&lt;/sup&gt; (0.162)</td>
<td>-0.396&lt;sup&gt;a&lt;/sup&gt; (0.124)</td>
<td>-0.416&lt;sup&gt;a&lt;/sup&gt; (0.153)</td>
<td>-0.385&lt;sup&gt;b&lt;/sup&gt; (0.155)</td>
<td>-0.346&lt;sup&gt;b&lt;/sup&gt; (0.136)</td>
<td>-0.504&lt;sup&gt;a&lt;/sup&gt; (0.172)</td>
<td>-0.307&lt;sup&gt;b&lt;/sup&gt; (0.150)</td>
<td>-0.301&lt;sup&gt;c&lt;/sup&gt; (0.156)</td>
<td>-0.202 (0.190)</td>
<td>0.419&lt;sup&gt;a&lt;/sup&gt; (0.148)</td>
</tr>
<tr>
<td>POG</td>
<td>0.467&lt;sup&gt;a&lt;/sup&gt; (0.133)</td>
<td>0.487&lt;sup&gt;a&lt;/sup&gt; (0.142)</td>
<td>0.516&lt;sup&gt;a&lt;/sup&gt; (0.142)</td>
<td>0.523&lt;sup&gt;a&lt;/sup&gt; (0.136)</td>
<td>0.531&lt;sup&gt;a&lt;/sup&gt; (0.110)</td>
<td>0.563&lt;sup&gt;a&lt;/sup&gt; (0.119)</td>
<td>0.592&lt;sup&gt;a&lt;/sup&gt; (0.110)</td>
<td>0.602&lt;sup&gt;a&lt;/sup&gt; (0.099)</td>
<td>0.499&lt;sup&gt;a&lt;/sup&gt; (0.131)</td>
<td>0.519&lt;sup&gt;a&lt;/sup&gt; (0.131)</td>
<td>0.518&lt;sup&gt;a&lt;/sup&gt; (0.130)</td>
<td>0.481&lt;sup&gt;a&lt;/sup&gt; (0.145)</td>
<td>0.485&lt;sup&gt;a&lt;/sup&gt; (0.129)</td>
</tr>
<tr>
<td>PDROP</td>
<td>0.014 (0.023)</td>
<td>0.017 (0.026)</td>
<td>-0.047 (0.033)</td>
<td>-0.055 (0.043)</td>
<td>0.013 (0.018)</td>
<td>0.023 (0.020)</td>
<td>0.013 (0.018)</td>
<td>0.023 (0.020)</td>
<td>0.049&lt;sup&gt;b&lt;/sup&gt; (0.023)</td>
<td>0.049&lt;sup&gt;b&lt;/sup&gt; (0.021)</td>
<td>-0.033 (0.046)</td>
<td>0.049&lt;sup&gt;b&lt;/sup&gt; (0.021)</td>
<td></td>
</tr>
<tr>
<td>SFAIL</td>
<td>1.622&lt;sup&gt;a&lt;/sup&gt; (0.292)</td>
<td>1.544&lt;sup&gt;a&lt;/sup&gt; (0.301)</td>
<td>1.556&lt;sup&gt;a&lt;/sup&gt; (0.285)</td>
<td>1.476&lt;sup&gt;c&lt;/sup&gt; (0.315)</td>
<td>1.553&lt;sup&gt;a&lt;/sup&gt; (0.390)</td>
<td>1.865&lt;sup&gt;a&lt;/sup&gt; (0.362)</td>
<td>1.857&lt;sup&gt;a&lt;/sup&gt; (0.369)</td>
<td>2.172&lt;sup&gt;c&lt;/sup&gt; (0.469)</td>
<td>2.489&lt;sup&gt;a&lt;/sup&gt; (0.371)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STDN</td>
<td>-0.141&lt;sup&gt;a&lt;/sup&gt; (0.032)</td>
<td>-0.148&lt;sup&gt;a&lt;/sup&gt; (0.037)</td>
<td>-0.150&lt;sup&gt;a&lt;/sup&gt; (0.037)</td>
<td>-0.151&lt;sup&gt;a&lt;/sup&gt; (0.036)</td>
<td>-0.083&lt;sup&gt;b&lt;/sup&gt; (0.039)</td>
<td>-0.030 (0.051)</td>
<td>-0.083 (0.039)</td>
<td>-0.030 (0.051)</td>
<td>-0.083&lt;sup&gt;b&lt;/sup&gt; (0.039)</td>
<td>-0.030 (0.051)</td>
<td>-0.083&lt;sup&gt;b&lt;/sup&gt; (0.039)</td>
<td>-0.030 (0.051)</td>
<td></td>
</tr>
<tr>
<td>EMGI</td>
<td>1.819e-04&lt;sup&gt;a&lt;/sup&gt; (5.7e-05)</td>
<td>1.561e-04&lt;sup&gt;a&lt;/sup&gt; (5.39e-05)</td>
<td>1.591e-04&lt;sup&gt;a&lt;/sup&gt; (5.46e-05)</td>
<td>1.521e-04&lt;sup&gt;a&lt;/sup&gt; (5.74e-05)</td>
<td>1.247e-04&lt;sup&gt;a&lt;/sup&gt; (5.29e-05)</td>
<td>0.062 (0.049)</td>
<td>0.072&lt;sup&gt;c&lt;/sup&gt; (0.041)</td>
<td>0.095&lt;sup&gt;b&lt;/sup&gt; (0.041)</td>
<td>0.113&lt;sup&gt;a&lt;/sup&gt; (0.038)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNPL</td>
<td>0.3044</td>
<td>0.3311</td>
<td>0.3415</td>
<td>0.3576</td>
<td>0.6159</td>
<td>0.6252</td>
<td>0.6341</td>
<td>0.6401</td>
<td>0.6619</td>
<td>0.6760</td>
<td>0.6671</td>
<td>0.6916</td>
<td>0.6768</td>
</tr>
<tr>
<td>Overall R²</td>
<td>0.3044</td>
<td>0.3311</td>
<td>0.3415</td>
<td>0.3576</td>
<td>0.6159</td>
<td>0.6252</td>
<td>0.6341</td>
<td>0.6401</td>
<td>0.6619</td>
<td>0.6760</td>
<td>0.6671</td>
<td>0.6916</td>
<td>0.6768</td>
</tr>
<tr>
<td>N</td>
<td>288</td>
<td>256</td>
<td>256</td>
<td>256</td>
<td>278</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>182</td>
<td>182</td>
<td>184</td>
<td>184</td>
<td>184</td>
</tr>
</tbody>
</table>
Overall, all of the regressions support the hypotheses (H1 – H4) of this paper. The specifications that included only the control variables are simply not able to explain the dropout situation in Mexico, whereas the regressions including the macroeconomic variables present a much clearer picture of the situation. It is also particularly interesting that EMG1 is actually the most important determinant factor of dropping out at the secondary level in Mexico.

**Primary School Analysis and Comparisons**

Overall, as in the case of secondary level dropout rates, the economic variables are essential in explaining dropout rates at the primary level. Moreover, once again, these key variables are among the most significant explanatory determinants of PDROP, although less so than in the case of secondary schools. EMG1 and SLR, for instance, both exhibit among the largest coefficient/std. error (C/SE) ratios and are significant at the 99% confidence level across the board. UNPL also performs well in explaining dropout rates at the primary level, although its significance levels are relatively low. MAQ was not significant in these regressions, although it was very close to significance. Of the control variables, only SDROP consistently exhibits high levels of significance and large C/SE ratios. Income, measured in terms of GDPV and GDPC, is simply not an important factor.

Among the most interesting findings in the primary school analysis is the reversal in the direction of the effects of many of these variables, including EMG1, SLR, and UNPL. Unlike in the case of secondary level analysis, the effect of EMG1 on PDROP is actually negative. This finding suggests that there exists both income and substitution effects from EMG1 on dropout rates, depending on the level of schooling. In the case of secondary school, because students themselves migrate, the positive effect of EMG1 on dropout rates is a manifestation of a substitution effect (school substituted by work abroad). In the case of primary school, because children are not a large percentage of the emigration population and since they do not migrate for work, emigration does not directly affect dropout rates. However, emigration does have an indirect negative effect on children through income. Children (and communities, generally) whose parents (denizens) emigrate to work in the United States send back remittances and these increase the income of children (communities) in Mexico (Woodruff and Zenteno 2001). If previous theories that suggest income is a large determinant of dropping out are in fact even partially valid, then this would explain the large negative effect of EMG1 on PDROP.

The negative effect of SLR on dropout rates also seems to support the existence of both an income and a substitution effect. In the case of PDROP, because children are not regular participants in the labor market, the only effect SLR can have on their dropping out of school or not is through its effect on income. In this analysis, we can see that, in fact, SLR does have a strong and significant negative effect on dropout rates at the primary level. The effect of unemployment on PDROP, also suggests that at the primary level there does exist an income effect on dropout rates. Based on the overall impact of EMG1, SLR, and UNPL this income effect is of a smaller magnitude than the substitution effect experienced in secondary schools. In relative terms, the economic variables are stronger determinants of SDROP than PDROP.
V. Conclusions

Using panel data and a random effects model, this paper analyzed the effect of macroeconomic labor-market variables on dropout rates in Mexico while controlling for the traditional determinants cited in the literature on dropout rates. Unlike other papers, this study uses panel data from Mexico and focuses exclusively at the macro level of analysis in order to assess the effects of these broader variables at the regional level. The results of the secondary school regression analysis support all of our hypotheses, namely, that the effect of international emigration, maquiladora presence, and wages (and unemployment in special cases) on secondary school dropout rates is large and highly statistically significant. We can therefore conclude that these factors are the main reason why the poorer southern federal entities outperform their wealthier northern counterparts, initially a confounding result. Our analysis also employed regression analysis at the primary school level, which supported our initial findings and demonstrated the existence of income and substitution effects from the macroeconomic variables on dropout rates at the primary and secondary school level, respectively.

Some limitations of this analysis are the use of unbalanced data and several sub-optimal proxies, unavailability of more qualitative controls, unavailability of micro-level data at the subnational level, and potential for endogeneity and omitted variable bias in some cases. Future analyses that look at dropout rates starting in the year 2000 should have a much more complete set of variables at the level of federal entity, including micro-level data. It would also be interesting to carry a similar study using cross-country data to see whether these macroeconomic labor-market variables affect dropout rates everywhere or only in specific contexts of development.

References

Barnes, Donna, “Causes of Dropping Out from the Perspective of Education Theory,” in Laura Randall and Joan B. Anderson (Eds.), Schooling For Success: Preventing Repetition and Dropout in Latin American Primary Schools, pp. 14-22, 1999


BIE, Bank of Economic Information on Mexico, Base de datos, 2006 http://dgcnesyp.inegi.gob.mx/?c=72

Bracho, Teresa, “Basic Education in Mexico: An Overview,” in Laura Randall and Joan B. Anderson (Eds.), Schooling For Success: Preventing Repetition and Dropout in Latin American Primary Schools, pp. 103-118, 1999


CEPAL, Economic Commission for Latin America and the Caribbean, Panorama social de América Latina, (Santiago de Chile: UN CEPAL), 1993


King, R.H., The Labor Market Consequences of Dropping Out of High School, Center for Human Resource Research, Ohio State University, 1978


MPI, Migration Policy Institute, Migration Information Source, 2006 http://www.migrationinformation.org/


Psacharopoulos, George and Ana Maria Arriagada, “The Determinants of Early Age Human Capital

Rodríguez-Oreggia, Eduardo, “Institutions, Geography, and the Development of Regional Returns to Schooling in Mexico,” (Ciudad de México: Instituto de Investigaciones sobre Desarrollo Sustentable y Equidad Social, Universidad Iberoamericana), 2005


Sautu, Ruth, “Poverty, Psychology, and Dropouts,” in Laura Randall and Joan B. Anderson (Eds.), Schooling For Success: Preventing Repetition and Dropout in Latin American Primary Schools, pp. 23-32, 1999

STPS, Secretariat of Labor and Social Welfare of Mexico, Estadísticas laborales 2006 http://www.stps.gob.mx/


Bankruptcy and Industry Performance: Is Government Protection Hurting Returns in the Airline Industry?

Vikram Shankar
Georgetown University

How does the financial performance of an industry change when one of the competitors in the market files for bankruptcy protection? Standard competition models tell us that in a world with no protection for distressed firms, market shares, revenues, and profits would increase for a failing firm’s competitors, as they capitalize on the exit of a rival. These models do not take into account a firm’s financial condition or the implications of government protection. Hiding from creditors under the umbrella of the government, via Chapter 11 bankruptcy protection, would pose the question of how the financial performance of surviving, stable firms is affected. Notions of large-scale bankruptcies and lagging financial performance are deeply tied to the airline industry in the United States, which has lost over $40 billion and seen four out of its six major legacy carriers file for bankruptcy since 2002. This paper seeks to evaluate the effects that bankruptcy protection has on firms’ financial performance, as measured by stock returns.

This question is of importance given the trends in the industry over the past few decades. The domestic airline industry has evolved considerably since being deregulated in the late 1970’s. Benefits to consumers have been evident, as fares in real terms have plummeted since government intervention ended. Firms, on the other hand, have not experienced such prosperity, as free entry into the market has increased the number of competitors, forcing fares down and cutting profit levels. This has especially been the case in recent years, in which recession, rising fuel costs, terrorism concerns, and continued fierce competition have put many of the big names in the business into distress, and ultimately bankruptcy. Major bankruptcies coupled with large industry losses have caused many to question the merits of “egregious” bankruptcy statutes, which are criticized for protecting unprofitable, inefficient capacity and for not penalizing firm failure. These critics argue that in a fully deregulated industry, high-cost, inefficient firms should not be bailed out, and that the benefits to consumers from lower fares do not offset the losses to firms from lagging financial performance. This paper seeks to assess the validity of these claims.

Airline bankruptcies are nothing new, but the current phenomenon of major carriers falling into Chapter 11 is relatively novel. Most bankruptcies before 2002 were by small
carriers who were likely to liquidate their assets and exit the industry. In this new era of the industry, large bankrupt carriers are highly unlikely to liquidate, as the government provides them protection and allows them to restructure their debt obligations to creditors. Past literature has discussed the effects of bankruptcy throughout industries and also within the airline industry, but has not shed any light on these recent events. Thus, this study aims to not only test the effects of bankruptcy protection on financial performance, but to update previous findings to include recent major events.

The research presented seeks to provide evidence on one of two competing theories identified by Lang and Stulz (1992): the contagion effect hypothesis and the competition effect hypothesis. In brief, the former states that negative news on an individual firm is harmful to the entire industry, as it lures investors away from similar firms. This impact is seen in financial markets by decreases in returns during times of distress and increases in returns during times of resurgence. The competition effect states what classic economic models would: that a firm’s failure should shift demand over to other firms, thereby increasing revenues and profits for survivors. This impact is seen in markets by increases in returns during times of distress and decreases in returns during times of resurgence. Past literature has found instances of both being true, and this study provides another attempt at distinguishing them.

Using event study methodology as well as a dummy variable approach, the following analysis establishes certain event windows surrounding Chapter 11 filings and emergences and seeks to evaluate changes in returns. Rather than use individual airline stocks, which come and go quickly over time with entry and exit both into the industry and into bankruptcy, the study employs the use of several airline indices, each holding different universes of airline stocks. This evaluation of bankruptcy and emergence effects on stocks of different capitalizations further distinguishes this study from other literature.

The research suggests a small, although statistically insignificant, confirmation of the contagion effect hypothesis in all but one of the airline indices. The small-cap, large-cap, and overall composite indices all exhibit patterns consistent with the contagion effect, as they show negative abnormal returns (as low as −4.5%) during bankruptcy windows and positive abnormal returns (as high as 5.8%) during emergence windows. Mid-cap stocks exhibit the opposite behavior, consistent with the competitive effect, as they show positive abnormal returns (1.4%) during bankruptcy windows and negative abnormal returns (−1.4%) during emergence windows. These effects are found to be stronger when using a smaller window around the event date and replacing actual filing dates with major news stories that signal either bankruptcy or emergence. The effects are also found to be stronger in some individual cases of bankruptcy or emergence, particularly earlier instances, than in others. These results suggest that bankruptcy protection may be partially at fault for the industry’s losses but is not a major factor, as its effects on returns are largely insignificant.
DIVORCE NORMS, INTRAHOUSEHOLD BARGAINING, AND HOUSEHOLD OUTCOMES IN RURAL ETHIOPIA

Sarah Carroll
Stanford University

I. Introduction
In recent years, gender equality has gained significant attention from international development organizations (e.g. World Bank 2001). Recent economic research has suggested that increasing women’s rights could also help to accelerate development through several possible channels, including effects on fertility and corruption (see Kevane 2004 for an overview). Weak property rights present one of the most significant disadvantages for women across much of the developing world. The tenuous nature of these rights leaves women with little autonomy outside of marriage and, according to some evidence, little decision-making power within marriage as well.

In Ethiopia, one of the poorest countries in the world, statutory law guarantees husbands and wives an equal claim to family property at divorce. However, due to the influence of customary laws, women are often left with little or nothing in practice. These laws vary by region, ethnicity, and religion and give rise to very different levels of gender equality in property rights from place to place, facilitating a quantitative analysis of the effect of gender-biased property rights within households.

This paper uses a standard household bargaining framework to connect the tenuousness of rural Ethiopian women’s property rights with economic outcomes, including household budgets’ shares for food, education, medical care, children’s clothing, and cigarettes and alcohol; as well as individual children’s schooling outcomes. The analysis to follow makes use of extensive data on hypothetical property division at divorce that is included in the 1997 round of the Ethiopia Rural Household Survey (ERHS). I predict that divorce settlements that are more favorable towards women strengthen their bargaining positions within marriage and allow them to exert more influence in household decision-making. If the preferences of wives in this survey are similar to those analyzed in past literature, this increased influence should result in greater resource allocation towards children’s well-being and human capital accumulation.

II. Literature Review
For decades, conventional theory presumed that a household made its economic deci-
tions as a unit, maximizing a single utility function. In the early 1980s, economists pro-
posed an alternative idea: households consist of individual actors that bargain with each
other to maximize their own, individual utilities (Manser and Brown 1980, McElroy and
Horney 1981). Substantial empirical evidence supports this theory and suggests that
“income controlled by the husband and wife [has] significant and often substantially dif-
ferent effects on family behavior, whether measured by expenditure on categories of goods
and services or by outcomes such as child health” (Lundberg and Pollak 1996, p.144). If
this is true, policymakers might be able to affect economic outcomes by enacting policies
to change the balance of control over economic resources and other factors that affect
spouses’ bargaining power.

Many studies using data from developing countries have produced evidence that when
women’s bargaining power is stronger, households devote more resources towards growth-
enhancing expenditures, such as children’s education and nutrition (e.g. Hoddinott and
Haddad 1995, Quisumbing and Maluccio 2003, Doss 2005). Since many economic devel-
opment models rely on investment in education and other forms of human capital as a pri-
mary engine of growth, increasing women’s bargaining power might help to accelerate
economic development.

III. Data Description

The Ethiopia Rural Household Survey (ERHS) is a set of four-round panel data that,
in the 1997 round, spanned about 1,500 households consisting of nearly 10,500 individu-
als. Data exists for four waves (1989, 1994, 1995, and 1997), and by 1997, extensive sur-
veys were conducted in fifteen villages across the country. The villages represent
Ethiopia’s broad range of religious and ethnic groups and are chosen specifically to pro-
vide a cross-section of the country’s various farming systems; village selection is not
entirely random but the data provides a fairly comprehensive view of the country’s rural
agricultural and social systems.

<table>
<thead>
<tr>
<th></th>
<th>Young Children</th>
<th>Older Children</th>
<th>Land</th>
<th>House</th>
<th>Livestock acquired after marriage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Husband</td>
<td>18.29</td>
<td>41.38</td>
<td>51.38</td>
<td>55.86</td>
<td>22.53</td>
</tr>
<tr>
<td>Wife</td>
<td>57.58</td>
<td>6.90</td>
<td>3.28</td>
<td>4.49</td>
<td>2.85</td>
</tr>
<tr>
<td>Other*</td>
<td>13.27</td>
<td>36.09</td>
<td>43.81</td>
<td>38.85</td>
<td>66.21</td>
</tr>
<tr>
<td>N/A</td>
<td>10.86</td>
<td>15.63</td>
<td>1.53</td>
<td>0.80</td>
<td>8.41</td>
</tr>
<tr>
<td>Total</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

* “Other” refers mainly to outcomes in which the spouses split custody of children or divide the asset in question.

The survey’s inclusion of questions about property division at divorce makes it well
suited to a household bargaining analysis that uses divorce as the threat point. One key
characteristic of the ERHS questions is that they are purely hypothetical, which means that
spouses who have never been divorced can still respond. Specifically, survey-takers ask
household heads, “What would happen to custody of children and household assets in case, God forbid, you and your spouse mutually agree to divorce?” Possible responses for each type of asset include full ownership rights for either spouse, as well as a variety of intermediate arrangements in which the couple splits the asset: either half and half, according to custody of children, or by another system.

This data also highlights local variation in customary laws. Given how widely property division varies across the entire dataset, there is a relatively small amount of variation within each village, suggesting that local customary laws are enforced more rigidly than the more egalitarian, nationwide statutory laws.

IV. Methodology

In this paper, I develop a model addressing the potential effect of property rights inequality on spouses’ bargaining power and household decisions. Similar papers use pre-marriage assets as the primary indicator of bargaining power, but due to biased customary laws, most Ethiopian women might find themselves at a similar disadvantage when divorcing, regardless of how much wealth they brought to the marriage. If customary divorce settlements are independent of pre-marriage wealth, then the value of spouses’ pre-marriage assets will not hold much relevance as an explanatory variable.

I construct variables representing the expected distribution of a household’s land, home, livestock, and older children at divorce and use them as potential indicators of bargaining power. I choose the first three assets because they comprise the majority of most families’ wealth and are relatively durable. The effect of custody over children is more complicated—although parents might want to retain custody due to love for their children or due to their productive capacity, their care represents a significant strain on household resources. However, grown sons, in particular, provide a vital source of economic security for widowed mothers in many developing countries (Cain 1984). For this reason, I include older children in my model. In order to represent the spouses’ expected utility at divorce, I interact variables representing control over each asset at divorce with the asset’s value (or the total number of children). The variable representing asset control takes a value of one if the woman would claim all of the asset’s value at divorce, zero if the husband would claim the entire value, and one half if they would split it.

My analysis uses two types of dependent variables: household budget shares and individual educational outcomes for children. The budget share variables represent respondents’ estimates of household spending over the past four months on food, cigarettes, medical care, children’s clothing, and schooling as a proportion of total household expenditure during that time. Individual children’s educational outcomes are represented by the difference between the number of years of schooling they have completed and the cohort mean by age; i.e. a positive value means that the child has completed more schooling than his or her peers, on average.

I regress each dependent variable in turn on the four property rights indicators and a number of control variables. In my simplest budget share model, the controls include the natural log of household per-capita expenditure and household size, as well as a series of
variables representing the proportion of family members in given gender and age groups to control for differences in spending patterns based on household composition. I estimate a similar function that includes dummy variables for the reported religion and ethnicity of the household head. Due to the close link between customary divorce laws and religion, I expect controlling for ethnicity and religion to dilute the property rights effect and decrease the property rights variables’ statistical significance within the regression. The educational outcome model is similar, but includes a dummy variable for girls that I interact with the property rights indicators to isolate the specific effect of women’s bargaining power on girls’ educational outcomes.

V. Econometric Results

My results suggest that bargaining does occur within the households depicted in the ERHS data; however, some of the outcomes are different than expected. In the budget share regression excluding religion and ethnicity controls, the estimated coefficients on the housing rights and child custody variables are highly significant for the food, cigarettes and alcohol, and medical care budget shares. For example, a household with gender-equalized housing rights should spend 3.2 percentage points more of its budget on food than an otherwise identical household with male-biased rights. Likewise, while equalized housing rights decreased the budget share for cigarettes and alcohol by only 0.3 percentage points, the mean household budget share is only 0.42 percent. This effect is statistically significant and, in context, seems economically significant as well. Coefficients are less statistically significant when I include religion and ethnicity dummies, but the results are similar overall.

| Table 2: Budget Share Regression Results (no controls for religion and ethnicity) |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                | (1) Food        | (2) Cigarettes/Alcohol | (3) Medical Care | (4) Children’s Clothing | (5) Education |
| Land                           | -0.000          | -0.000           | 0.000           | -0.000         | -0.000         |
|                                | (0.26)          | (0.04)           | (0.45)          | (0.74)         | (0.65)         |
| House                          | 0.064           | -0.006           | -0.011          | -0.001         | -0.002         |
|                                | (3.42)***       | (3.50)***        | (2.87)***       | (2.87)***      | (1.22)         |
| Livestock                      | 0.000           | -0.000           | 0.000           | 0.000          | -0.000         |
|                                | (0.18)          | (1.64)           | (0.93)          | (0.95)         |                |
| Older Children                 | 0.011           | -0.001           | -0.002          | -0.001         | 0.000          |
|                                | (1.97)**        | (2.01)**         | (2.09)**        | (0.83)         | (0.53)         |
| Observations                   | 1018            | 1018             | 1018            | 1018           | 1018           |
| R-squared                      | 0.08            | 0.06             | 0.07            | 0.04           | 0.05           |

Absolute value of t-statistics in parentheses

*** significant at 1%, ** significant at 5%, * significant at 10%

The econometric results’ suggestion that housing and child custody rights are more important than land and livestock rights is interesting. Although land and livestock are
major sources of income and subsistence for most households, the need for housing might be more immediate—if a woman loses her claim on land, she will certainly suffer in the long run, but if she loses her home, she will suffer immediately. Likewise, perhaps older children contribute so much to economic security that they are more important to bargaining than the other assets. The specific importance of these two assets is a result that I hope to explore further in future research.

The negative coefficient for medical care is also surprising. If mothers care more about their children’s nutrition, it would make sense for them also to care more about their health. However, the negative effect on medical expenditures might just reflect women’s preferences in other areas, since the ERHS data does not specify whether the expenditure is for curative or preventative care. If this is true, perhaps better-nourished children simply require less medical care.

<table>
<thead>
<tr>
<th>TABLE 3: INDIVIDUAL SCHOOLING REGRESSION RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Land</td>
</tr>
<tr>
<td>House</td>
</tr>
<tr>
<td>Livestock</td>
</tr>
<tr>
<td>Older Children</td>
</tr>
<tr>
<td>Female*Land</td>
</tr>
<tr>
<td>Female*House</td>
</tr>
<tr>
<td>Female*Livestock</td>
</tr>
<tr>
<td>Female*Older Children</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>R-squared</td>
</tr>
</tbody>
</table>

Here, the gender gap in education is highly apparent. Without controlling for religion and ethnicity, the average girl obtains 0.337 fewer years of schooling than the average boy as a result of her gender alone, a very significant difference given an overall mean schooling level at age 15 of only 1.33 years. Increased housing rights for women have a positive effect on boys’ educational status, while increased land and child custody rights have a highly significant negative effect. Mothers’ housing rights have a strongly positive and significant effect on their daughters’ schooling—gender-equalized housing rights increase the average girl’s accumulated schooling by 0.44 years, although this effect is partially counteracted by the negative effect of child custody rights.

This might be linked with the old-age security hypothesis. Perhaps a woman who
enjoys secure housing rights does not need to worry too much about her future welfare; even independent of her husband and son, she will still have a place to live. If this is the case, then she might be freer to educate her daughter. However, independent of this effect, stronger child custody rights should increase the relevance of the old-age security hypothesis—a woman who knows that she will maintain a close relationship with her son even if she gets divorced might value him more as a future source of support.

IV. Conclusion
This paper’s results suggest that enforceable housing rights for women could help eliminate gender inequality in education. An undesirable trade-off does occur between boys’ and girls’ schooling. However, if combined with efforts to keep boys’ schooling relatively high, strengthening women’s housing rights could greatly accelerate the movement towards universal primary education and gender-equalized education. Likewise, programs or laws designed to increase women’s economic and social independence in old age would allow them to stop relying so heavily on their sons for future assistance, perhaps lessening the effects of child custody rights.

Many remaining questions demand further analysis. Most importantly, this analysis does not say anything about differences in the non-material consequences of divorce or about its social acceptability. Ethnographic evidence suggests that attitudes towards divorce vary widely, and in some cases, divorced women are shunned. This type of social and emotional penalty for divorced women worsens their bargaining threat point, and a more complete model would capture this effect.

The novelty of this analysis lies in its direct policy implications—the model’s design allows us to examine the effects of biased divorce settlements and easily tease out the specific effects of biased customary laws. Although enforcing egalitarian statutory laws would undoubtedly be difficult, the positive effects of enforcement on economic development could make it worthwhile.

References


ECONOMIC DEVELOPMENT AND THE GENDER WAGE GAP

Sherri Haas
Illinois Wesleyan University

I. Introduction

General wage inequality within countries is a topic that has received a great deal of attention in the economic literature. Differences in wages between men and women, particularly cross-national variation in the size of the gender gap, have not been studied as thoroughly. It was not until the latter half of the 20th century that the first full study of the role of women in economic growth and the effect of development on them was completed by Ester Boserup (1970).

Income parity between genders does not exist as of yet within any country, but the size of the wage gap varies considerably. Within this study, the greatest equity is found in Switzerland, which has a male-to-female wage ratio of 1.11, meaning that men make approximately 1.11 times as much as women. The greatest inequality is seen in Egypt, with a ratio of 3.84, implying that men make almost quadruple the wages of women. The difference in income between men and women is of great importance because it affects a very large number of people. The gender wage gap is not uniform cross-nationally, and if determinants of the gap size can be found, policy could be implemented to reduce the income disparities.

Economic inequality between men and women is an important problem deserving of in-depth study because of the large number of people it affects. Not only do current gender wage gaps affect women, they also likely affect their children and future generations. A number of studies have found evidence that women tend to spend income differently than men, with greater shares being put back into the household and more equally allocated between children of both sexes (Blau, Ferber, and Winkler, 2001). This in turn improves equality between genders if both girls and boys receive similar benefits and education.

This study will examine the relationship between the gender wage gap and the degree of economic development of a country as measured by the gross domestic product per capita. A second model uses the United Nations Human Development Index as a more comprehensive measure of development. It also analyzes the relationship of educational attainment and general wage inequality to the size of the gender wage gap.
II. Theory and Literature Review

This research is based upon the basic theory of supply and demand and on human capital theory with reference to the inverted-U curve postulated by Simon Kuznets.

Kuznets’ Inverted-U Hypothesis states that there is a relationship between the per capita income of a country and the amount of income inequality within it. When plotted it creates an inverted-U curve (Figure 1) such that as per capita income increases from subsistence a country will experience first increasing inequality and then decreasing inequality at later stages of development. Kuznets suggested that the pattern was the result of the inflow of people into urban areas with unequal income distributions as a country begins to develop, which would cause inequality to increase at first, but that inequality then decreases as the new members integrate into the labor force (Lantican et al, 1996).

![Figure 1: Kuznets’ Inverted-U](image)

This study suggests that an inverted-U pattern is also to be expected when measuring income inequality between genders within countries. As a country begins to develop the growing industrial sector increases the demand for male labor, while not providing greater opportunities for female workers, resulting in a larger wage gap between men and women. Boserup (1970) concluded that women are often active participants in “home industries” which produce hand made items for sale. As these industries gradually decline in importance in an economy and lose business to large scale manufacturing – which hire a larger share of men – women’s jobs are lost. In later stages of development, as the economy continues to grow and provide more service sector jobs, the labor force opportunities for women increase, thus decreasing the gender earnings gap.

Pampel and Tanaka found support for Boserup’s conclusion about women’s participation in the work force in their study of the effects of economic development on female labor force participation rates. Their cross-national study of 70 countries showed a curvilinear relationship, with female labor force participation rates first decreasing with regards to the measure of economic development, and increasing at greater levels of development. This would also suggest that female earnings first decrease relative to those of men and increase at later stages of development.

Human capital theory suggests that people are compensated in the work place based
on their abilities and skills (Borjas, 2000). Common measures used to get at the difficult to quantify concept of human capital include years of schooling, on-the-job training, and years in the labor force. Human capital theory would suggest that the greater the difference in educational attainment between men and women, the greater the wage difference will be.

Blau and Kahn found in their study of ten industrialized nations that the wage structure of a country is an important factor in determining the size of the gender wage gap. In the countries they examined, overall income inequality of a country accounted for a large portion of the gender wage gap. There is no reason to suppose that this factor would only be important in the industrialized countries, and so should also be included in the study.

Kidd and Shannon expanded on Blau and Kahn’s work, using the same methodology to compare the gender wage gaps of Australia and Canada. They also found that wage structure was significant, and point out that this is important because it “highlights the fact that the gender wage differential may differ between countries for reasons not specifically tied to gender” (1996). Evidence of the importance of the wage structure has also been found outside of the industrialized West. Brainerd’s study of seven countries of the former Soviet Union showed that the widening of the wage gap negatively affected women in all of them, although in five of the seven the losses were “more than offset by gains in rewards to observed skills and by an apparent decline in discrimination against women” (2000).

In sum, the literature suggests that an inverted-U pattern can be expected when plotting the gender wage gap against a measure of economic development. It is also expected that differences in human capital attainment between men and women will help explain the wage gap. Recent studies also propose that greater general wage inequality within a country will be correlated with higher gender wage inequality.

### III. Data

Cross-sectional data for the dependent variable, a ratio of male-to-female earnings, is available from the United Nations Human Development Report 2005: Human Development Indicators. The estimated earned income for men and women in US purchasing power parity dollars is provided for 154 of 177 countries based on the most recent year for which data are available between 1991 and 2003. While it is less than ideal to use data from different years, it is the best available, and it is also unlikely that the values for each country would have changed extremely in the time period.

GDP per capita is the variable of choice for previous literature on the Kuznets hypothesis, and the first model will use it as the measure of development. The data for this independent variable, GDP per capita, is also available in the Human Development Report 2005. The majority of the values are for the year 2003, however, for countries for which recent data were not available, the HDR calculated values based on economic regression.

It has been shown that GDP does not account for all aspects of development (Todaro and Smith, 2006). The United Nations Development Program created the Human Development Index as an alternative and more comprehensive measure of socioeconomic
development. Human Development Index values are used as the measure of development in the second model and also come from HDR 2005. The HDI is calculated using a combination of longevity, knowledge, and standard of living. Longevity is measured as the average life expectancy at birth, knowledge is a combination of adult literacy and mean years of schooling data, and standard of living is measured by real per capita income adjusted for its purchasing power within the country. HDI values range on a scale of 0 to 1, with 1 representing the highest level of development.

For each model, a scatter plot of the male-to-female wage ratio against the GDP per capita or HDI will be created to look for any obvious indication of a trend across countries. The greater the wage ratio, the larger the wage gap between genders in the country. Hence, a ratio of 1 represents gender wage equality, while a ratio of 2 means that men have an estimated earned income twice as great as women of the same country. As predicted by the theory, an inverted-U relationship is expected between the two variables.

The scatter plot of the male-to-female wage ratio against GDP per capita with a best fit quadratic line is depicted in Figure 2. While it is by no means a perfect inverted-U curve, there is a noticeable cluster of countries with low GDP per capita levels and lower wage ratios followed by higher wage ratios at slightly higher GDP levels. Countries with high incomes per capita – above approximately $25,000 – tend to have lower male-to-female wage ratios, with just a couple of high outliers.

Figure 3 shows a scatter plot of Human Development Index against the Male-to-Female Earnings Ratio with a quadratic trend line. There is clearly an inverted-U pattern present in the data. As development increases (measured by HDI) inequality between male and female earnings increases. At higher levels of development, however, wage inequality begins to decrease.
The percent of women in the labor force with different levels of education would be an ideal measure for operationalizing the education variable. The World Bank’s World Development Report reports this variable; however, it is missing data for a large percent of the countries under study. Adult literacy rates by gender are available from the Human Development Report and are used as an independent variable in the study as the male to female literacy ratio.

General wage inequality of a country can be measured in a number of ways. The Gini coefficient is one such measure commonly used for comparing income inequality across countries. A country’s Lorenz curve measures income actually received by households as a percentage of total national income. The equality line depicts perfectly equal income distribution. For example, if ten percent of the country earned ten percent of the income, twenty percent earned twenty percent of the income, and so on. The Gini coefficient is calculated by dividing the area between a country’s Lorenz curve and the perfect equality line by the total area under the equality line. This results in a coefficient ranging from 0, representing perfect equality, to 100, which indicates perfect inequality. Gini coefficients between 20 and 35 indicate relatively equitable distributions of income, while values between 50 and 70 signify highly unequal distributions (Todaro and Smith, 2006). The Gini coefficients for a large number of the countries included in the study are available from the Human Development Report. This variable is included because previous research has found that the gender wage gap varies with general wage inequality.

Descriptive statistics of the data set are given in Table 1. After removing 33 countries from the study due to lack of data on one or more variables, a total of 121 countries are included in the data set. Data are grouped according to the income categories specified by the World Bank: high income, gross national income per capita of $9,386 or more; upper middle income, $3,036-9,385; lower middle income, $766-3,035; and low income, $765 or less. One observation important to note is that the means of the dependent variable show a Kuznets-type relationship. The average value for low and high income countries is lower than that of the middle income groups.
There was large variation across countries in regard to each of the variables. The dependent variable ranged in value from near equality at 1.08 to significantly imbalanced at 3.84. The GDP per capita of the countries included in the study went from a low of $548 to a high of $37,738. The adult literacy rate ratio included a couple countries in which women had the advantage (low of .82), but in most nations men recorded higher rates, with a ratio high at 2.44. The Gini coefficient, where 0 represents perfect equality and 100 signifies perfect inequality, ranged from low inequality at 24.7 to high inequality at 70.7.

<table>
<thead>
<tr>
<th>Table 1: Descriptive Statistics for Data by Income Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable and Group</td>
</tr>
<tr>
<td>Male to Female Earnings</td>
</tr>
<tr>
<td>High Income</td>
</tr>
<tr>
<td>Upper Middle Income</td>
</tr>
<tr>
<td>Lower Middle Income</td>
</tr>
<tr>
<td>Low Income</td>
</tr>
<tr>
<td>GDP per Capita</td>
</tr>
<tr>
<td>High Income</td>
</tr>
<tr>
<td>Upper Middle Income</td>
</tr>
<tr>
<td>Lower Middle Income</td>
</tr>
<tr>
<td>Low Income</td>
</tr>
<tr>
<td>Gini Coefficient</td>
</tr>
<tr>
<td>High Income</td>
</tr>
<tr>
<td>Upper Middle Income</td>
</tr>
<tr>
<td>Lower Middle Income</td>
</tr>
<tr>
<td>Low Income</td>
</tr>
<tr>
<td>Male to Female Literacy</td>
</tr>
<tr>
<td>High Income</td>
</tr>
<tr>
<td>Upper Middle Income</td>
</tr>
<tr>
<td>Lower Middle Income</td>
</tr>
<tr>
<td>Low Income</td>
</tr>
</tbody>
</table>

IV. Empirical Method

My hypotheses to be tested are as follows:
1. Gender wage inequality will follow an inverted-U pattern with respect to the level of economic development of countries.
2. The disparity in educational attainment by gender will be positively related to the size of the gender wage gap.
3. The degree of general wage inequality in a country will be positively related to the size of the gender wage gap.

Regression analysis will be used to test the importance of each independent variable on the gender wage gap. In order to test the inverted-U hypothesis, both the GDP per capita and the square of the GDP per capita will be included as independent variables. It is expected that the GDP per capita will be positively related, while the square of the vari-
able will be negatively related to the wage ratio. Table 2 lists the variables with their descriptions and expected sign of their coefficients.

The equations to be tested are as follows:

Model 1: \( \text{M} \text{Earning/FEarning} = \alpha_1 + \alpha_2 (\text{GDP per capita}) + \alpha_3 (\text{GDP per capita})^2 + \alpha_4 (\text{MLit/FLit}) + \alpha_5 (\text{Gini}) + \epsilon \)

Model 2: \( \text{M} \text{Earning/FEarnings} = \alpha_1 + \alpha_2 (\text{HDI}) + \alpha_3 (\text{HDI})^2 + \alpha_4 (\text{MLit/FLit}) + \alpha_5 (\text{Gini}) + \epsilon \)

Where \( \text{ML}_{it} / \text{FL}_{it} \) is the ratio of the male-to-female adult literacy rates, and, as a measure of educational attainment, is predicted to be positively related to the male-to-female wage ratio. Gini, as a measure of general wage inequality, is also predicted to be positively related to the dependent variable.

**Table 2: Variable Descriptions and Expected Signs**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Expected Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mearn/Fearn</td>
<td>Estimated yearly earnings of males divided by those of females</td>
<td></td>
</tr>
<tr>
<td><strong>Independent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP per capita</td>
<td>Gross domestic product per capita in PPP US$ for the most recent year 1991-2003</td>
<td>+</td>
</tr>
<tr>
<td>HDI (Model 2)</td>
<td>Human Development Index value from the 2005 Human Development Report</td>
<td>+</td>
</tr>
<tr>
<td>((\text{GDP per capita})^2)</td>
<td>Above variable squared</td>
<td>-</td>
</tr>
<tr>
<td>((\text{HDI})^2) (Model 2)</td>
<td>HDI value squared</td>
<td>-</td>
</tr>
<tr>
<td>M/Lit/FLit</td>
<td>Adult male literacy rate divided by the adult female literacy rate</td>
<td>+</td>
</tr>
<tr>
<td>Gini</td>
<td>Gini coefficient as reported by the World Bank</td>
<td>+</td>
</tr>
</tbody>
</table>

**V. Results**

Model 1:

The results of the OLS regression analysis (Table 3), while not robust, do provide tentative support for the hypotheses. All coefficients had the sign predicted by the literature. Only the Gini variable coefficient was significant at the .01 level with a value of .005. However, GDP per capita and GDP per capita squared were close to being significant at the .05 level. The ratio of male-to-female literacy rates was not significant and also had a small coefficient. Overall, the variables predict only a small amount of the variance in the dependent variable, with an R square value of .118.
Model 2:
The results of regression analysis of Model 2 that are presented in Table 4 show that each variable is significant and has the expected sign. The R square value of .298 is over twice that of Model 1, and indicated that Model 2 explains approximately 30 percent of the variation in the dependent variable. Figure 4 plots the male-to-female earnings ratio values predicted by the regression results for Model 2. In calculating the predicted values the mean values for MLit/FLit (1.18) and Gini (40.33) were used, and HDI was allowed to vary. The simulation illustrates that the model does in fact predict an inverted U curve which peaks at an HDI value of approximately .79.

### Table 3: Summary Results of OLS Regression Analysis – Model 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Significance</th>
<th>t - Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.901</td>
<td>0.015</td>
<td>2.47</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>3.95E-05</td>
<td>0.067</td>
<td>1.852</td>
</tr>
<tr>
<td>(GDP per capita)^2</td>
<td>-1.18E-09</td>
<td>0.055</td>
<td>-1.938</td>
</tr>
<tr>
<td>MLit/FLit</td>
<td>0.281</td>
<td>0.135</td>
<td>1.505</td>
</tr>
<tr>
<td>Gini</td>
<td>0.015</td>
<td>0.005</td>
<td>2.894</td>
</tr>
<tr>
<td>R Square</td>
<td>0.298</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>121</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4: Summary Results of OLS Regression Analysis – Model 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Significance</th>
<th>t - Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-4.320</td>
<td>0.000</td>
<td>-4.466</td>
</tr>
<tr>
<td>HDI</td>
<td>11.129</td>
<td>0.000</td>
<td>5.316</td>
</tr>
<tr>
<td>(HDI)^2</td>
<td>-7.033</td>
<td>0.000</td>
<td>-4.772</td>
</tr>
<tr>
<td>MLit/FLit</td>
<td>1.121</td>
<td>0.000</td>
<td>4.933</td>
</tr>
<tr>
<td>Gini</td>
<td>0.021</td>
<td>0.000</td>
<td>4.491</td>
</tr>
<tr>
<td>R Square</td>
<td>0.298</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>121</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: Model 2 - HDI
Supplementary Analysis – Model 3:

Model 2 explained a significant amount of variance in the gender wage gap, and supported each of the hypotheses tested. A large amount of variance remained to be explained. Examining the countries which continued to be outliers through residual analysis suggested the possible significance of religious or cultural factors in determining the size of the wage gap. The initial models did not include any additional factors which may preclude women from fully participating in the work force and therefore result in lower earnings.

A third model was created with dummy variables added to account for the potential influence of culture on the magnitude of the gender wage ratio. Within the dataset, 34 countries were classified as predominately Roman Catholic, and 27 as predominately Muslim. Classification was based on the country having fifty percent or more of the population belonging to the religion according to the CIA World Factbook. There were no other religions for which the sample size was large enough to be included as a variable. The minimum, maximum, and mean values of the dependent variable are reported in Table 5 for the countries classified as Roman Catholic or Muslim.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/F Earnings Ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roman Catholic</td>
<td>34</td>
<td>1.39</td>
<td>3.7</td>
<td>2.22</td>
</tr>
<tr>
<td>Muslim</td>
<td>27</td>
<td>1.38</td>
<td>3.84</td>
<td>2.23</td>
</tr>
</tbody>
</table>

Results of the regression for Model 3 are summarized in Table 6. All original variables remained highly significant, although the coefficient values decreased somewhat for each. Model 3 explains almost 40 percent of the variance in the dependent variable, as seen by the R Square value of .396. The dummy variables for both Roman Catholic majority and Muslim majority were highly significant and positively related to the dependent variable. A country having either a Roman Catholic or Muslim majority increases the predicted size of the gender wage gap. Figure 5 plots the predicted values based on the results of Model 3 using the mean values for the MLit/FLit and Gini variables, and assuming a value of 0 for each of the dummy variables (meaning neither Roman Catholic nor Muslim).
VI. Conclusion

The results of this study provide support for the presence of a Kuznets variety inverted-U curve between GDP per capita and gender wage inequality across nations. GDP per capita was positively related to the size of the gender wage gap while the square of it was negatively related. This shows that the relationship holds at least somewhat that as economic development increases the size of the gender wage gap increases, but at high levels of per capita income the difference in pay decreases.

The results of Model 2 using Human Development Index values as the measure of development are more robust that those using the traditional Kuznets variable, GDP per capita. All four variables are significant at the highest level, and the R square value is over double that of the first model. Model 3 explains an additional ten percent of the variance in the dependent through the addition of dummy variables for the Roman Catholic and Muslim religions. Plots of the values predicted by Models 2 and 3 both have a clear curve which peak at an HDI value of approximately .79. The results suggest that the decrease in gender wage inequality is not expected to be seen until countries reach development levels close to .80 on the Human Development Index.

The role of human capital in determining the gender wage gaps as measured by the ratio of literacy rates did not prove significant in Model 1, but was significant in Models 2 and 3. The gap between men and women in adult literacy was positively related to the size of the gender wage gap. While differences in literacy rates may serve as a reasonable proxy for human capital differences in some countries, it is possible that in a large number of countries literacy is so high for both genders that it does not show much of the difference. Research using a different measure of human capital differences which allows for more of the actual human capital difference to be observed may prove to have even more significance on the gender wage gap.

The importance of overall wage inequality (as measured by the Gini coefficient) on gender wage inequality was seen to be significant in this study. One policy implication that can be seen stems from the significance of the Gini coefficient variable in the analysis. As suggested by Blau and Kahn (1996) and supported by Kidd and Shannon (1996) as well as Brainerd (2000), general wage inequality within a country is positively related to gender wage inequality. One way in which a government could take action to reduce

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Significance</th>
<th>t - Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-3.537</td>
<td>0.000</td>
<td>-3.809</td>
</tr>
<tr>
<td>HDI</td>
<td>9.44</td>
<td>0.000</td>
<td>4.661</td>
</tr>
<tr>
<td>(HDI)^2</td>
<td>-5.974</td>
<td>0.000</td>
<td>-4.173</td>
</tr>
<tr>
<td>MLit/FLit</td>
<td>0.918</td>
<td>0.000</td>
<td>4.04</td>
</tr>
<tr>
<td>Gini</td>
<td>0.019</td>
<td>0.000</td>
<td>4.049</td>
</tr>
<tr>
<td>Roman Catholic</td>
<td>0.371</td>
<td>0.000</td>
<td>3.619</td>
</tr>
<tr>
<td>Muslim</td>
<td>0.353</td>
<td>0.004</td>
<td>2.979</td>
</tr>
<tr>
<td>R Square = 0.396</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N = 121</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 6: Summary Results of OLS Regression Analysis – Model 3**
the disparity in incomes between genders would be to work to lower the overall income gap, or at a minimum keep it from growing.

The results of Model 3 suggest an avenue for further research. Dummy variables for both Roman Catholic and Muslim (defined as fifty percent or greater of the population) were statistically significant and positively correlated with the size of the gender wage gap. Model 3 was run only as a preliminary look at the relevance of religion on the gender wage gap, and significant further research could be done on the subject.

References


I. Introduction

In the last 24 years, international official remittance flows have almost tripled, rising from 43.3 billion US dollars in 1980 to 125 billion in 2004. Currently, remittances have a volume roughly double that of official development assistance flows, becoming the second largest source of external financing in developing countries, only after foreign direct investment. This rapid growth has invited significant attention from the international community and motivated several studies on the role of remittances on economic development, but there is yet no established consensus even on whether remittances have a positive or negative macroeconomic impact on recipient countries.

As a new field of study, methods for measuring and interpreting the role of remittances on the economic development of recipient countries have been varied and often problematic. Rather than trying to find a direct link between remittances and economic growth, this paper focuses on a channel through which remittances may have a macroeconomic effect on a recipient country – through their effect on national consumption and savings patterns.

II. Literature and Theories

The most popular method for measuring the impact of remittances on economic development is through studying the relationship between remittances and per capita GDP or per capita growth. However, the idea that the national rate of investment, or savings, positively affects economic growth is widely accepted in the economics field, as the Solow (1956) Growth Model argues by showing that countries with higher savings rates, and, consequently, lower shares of consumption-to-GDP ratios, have higher output and productivity rates. Since output and productivity rates are clearly two factors that contribute to economic development, it is useful to conduct a study on the impact of remittances on consumption and savings patterns to assist in a better understanding of the role of this huge international financial flow in economic development.

While a more thorough literature review is presented in the full-length version of this
paper, it is clear that the empirical results on the economic impact of international remittances on recipient countries has produced numerous conflicting responses as well as highly informative and reliable results using survey data which are, however, unfortunately limited to only a few countries or regions and cannot be safely extrapolated to represent world trends on remittance allocations. In the studies assessed, different assumptions were made and distinct theories were chosen as their basis. From the inexhaustible list of theories presented throughout the literature, this study chooses to rely on two of them for their simplicity and general acceptability in economics: Friedman’s Permanent Income Hypothesis and the endogeneity of savings with respect to income. These two concepts provide some understanding of the behavior of remittance recipient households and should be considered in the design of an empirical model as well as in the interpretation of empirical results, since it may be extrapolated into a country-wide understanding of the macroeconomic role of remittances.

While the role of the permanent income hypothesis in assisting in the understanding of remittance allocation is fairly simple, dependent on whether the recipient households perceive remittance inflows to be permanent or temporary income flows, it is worth delving deeper into the second theory and looking into how the potentially endogenous relationship between savings and income may affect remittance allocation decisions. When households have very low income levels, their savings rates are also very low and do not increase rapidly with a change in income, since there are minimum subsistence costs that households must cover before being able to put money aside for savings. However, as households’ incomes increase beyond subsistence levels, savings rates may increase dramatically, possibly because of a strong desire for financial stability or because financial success appears to be a more realistic long-term goal and thus one for which it is worth saving more, besides other potential explanations.

By accepting this description of the relationship between savings and income, however, one is still not able to pinpoint the savings rate of remittance recipient households. At the same time that these households are unlikely to be on the upper half of the income scale, it is too hasty to assume that they are on the far left of the graph either. International immigration is costly for households, since it requires visas, or forged documentation, as well as transportation and initial living costs, besides financial and physical risks, among other costs. Thus, recipient households are comprised of those families who can temporarily incur these costs in order to send members abroad. If these costs are large enough, remittance recipient households may actually be further towards the middle of the graph, where savings rates increase significantly with income.

III. Model and Data

Taking the theories briefly explained above into consideration, the general form of the model
designed to measure the macroeconomic impact of remittance receipts through consumption and savings rates is as follows:

\[
\begin{align*}
\ln(\text{Consumption}_{it}) &= \beta_1 + \beta_2 \ln(\text{Remittances}_{it}) + \beta_3 \ln(\text{GDP}_{it}) + \beta_4 \ln(\text{GDPpc}_{it}) \\
&+ \beta_5 \text{Inflation}_{it} + \beta_6 \text{Female}_{it} + \beta_7 \text{LowInc}_{it} + \beta_8 \text{LowMidInc}_{it} + u_i + \varepsilon \\
\ln(\text{Savings}_{it}) &= \beta_1 + \beta_2 \ln(\text{Remittances}_{it}) + \beta_3 \ln(\text{GDP}_{it}) + \beta_4 \ln(\text{GDPpc}_{it}) \\
&+ \beta_5 \text{Inflation}_{it} + \beta_6 \text{Female}_{it} + \beta_7 \text{LowInc}_{it} + \beta_8 \text{LowMidInc}_{it} + u_i + \varepsilon
\end{align*}
\]

and the instrumental equation is:

\[
\ln(\text{Remittances}_{it}) = \alpha_1 + \alpha_2 \text{NetMigration}_{i(t-5)} + \mu
\]

The key independent variable, remittances, was taken directly from a WDI Online database, while the dependent variables consumption and savings were calculated from consumption-to-GDP and savings-to-GDP ratios and GDP. Other variables included in the model were GDP, which controls for the relationship between the economic size of a country and its ability to save and consume, as well as GDP per capita, accounting for the income level of a household and its ability to save and consume.

A variable to control for the generally accepted relationship between savings and interest rates had to be taken into consideration as well. However, there is no extensive and reliable data on interest rates, nor on its closest alternative measure, the almost risk-free rate offered on national Treasury bills. Thus, inflation rates were used as a proxy for their tendency to move with interest rates. The appropriateness of this substitution, however, is discussed in the results section of this synopsis.

The fifth independent variable included in the empirical study was the proportion of females in the population, since the literature review suggested a variation in remittance allocation patterns based on the gender of the head of the recipient household. Thus, female is actually a proxy for the proportion of households headed by females, since the use of this variable requires the assumption that the proportion of female-led households in a country is similar to the percentage of females in the country.

The last two variables are dummy variables indicating the classification of a country as a low income country or low-middle income country by the World Bank, which defines low income countries to be those with a 2004 per capital Gross National Income of up to $825, and low-middle income countries as those with per capita GNIs from $826 to $3255. These two variables account for the possible nonlinear relationship between savings and income, to verify that there is a minimum threshold that must be surpassed before savings and consumption rates can be significantly affected by income.

An instrumental equation using net migration was also used to adjust for the possibility of the endogeneity of remittances with respect to consumption, as was suggested by the literature arguing for the countercyclical nature of remittances. Results from the Hausman test confirmed the endogeneity of remittances with respect to both savings and consumption, which was corrected by using net migration from 5 years prior to each observation as an instrument.
IV. Results

Using data available for 42 remittance recipient countries, per capita GDP, inflation, and the low-middle income dummy variable proved to be statistically insignificant for consumption and savings levels, and the percentage of females in the population were found to be only weakly significant for consumption and not significant at the 5%-level for savings. Coefficients for the other variables, however, were found to be statistically robust.

While the strength of the coefficient of the GDP measurement was perhaps unsurprising, the statistically significant coefficients for remittances and the low income dummy variable deserve further comments. The empirical analysis showed the elasticity of consumption with respect to remittances as 16%, and -37% to be the elasticity of savings with respect to remittances. These results, which are of course based only on official international remittances, as already mentioned above, indicate that remittance receipts are in fact consumption-biased. Although it may be unsurprising that the relationship between remittances and consumption is positive, since it would be unreasonable to expect a remittance recipient household to save all of the remitted money, the inverse relationship between remittances and savings is certainly revealing.

The final statistically robust independent variable, the dummy variable classifying a country as low income, with a 2004 GNI per capita of less than $825, also carries an interesting coefficient. Low income countries are shown to have a consumption rate 29% higher than non-low income countries, and save 52% less than countries with GNI per capita over $826. Although it may be expected that poorer countries have a higher consumption rate and lower savings rate than richer countries, it must be noted that the low-middle income dummy variable as well as the per capita GNP variable were shown to be statistically insignificant, indicating that having a GNI per capita under $825 or more than that is more significant than each dollar increase in per capita GNP or having under $3255 or more. Thus, there is in fact a threshold determining consumption and savings rates, pointing to a non-monotonic relationship between income and allocations.

V. Conclusion and Further Study

The empirical analysis summarized in this synopsis presents results indicating a bias of the allocation of remittance receipts towards consumption rather than savings. While the positive elasticity of consumption with respect to remittances might not be surprising, when combined with the finding of an inverse relationship between remittance receipts and savings, it may present an interesting insight into the way through which remittances contribute to economic development. Policy planners and government leaders may find the negative relationship between savings and remittances to be somewhat discouraging given its large volume and, consequently, potential for huge macroeconomic impacts if channeled into investment. However, the allocation of remittances towards consumption does not necessarily paint a dismal picture, particularly if it tends towards the consumption of productivity-enhancing goods, such as machinery or even food, or if consumption is biased towards domestically produced goods. This particular allocation of remittance
receipts within different consumption options is a potential extension of this paper and certainly an area worthy of further study.

Another possible contribution to this paper concerns the two variables chosen for this empirical analysis that were found to be statistically insignificant despite literature indicating otherwise. The proxy chosen for interest rates, inflation, was found to be statistically insignificant for both savings and consumption measures, although it is difficult to negate the positive relationship between savings and interest rates, most likely pointing not to an error in the model itself, but to the choice of an inadequate proxy. Measuring interest rates, or finding an adequate proxy for such, has been a stumbling block for several empirical analyses, not exempting this paper, so that if a more adequate measure is found, it would be important to conduct this empirical analysis again. The same should be done once an adequate measure of the percentage of female-headed households becomes available, this time to verify whether being led by a female household head does not actually affect allocation patterns of remittance receipts as the results of this empirical study indicates, or if the proxy chosen for this study, the percentage of females in the country population, is actually an inadequate substitute. Finally, it must again be noted that the empirical study described in this paper was conducted using values for only official international remittances rather than total remittances. This omitted variable bias in the models presented may be causing a distortion on the coefficients, so that they appear larger than they would be if total remittance receipts were measurable and included. While this proves to be problematic in providing an accurate picture of the macroeconomic impact of overall remittances, there is currently no standardized database using household survey data on remittances for various countries. This problem can be resolved in two alternative ways: either by furthering data collection on remittances through household survey data as is being studied in joint efforts by the World Bank, IMF, and other organizations, or encouraging a greater proportion of remittances to be sent as official remittances. It is important to note, furthermore, that encouraging the remittance of monies through official channels rather than informal mechanisms helps not only in a more accurate study of the role of remittances on economic development, but also, and more importantly, may encourage a trend towards allocating remittance receipts towards savings. Receiving remittances in banks or credit unions, for example, exposes recipients to the banking system, which sometimes even opens a bank account. The exposure to savings, credit, and insurance products, then, furthers the potential macroeconomic contribution of remittances by increasing the number of households with bank accounts and, possibly, savings deposits.
References


MULTILATERAL VS. BILATERAL AID: DIFFERING EFFECTS ON GROWTH?

Jennifer Hardy
Georgetown University

I. Introduction

This paper addresses the aid effectiveness question by combining several strands of existing growth literature. Given the large volumes of official development assistance that flow across the globe on an annual basis, both donor and recipient countries have a vested interest in determining whether those funds actually contribute to economic growth and poverty reduction, among other goals. These are two separate questions, and the analysis below will focus on the aid/growth relationship. Three open questions in the literature are addressed: whether multilateral aid and bilateral aid have differing impacts; whether there is a non-linear relationship between aid and growth; and whether estimates of aid effectiveness are affected by endogeneity. Based on the differing motivations between allocation of aid from multilateral institutions and aid provided directly by bilateral donors, multilateral aid is expected to have a more positive impact on growth, although this effect may be obscured by biased estimates due to reverse causality. This hypothesis will be tested empirically.

II. Literature Review

Economists have hotly debated the role of aid in growth, both theoretically and empirically, for years. Peter Boone (1996) notes that “Foreign aid programs were launched long before there was compelling theory, or compelling evidence that proved they could work.” His assertion is borne out by the continuing dispute over aid’s true effectiveness.

Much past research focuses on the straightforward aid – growth relationship, with some studies taking account of potential reverse causality. Models in this strand take the form

\[ g_{it} = \alpha + \mu h_{it} + \gamma x_{it} + u_{it} \]

where h is the aid share and x is a vector of country-specific control variables. Overall, the small positive coefficients observed are not significant. One recent model that has been hotly debated is the “good policy” model first advocated by Burnside and Dollar (1996), which purported to show that aid has a significant positive effect in countries characterized by fiscal surplus, low inflation, and trade openness. An index constructed from these
measures and interacted with the share of foreign aid in GDP had a positive coefficient in growth regressions. However, their findings have since been shown to be highly sensitive to choice of data set and model specification (see, for example, Easterly, Levine and Roodman [2004]).

Recent studies have focused attention on the problem of endogeneity in estimates of the aid-growth relationship. If poorly-performing countries – those with low per capita GDP or, more specifically, low levels of per capita GDP growth – are disproportionately targeted as aid recipients, a reverse causality bias could lower the true coefficients on aid terms in growth relationships. This appears to be supported by a significant negative correlation between per capita GDP levels and aid flows. Instrumental variables are the remedy of choice for this problem.

Ram (2003), who attempts to explain the lack of a significant relationship between aid and growth in past literature, decomposes aid into bilateral and multilateral components, and then using the Burnside and Dollar model to estimate coefficients on these variables. His motivation is based on the differential allocation of aid identified by Maizels and Nissanke. Ram further postulates that bilateral donors have developed special relationships with their aid recipients and devote their resources to more appropriate projects; he is therefore unsurprised to find a positive, significant coefficient on bilateral aid an a negative, equally significant coefficient on multilateral aid.

III. The Model

Based on Ram’s model, modified in response to issues highlighted in the literature, my model for the impact of aid on GDP growth is:

\[
\ln(GDPPC)_{it} = \beta_0 + \beta_1 \ln Y_{0it} + \beta_2 \text{EDUC}_{0it} + \beta_3 \text{BLAID}_{it} + \beta_4 \text{MLAID}_{it} + \beta_5 \text{BLAID}_{2it} + \beta_6 \text{MLAID}_{2it} + \beta_7 \text{INVSH}_{it} + \beta_8 \text{POLRISK}_{it} + \beta_9 \text{ECORG}_{it} + \beta_{10} \text{ETHNF}_{it} + \beta_{11} \text{OPEN}_{it} + u_i + e_{it}
\]

where $\Delta \ln(GDPPC)$ is the change in the log of GDP per capita in country $i$ during time period $t$, $Y_0$ is initial per capita GDP for the time period, EDU0 is a measure of initial human capital, BLAID is average real net bilateral aid as a share of real GDP, BLAID2 is BLAID squared, MLAID is average real net multilateral aid as a share of real GDP, and MLAID2 is MLAID squared. The first two variables incorporate the predictions of the neoclassical growth literature that human capital contributes to increased growth and that poorer countries exhibit conditional convergence. BLAID, BLAID2, MLAID and MLAID2 are the dependent variables of interest. Control variables include INVSH, the share of investment in real GDP averaged over the time period, to account for accumulation of capital; POLRISK, the International Country Risk Guide index of political risk, at its 1984 level; ETHNF, a measure of ethnolinguistic fractionalization; and OPEN, a crude measure of trade openness calculated as the ratio of total imports and exports to GDP. The time period for the model is 1970-2000; observations are carried out on a country $i$ during
a time period $t$. Ten-year time periods were used to minimize the effects of business cycles on growth. Observations are available on at least some variables for 68 countries, and no variable has fewer than 44 groups with observations.

IV. Data

Charts 1 and 2, below, show that bilateral aid and multilateral aid are both negatively correlated with contemporaneous growth; the trend is slightly stronger for multilateral aid.

Two-stage least squares is employed to account for the possibility of endogeneity. For bilateral aid, a crude instrument to predict fitted values incorporates country-specific real bilateral aid in period $t-1$, aggregate real bilateral aid in period $t$, and initial country population in period $t$. This construction relies on the observations that aid flows to individual countries are correlated but not constant over time, that the total level of aid provided during a given time period necessarily affects the level of aid received by individual countries, and that small countries tend to receive more aid per capita than large ones. An instrument for multilateral aid similarly relies on lagged real multilateral aid, current aggregate real multilateral aid and initial country population in period $t$. The aid data for these variables are again collected from the DAC database and the population data from the World Development Indicators.

V. Empirical Results

The results from the initial regression and regressions incorporating additional control variables are summarized in Table 1.
### Table 1  
**Dependent variable: Change in ln(GDP per capita)** (standard errors in parentheses)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Regression 1</th>
<th>Regression 2</th>
<th>Regression 3</th>
<th>Regression 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yo</td>
<td>-0.008 (.011)</td>
<td>-.028 (.014)</td>
<td>-.11 (.032)</td>
<td>-.00003 (.00001)</td>
</tr>
<tr>
<td>EDUC</td>
<td>-.37a (.11)</td>
<td>-.65c (.15)</td>
<td>-1.42 (.32)</td>
<td>-.006a (.001)</td>
</tr>
<tr>
<td>BLAID</td>
<td>-7.5 (2.48)</td>
<td>-4.95 (3.57)</td>
<td>-4.79 (6.25)</td>
<td>-.041 (3.45)</td>
</tr>
<tr>
<td>MLAID</td>
<td>-10.3a (3.69)</td>
<td>-4.6 (5.59)</td>
<td>-.85 (7.62)</td>
<td>-.015 (5.51)</td>
</tr>
<tr>
<td>BLAID</td>
<td>2.7 (10.96)</td>
<td>.14 (22.18)</td>
<td>.29 (37.8)</td>
<td>.0837 (18.6)</td>
</tr>
<tr>
<td>MLAID</td>
<td>42.6b (21.30)</td>
<td>.26 (30.6)</td>
<td>.035 (41.9)</td>
<td>.092 (29.8)</td>
</tr>
<tr>
<td>INVSH</td>
<td>.14a (.005)</td>
<td>.016c (.009)</td>
<td>.0125b (.005)</td>
<td></td>
</tr>
<tr>
<td>POLRISK</td>
<td>.003 (.003)</td>
<td>.004b (.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECORG</td>
<td>.005 (.004)</td>
<td>.034 (.024)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETHNF</td>
<td>-.004 (.105)</td>
<td>.04 (.124)</td>
<td>-.07 (.098)</td>
<td></td>
</tr>
<tr>
<td>OPEN</td>
<td>-.001 (.001)</td>
<td>-.007a (.002)</td>
<td>-.0008 (.0007)</td>
<td></td>
</tr>
<tr>
<td>SSA</td>
<td>-1.12 (.095)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAME</td>
<td>.036 (.087)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEA</td>
<td>.120 (.088)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.03a (.12)</td>
<td>.75a (.24)</td>
<td>2.28a (.35)</td>
<td>.761a (.197)</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>N</td>
<td>197</td>
<td>116</td>
<td>116</td>
<td>125</td>
</tr>
<tr>
<td>R² Within:</td>
<td>0.22</td>
<td>0.31</td>
<td>0.48</td>
<td>0.33</td>
</tr>
<tr>
<td>Between:</td>
<td>0.22</td>
<td>0.55</td>
<td>0.02</td>
<td>0.65</td>
</tr>
<tr>
<td>Overall:</td>
<td>0.20</td>
<td>0.37</td>
<td>0.01</td>
<td>0.42</td>
</tr>
</tbody>
</table>

a: significant at the 1% level; b: significant at the 5% level; c: significant at the 10% level

We then turn to the suggested instruments for BLAID and MLAID identified in the data section of this paper. Table 2 presents the first stage of two-stage least squares.
### Table 2: First Stage of 2SLS: Random-effects panel regression

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BLAIDlag</td>
<td>.93A</td>
<td></td>
<td>.134A</td>
<td>(.06)</td>
<td>(.011)</td>
<td></td>
</tr>
<tr>
<td>BLAIDagg</td>
<td>-.16</td>
<td>(-.24)</td>
<td>-.03</td>
<td>(.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLAIDlag</td>
<td>1.30A</td>
<td></td>
<td>.149A</td>
<td>(.60)</td>
<td>(.009)</td>
<td></td>
</tr>
<tr>
<td>MLAIDagg</td>
<td>-2.07C</td>
<td></td>
<td>-.184</td>
<td>(1.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIDlag</td>
<td></td>
<td>1.11A</td>
<td>.277A</td>
<td>(.060)</td>
<td>(.018)</td>
<td></td>
</tr>
<tr>
<td>AIDagg</td>
<td></td>
<td>-.413</td>
<td>-.009</td>
<td>(.317)</td>
<td>(.009)</td>
<td></td>
</tr>
<tr>
<td>POP</td>
<td>-.21</td>
<td>-.12</td>
<td>-.39</td>
<td>-.062</td>
<td>-.00002</td>
<td>-.0001</td>
</tr>
<tr>
<td>Constant</td>
<td>.025</td>
<td>.016</td>
<td>.046</td>
<td>.0023</td>
<td>.0001</td>
<td>.005</td>
</tr>
<tr>
<td>N</td>
<td>197</td>
<td>197</td>
<td>197</td>
<td>197</td>
<td>197</td>
<td>197</td>
</tr>
<tr>
<td>R² Within</td>
<td>0.09</td>
<td>0.41</td>
<td>0.25</td>
<td>0.05</td>
<td>0.42</td>
<td>0.23</td>
</tr>
<tr>
<td>Between</td>
<td>0.91</td>
<td>0.92</td>
<td>0.92</td>
<td>0.83</td>
<td>0.79</td>
<td>0.84</td>
</tr>
<tr>
<td>Overall</td>
<td>0.56</td>
<td>0.71</td>
<td>0.66</td>
<td>0.45</td>
<td>0.62</td>
<td>0.58</td>
</tr>
</tbody>
</table>

A: significant at the 1% level; B: significant at the 5% level; C: significant at the 10% level
VI. Conclusion

This paper seeks to address three main questions raised by the aid effectiveness literature. First, do the marginal impacts of multilateral and bilateral aid on growth differ? In other words, are the coefficients on these variables in a growth regression significantly different? The coefficients on BLAID and MLAID are not found to be significantly different once control variables are added to the equation; nor are the coefficients on the squared terms for these variables. The coefficient on MLAID is significant and negative, while that
on MLAID2 is significant and positive in a basic regression with initial GDP per capita and human capital as the only control variables; the coefficients on BLAID and BLAID2 are not significantly different from 0. Adding more variables eliminates the significance of all aid terms. I fail to replicate Ram’s finding that the coefficients on bilateral and multilateral aid are significantly different once control variables are included in the regression. One possible explanation could be that the coefficients on effective development assistance vary between multilateral and bilateral aid (Ram used EDA while I employed ODA). However, a parsimonious likelihood-ratio test failed to reject the null hypothesis that the constrained model was insignificantly different from the unconstrained model (see appendix). Given that my variables were otherwise similar to Ram’s, with the inclusion of squared terms, it seems that his choice of a cross-section rather than panel regression may have been the determining factor in these very different findings.

Second, does inclusion of squared terms better explain the relationship between aid and growth? That is, is there a nonlinear relationship (either decreasing or increasing) that produces a better fit for the model? Regression 1 found a negative coefficient on MLAID and a positive coefficient on MLAID2, suggesting the exact opposite of the decreasing returns found by Hansen and Tarp, among others. This effect is not robust and diminishes as additional control variables enter the regression. When two-stage least squares was used, both linear and squared aid terms became significant (at the 5% or 1% level). This significance, which matches the predicted effect of decreasing returns to aid due to limited investment opportunities or inability to absorb unlimited capital flows, indicates that future investigation of the aid-growth relationship should continue to consider nonlinear models.

Finally, does endogeneity affect estimates of aid effectiveness? Instruments were suggested and tested for the level of bilateral aid and multilateral aid received by a country. Two-stage least squares using these fitted values offers estimates that are significant and show the positive yet decreasing returns to aid (of both kinds) previously found by Hansen and Tarp and Dalgaard and Hansen. The chosen instruments (lagged aid, aggregate aid and population) attempt to provide both time-series and cross-country variation while estimating variables that are uncorrelated with the error term in the growth regression.

Overall, the findings of this paper largely reject Ram’s assertion that past empirical work should be re-estimated using decomposed aid terms, as well as his contention that endogeneity is not a matter of concern in the aid effectiveness literature. Researchers’ time and resources might be better spent identifying ways in which aid can be made non-fungible or used to support institutional growth and political stability, which also appear to have positive effects on economic growth. The diminishing returns to aid indicate that simply increasing aid flows is not the answer to disappointing development performance; at best, my findings indicate that aid can support growth when it makes up a significant but minor portion of GDP.
References


Center for Global Development, “Data Sets and Resources.” http://www.cgdev.org/section/topics/dataset


MULTILATERAL VS. BILATERAL AID


USAID Global Education Database, http://qesdb.usaid.gov/cgi-bin/broker.exe?_program=gedprogs.ged_theme_une.sas&_service=default

World Development Indicators Online, access provided by Georgetown University Library. http://0-devdata.worldbank.org.library.lausys.georgetown.edu/dataonline/
A Natural Experiment on Reference-Dependent Model – A Test of Alternative Theories: Traditional Expected Utility Theory vs. New Prospect Theory

Xiaoti Zhang
University of Warwick

ABSTRACT

The tendency for people to value outcomes in terms of gains or losses is all pervasive and naturally extends to the work place, where workers usually have personal goals in mind (in terms of gaining or losing rewards) when they face an incentive reward scheme and decide how hard to work as a result.

I. Introduction

This project focuses on investigating the existence of loss aversion in the labour supply market. When people make labour supply decisions, do they value a loss more than they would a gain (reference-dependent model) or do they give equal weight to them both (traditional expected utility theory)?

The method is to run a natural experiment in the workplace by introducing a constant number of daily reward streams to serve as a reference level and studying the labour supply response to this by observing the changes of workers’ efforts during the experiment.

The exact reward scheme follows rank-order tournament structure, where we give the bonus to the best performance shop assistant in every session. The basic idea is that under expected utility model, Nash equilibrium exists where gain from the introduction of the rank-order tournament is equal to loss from its removal. However, under the reference-dependent model, Nash equilibrium exists where loss is more than gain.

Thus, as each model leads to strikingly different predictions in terms of experiment results, this experiment will serve as a test of the validity of these alternative theories: traditional expected utility theory vs. new prospect theory (i.e. reference-dependent model).
**II. Theoretical Model**

**2.1 Basic Model Setup**

EUT:  \( U(i,t) = W(e_{i,t}) - C(e_{i,t}) \); RDM:  \( U(i,t) = \frac{W(e_{i,t})}{W(e_{i,t})} - C(e_{i,t}) \)

Where  \( U(i,t) \) - Utility in period t for individual i (concave function)  
\( W(e_{i,t}) \) - Wage level in period t for individual i (non-convex function)  
\( U(i,t) \) - Cost associated with chosen effort level for individual i at period t (convex)

For simplicity, we take the simplest forms of both wage and cost function. Thus,  
\( W(e_{i,t}) = e_{i,t} \) (Linear form) and  
\( C(e_{i,t}) = e_{i,t}^2 \) (quadratic form)

**2.2 Reference-Dependent Model and Expected Utility Model Results**

**TABLE 1: THE NASH EQUILIBRIUM OUTCOMES UNDER DIFFERENT PERIODS**

<table>
<thead>
<tr>
<th>Equilibrium/Periods</th>
<th>Period One</th>
<th>Period Two</th>
<th>Period Three</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref. Dependent Model</td>
<td>( e_{i,1} = \frac{1}{\sqrt{2}} )</td>
<td>( e_{i,2}^* = \frac{1 + b}{\sqrt{2}} )</td>
<td>( e_{i,3}^* = \frac{1}{\sqrt{2}(b+1)^2} )</td>
</tr>
<tr>
<td>Expected Utility Model</td>
<td>( e_{i,1}^* = \frac{1}{2a_i} )</td>
<td>( e_{i,2}^* = \frac{1 + b}{2a_i} )</td>
<td>( e_{i,3}^* = \frac{1}{2a_i} )</td>
</tr>
</tbody>
</table>

Under the Reference Dependent Model:

The Gain= \( (e_{i,2}^* - e_{i,1}^*) = \frac{b}{\sqrt{2}} \) and Loss= \( (e_{i,2}^* - e_{i,3}^*) = \frac{(b+1)^3 - 1}{\sqrt{2}(b+1)^2} \)

So, under the Reference Dependent Model, Loss of effort due to removal of bonus is bigger than the increase of effort due to introducing bonus.

Under the Expected Utility Model, the increase of effort due to introducing the incentive scheme must be the same as the decrease of effort due to taking off the scheme.

**III. Experimentation Methodology**

**3.1 Experiment goal**

The goal of the experiment is to detect the existence of loss aversion by comparing the effort levels under period one and period three and the diminishing sensitivity property by checking the changes of effort levels within period two. Hence this will differentiate the Reference-Dependence Model from the Expected Utility Model.

**3.2 Experiment design**

The experiment duration is 8 weeks, split into three periods of 2, 4 and 2 weeks respectively. In period two, we introduce a constant number of daily reward streams into the shop. The reward amount is £5 per session. We run two independent rank-order tour-
nament games everyday, separating them by the 1pm lunch break.

The manager provides two workers with the same instruction at the beginning of each game. The order is:

“We are running an incentive scheme this morning/afternoon. A bonus worth £2.50 will be given at the end of this morning/afternoon session to the best performing shop assistant. Performance is assessed according to the individual accumulative sales figure on the Shop Assistant Productivity Table at the end of each session.”

The structure of the bonus follows the Rank-Order Tournament Structure. The advantages are that it can not only serve the purpose of this paper (i.e. generating reference level) but also avoids some severe problems inherited from other forms of bonus scheme. For example, the teamwork rewards scheme would introduce free-riding problem. Importantly, we take the bonus scheme off at the end of period two without prior notice to the shop assistants, because it is vital to avoid adjusting the reference level.

IV. Experimental Results and Data Analysis

There are three main hypotheses that I would like to investigate:

Hypothesis One: The average effort in a bonus period is significantly higher than in periods without bonus. This hypothesis refers to the existence of bonus effect.

Hypothesis Two: The average effort after taking off rewards (period 3) is significantly lower than during the period pending the rewards (period 1), which corresponding to loss aversion property in the Reference-Dependent Model.

Hypothesis Three: The average effort level of a high ability type exerts a higher degree of loss aversion in absolute values.

The main findings of the experiment are: the average effort in a bonus period is significantly higher than in periods without bonus and that the average effort after taking off rewards is significantly lower than during the period pending the rewards. The latter corresponds to the loss aversion property in the reference-dependent model (prospect theory). Therefore, the main results are in support to the new prospect theory and in contrast to the traditional expected utility theory.

V. Conclusion

In conclusion, loss aversion in the labour supply market is an important phenomenon, both because of its economic magnitude and because of the insights it suggests about how psychologically (bounded rationality) feelings motivate economic behaviour in general.

The original contribution that this project makes is that it provides a recipe for doing convincing behavioural economics ‘in the field’. The simple rank-order tournament game structure has been integrated with reference-dependent model and act as a framework to generate a reference level. The hope is that the principles that govern this simple experimental case will be able to be applied in a recognizable form to more complex situations.
References:


THE DETERMINANTS OF PARTICIPATION AND SUCCESS AT OLYMPIC GAMES: A CROSS-COUNTRY ANALYSIS

Oyebanke Oyeyinka
Carleton College

ABSTRACT
This paper analyzes factors affecting participation, success and gender inequality at the Olympic Games. The main factors considered are economic, political and religious. The study involves 206 countries and analyzes all 14 Games between 1952 and 2004. Most studies have considered the factors affecting success at the Games but very few have analyzed participation and gender inequality at the Games. Political instability and unrest have previously not been explicitly studied while religion as reflective of culture is novel to the literature on Olympic participation and success. The results of this study show that economic factors were a significant determinant of both success and participation at the Games. Political instability specifically ethnic wars and genocide acted as a deterrent to countries attending the Games. Religion also had an effect on participation and success at the Olympics, but the effect did not differ much between the sexes.

“Not to win but to take part, just as the most important thing in life is not the triumph but the struggle.” – The Olympic Creed.

I. Introduction
The Olympic ideal is an admirable and idealistic portrayal of a significant world event; however, the reality of participation is different. The literature on the Olympics, as well as other sports, shows that countries and individuals participate and they do so primarily to succeed - win. Success at games translates to winning medals. This is in fact borne out by a survey on the current Winter Olympics taking place in Torino, Italy.

Different surveys using traditional methods as well as computer-based techniques are being used to assess the likely chances of individuals and countries to win medals. Recently, a mail entitled “Winter Olympics(R) Survey: Who Will Bring Home The Medals?” was sent by a National Panel Survey to various mailboxes. The options given on
this survey were – United States, Russia and Canada. The questionnaire suggested that Olympic medal winners are predictable enough for a spectator to tell who the overall winners will be even before the end of the Games. Although speculation on medal success might be considered a part of the Games, it reflects tendencies to concentrate on winning and less efforts being directed towards encouraging countries to participate. Economists have therefore tended to formulate the problem of participation on the propensity to participate and succeed, and significantly, much of the literature on the Olympics is directed towards the winning of gold medals.

This study follows this tradition. It poses a simple question: what factors determine participation and success at the Olympics games? We answer the question by formulating and testing three hypotheses on data based on 14 Summer Olympics. They follow the proposition that the determinants of participation and success are complex and the variables comprise a set of economic, political and religious factors.

The paper is organized as follows: the next section reviews the literature on the historical evolution of the Olympics Games explained through political, economic and religious lenses. This is followed by a review of the literature and theoretical frameworks of past works. Section five describes the data and the sources, while I present the empirical findings as well as direction for future research in the discussions in sections six and seven. Section eight presents the conclusions.

2.0 Background: Historical Participation at the Olympics

The Summer Olympics has seen increased participation in recent years. This increase in international participation has caused economists such as Johnson and Ali (2000), to observe country specific factors and world changes that stimulated a 54% increase from 93 nations who attended the 1964 Games held in Tokyo to 197 nations participating in Atlanta in 1996. The table below however shows that although participation by nations has increased over time, huge inequalities exist between the numbers of athletes sent by individual countries. The change in continental participation by countries and individuals from the countries are shown in the table below.

<table>
<thead>
<tr>
<th>Continents</th>
<th>Africa</th>
<th>America</th>
<th>Asia</th>
<th>Europe</th>
<th>Oceania</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years</td>
<td>1896</td>
<td>1908</td>
<td>1924</td>
<td>1936</td>
<td>1956</td>
</tr>
<tr>
<td></td>
<td>Nations</td>
<td>Nations</td>
<td>Nations</td>
<td>Nations</td>
<td>Nations</td>
</tr>
<tr>
<td></td>
<td>Athletes</td>
<td>Athletes</td>
<td>Athletes</td>
<td>Athletes</td>
<td>Athletes</td>
</tr>
<tr>
<td>1896</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1908</td>
<td>1</td>
<td>11</td>
<td>2</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>1924</td>
<td>2</td>
<td>55</td>
<td>10</td>
<td>524</td>
<td>3</td>
</tr>
<tr>
<td>1936</td>
<td>2</td>
<td>76</td>
<td>12</td>
<td>690</td>
<td>5</td>
</tr>
<tr>
<td>1956</td>
<td>7</td>
<td>101</td>
<td>17</td>
<td>662</td>
<td>18</td>
</tr>
<tr>
<td>1968</td>
<td>24</td>
<td>284</td>
<td>32</td>
<td>1541</td>
<td>21</td>
</tr>
<tr>
<td>1980</td>
<td>21</td>
<td>485</td>
<td>15</td>
<td>530</td>
<td>14</td>
</tr>
<tr>
<td>1992</td>
<td>45</td>
<td>739</td>
<td>39</td>
<td>1895</td>
<td>34</td>
</tr>
<tr>
<td>2004</td>
<td>52</td>
<td>816</td>
<td>42</td>
<td>1945</td>
<td>44</td>
</tr>
</tbody>
</table>
From the table above, we see that Africa has the most countries that participated in 2004, yet the number of athletes that were actually sent to the Games is comparatively less than that of other continents. Asia, like Africa, started with no countries at the Games, and as at 2004, sends up to 6 times as many athletes as Africa.

Although national participation has been on the increase since the onset of the Olympics, success at the Games is still concentrated among a few nations. Hoffman et. al. (2004) note that between 1960 and 2000, no Asian or African country, excluding Japan, China and South Korea, have made it as one of the top twenty winning teams. Also, of the 241 nations and territories that attended the Games in 1996, less than half of them had ever won a medal in any Olympic event. Conversely, all African nations combined share a little over 2 percent of all the medals ever bestowed (Johnson and Ali, 2000). Maps 4-6 once again depict the change in medals won at the Games between 1896 and 1996 (please refer to appendix for maps).

An issue even less addressed when analyzing the question of Olympic participation is that concerning the differential representation of male and females at the Games. The specific question then is whether attendance by female athletes is at par with that of their male counterparts. Olympic statistics shows that it is not. There are variations in the number of male and female athletes sent in by different countries. The ratio of female to male participation has contracted considerably from about 0:1 in 1986 to .38:.62 in 2000. Yet, there remains some gap to be bridged before equal representation by both sexes can be attained.

Economists have attempted to analyze what guarantees success and participation at the Olympics. Few have looked into what spurs or constrains male versus female athletes from participating at the Games. The obvious factors can be divided into the structural or Olympic specific factors, economic factors, political factors and religious factors.

II. The debate over the sexes and factors affecting their participation in sports

Structural Factors

The Olympics can be traced as far back as 776BC. Then they were dedicated to Olympian gods. They continued for 12 centuries until they were banned for being a “pagan cult” in 393 A.D. A French man, Pierre de Coubertin, after much thought, decided to revive them. On the 6th of April 1896, the Games were re-opened in Athens. This modern Olympics sought to contribute to “building a peaceful and better world by educating youth through sport practiced without discrimination of any kind, in a spirit of friendship, solidarity and fair play.” Thus the Olympics aimed to serve as a unifying event, intended to transcend nationalities, race, and sex. From the onset, there existed some idealist barriers to this goal of solidarity and fair play. Specifically, Pierre de Coubertin, was biased against women participating in the Games. In 1912, de Coubertin defined the Games as:

“the solemn and periodic exaltation of male athleticism, with internationalism as a base, loyalty as a means, art for its setting and female applause as a reward.” He further believed that “a woman’s glory rightfully came through the number and quality of children she produced, and that as far as sports were concerned, her greatest accomplishment was to encourage her sons to excel rather than seek records for herself” (DeFrantz, 185).
Not surprisingly, the views held by the IOC and the new founder, delayed participation by women in the Games. Women were included in the swimming events for the first time in 1912, after the International Swimming Federation actively promoted women’s involvement. The opportunity to compete in the track and field events was also held off until the 1928 Olympic Games in Amsterdam (DeFrantz, 186). Thus we catch a glimpse of a huge factor which has caused a delay in comparable participation by female and male participants, at the Games, and causes female athletes to play catch-up to male athletes in this 21st century.

The organization of the Olympics is done by an international, non-governmental, non-profit organization called the International Olympic Committee (IOC). The IOC has the primary responsibility of supervising the organization of the Summer and Winter Olympics. The IOC also guarantees the smooth running of the Games in accordance to the Olympic charter. At a national level, National Olympic Committees (NOCs) exist to promote the fundamental principles of the Olympics, within the framework of sports. The NOCs are further responsible for selecting the Olympic teams and participants from their respective nations.

A nation can compete at the Games only after receiving an invitation from the IOC. Although all countries are eligible to participate, a country may not be extended an invitation if they are in violation of the Olympic charter. This rule was applied during the apartheid era in South Africa. Here, we see that the international community may use the Olympics as a way to cry out against a situation in countries. Participation at the Olympics by a country can therefore be hindered if nations leverage the Games to fight a war that they may be unable to on the ground or with guns.

For an event to be included in the program of the Summer Olympic Games, it has to be practiced by men in at least 50 countries and on 3 continents, and by women in at least 35 countries and on 3 continents. Current Olympic Games have approximately 400 events, encompassing its 35 sports, 28 of which are summer sports, while 7 are played in the winter. In the past, certain games including cricket have been dropped from the Olympics. As hosting countries often propose new sports, they and their neighbors or those with similar demographics, may win medals in these sports, giving them an advantage over other participants.

An individual participates in the Olympics only upon approval by an NOC, after being nominated by a national federation. Although these competitors must be adequately prepared for high level competition, the IOC also admits athletes from countries where athletic standards are comparatively lower. This allows all members to participate in the Games (Sutherland 1996). The IOC holds the right to refuse a nominated individual entry into the Olympics, based on its own discretion, without indication of grounds. Its Executive Board also determines the number of all participants at the Games and a limitation is placed on how many participants each country can bring.

An individual who is a national of two or more countries at the same time may represent either one of them. However, having represented one country at the Olympic Games, or in any other competition recognized by the International Federations (IFs), the athlete
must allow a lapse of 3 years before representing another country. There is also no age limit of competitors in the Games, other than that prescribed in the competition rules of an IF, as approved by the IOC Executive Board. This rule has often caused countries to lose their best athletes to other countries, reducing participation and success in the country formerly represented by the athlete, and increasing success at the country the athlete now represents.

**Economic Factors**

The role economics plays in the Olympics has been studied in various forms. Past studies range from cost/benefit analysis of hosting the games, to the assessment of gender equality in national employment as a determinant of a nation’s performance at the Olympics (Klein 2004). Consequently, the individual or the nation participating has to consider the cost of investing in the Games and the returns thereof. Johnson and Ali (2000) noted that the per capita cost of a bronze or silver medal is $1700, while that for the gold is $4750. It is doubtful that individual players will be willing and able to contribute such funds to win a bronze or silver medal at the Games. The price for the gold medal is even more out of the range of most individual participants’ personal income. During Athens 2004, Kenya’s GDP per Capita, measured by purchasing power parity was $1100, while that of the United States was $40,100. Indisputably, the Kenyans would invest fewer resources to the Olympics than the Americans.

Underlying economic theories including that of opportunity cost, utility maximization also surround the Games, but these theories are discussed in-depth in section 4.0 of this paper.

**Political Factors**

The Games have been used as a platform for making political statements, despite the intent of solidarity of the Olympics. In a sense, the portrayal of nations’ grievances on the sport field is not exclusive to the Olympics. Sports have been used as an avenue for nations to foster peace or vent their discontentment. Unfortunately, participants in the Olympics and winners thereof have been forced to suffer the brunt of political pressures. A classic example is seen in the Berlin Olympics in 1936, which Hitler used to further the supremacy of the Aryan race. Due to his animosity towards the Jews, some talented Jews decided not to participate in the Games. Again in 1972, Palestinian terrorists took several Israeli athletes hostage when they penetrated the Munich Olympic village. They intended to force the hand of the Israeli government to release 200+ Palestinian prisoners. The kidnapping turned sour and all the athletes were killed, caught in the middle of a political fight. The Olympics in Berlin and Munich depict how politics has played a role in preventing athletes from participating and perhaps winning in the Games.

Consequently, the Olympics have also served as a platform for signaling significant political events. In 1940 and 1944, no Olympics were held due to World War II. Three Games have experienced widespread boycotts that considerably reduced participation in the Olympics (Bernard and Busse, 2000). These Games were that held in Montreal in 1976, boycotted by many African nations, the Moscow Games in 1980 due to a U.S. led boycott and the Soviet Union’s retaliation of the 1980 Games during the Los Angeles
Games in 1984. These boycotts all led to depletions in participants in certain sports.

Past studies have accounted for these boycotts and other political influences like the ban on South Africa and Libya from the Games, when they failed to conform to Olympic standards (Ali and Johnson 2000). This study will look at political stability as a deterrent to participation and success in the Olympics. World War II prevented two summer Olympic Games. Have any other wars prevented nations from attending the Olympics? This study will observe if political unrest has served to reduce countries’ success and participation at the Games. Perhaps, political instability could explain why neither Sudan, Somalia nor the Democratic Republic of Congo have ever won a medal in the Olympics, and why Vietnam only recently won it’s first medal in the 2000 games in Sydney.

Religious Factors

Another concept, which is historically associated with past and present day Olympics, is religion. From its onset in 776 BC, the Olympics were staged to honor gods in Olympia, Greece. The gods were in turn believed to bestow physical prowess on the athletes. Some believe that this spiritual dimension of the Games have been lost. But has it? Modern day Olympians do not explicitly acknowledge religion as central to the Games, yet the symbolic anthem of the Olympics, along with the cultures and beliefs of the individual athletes are inherent in the modern Games.

At the 2004 Games in Athens, the participation of two women caused stirs in the Olympic Games. These women were Rubab Raza, the first woman to swim for Pakistan and Friba Razayee, the first woman in history to compete for Afghanistan at the Games. Afterwards, the BBC sent out an opinion poll on why the participation of Muslim women in the Olympics had increased. Responses were varied. Some argued that in the past, Muslim women tended to pick family over sports.Still others stated that Islam teaches proper dress codes. A sport falling within the prohibited dress code was then not allowed. Apart from the dress codes for Islam, other religions may hold parallel practices, which may prevent their participation and consequently success at the Olympics. Again, it may be that it is governmental decisions coupled with religious practices discouraging grass root participation in sports, which is then reflected in a lack of participation by nations at the international level.

In recent times, we also see the Israeli plight. Ever since Israel’s founding, some Muslim nations have refused to compete against the Jewish state. In the Olympic Games held in Athens in 2004, an Iranian judo champion quit the Olympics entirely instead of competing against an Israeli. His one-man boycott earned him encomiums from his President - Mohammed Khatami. This boycott is however not an isolated incidence. Israel continues to see animosity towards her being played on the sports field. Even the mentally impaired have been subjected to this treatment, as viewed during the Special Olympics held in Ireland in 2003. The International Olympic Committee (IOC) is yet to reprove these acts. In the meantime, a handful of potential Olympic medalists and participants fail to prove themselves competitively. In this study, the impact religion has in determining success and participation in the Olympics will be analyzed. Having provided insight into some of the issues surrounding modern day Olympics, we now look at studies to access
the contributions of the academics to observations made.

III. Literature Review

Economists have in the past investigated factors that affect participation and success at the Olympics. These economists define success by the number of medals won by a country, while participation is collated on the national level. Research carried out specifically on economic, political and religious influences on success and participation however varies across studies. The influence of economic institutions has mainly been captured by GDP per capita, and has produced consistent, positive and significant results. The influence of political institutions is frequently equated with the influence of government types. One of such measures assesses the advantages incurred by being a socialist country. Political proxies like this socialist country factor have also proved significant in determining success and participation at the Games. However, a study on religious institutions in the context of the Olympics is yet to be found.

Studies on Olympic success date as far back as the 70s. Examples of such studies are that of Levine (1972) and Ball (1972). These past studies mainly used correlations to investigate the factors affecting Olympic performance. Levine’s (1972) study found that Olympic success is highly correlated with a number of economic, demographic, social and political variables. However, GDP was the major determinant of Olympic success. Between the 70s and the 90s, literature on Olympic success and participation became non-existent. Kuper and Sterken (2003) speculate that this could be due to be disturbances in the Games caused by the Cold War. In recent years, there have been further studies on Olympic success by Bernard and Busse (2004), Tcha (2004), and participation Johnson and Ali (2000), at the Olympics. The most referenced papers are that of Bernard and Busse (2004) and Johnson and Ali (2000). These two papers, along with that of Tcha (2004) will guide my study.

Johnson and Ali (2000) considered factors affecting participation and success at the Games, and further highlighted differing impacts of their explanatory variables on male and female athletes. In their study, the ordinary least squares regressions were used to carry out their analysis. On the other hand, Bernard and Busse (2004) considered only the factors affecting medal success, using a Cobb-Douglas function in their study. The Cobb-Douglas was used based on the assumption that countries “produce” medals, but the relationship between the explanatory variables and medal wins were not linear. Johnson and Ali (2000) base their analysis on 12 Games, from 1952 to 1996, while Bernard and Busse (2004) restricted their observations on 10 Games, specifically from 1960 to 1996. The studies by Johnson and Ali (2000) and Bernard and Busse (2004) show consistent results.

Both Bernard and Busse (2000) and Johnson and Ali (2000, 2002) used GDP per capita as a measure of national wealth and productive capacity. It reflects the nation’s ability not only to train its athletes, but also to send them to the Games. GDP per capita also serves as a measure of resources per person. Both studies showed GDP per capita to have a positive and significant influence on success and participation. Bernard and Busse (2004) concluded that if a country were to double its total GDP, it could expect an increase
by 1-1.5% of the total medals won. Johnson and Ali (2000) also found that GDP per capita played a significant role in determining participation, specifically finding that richer nations send approximately 8 more athletes at the average for each thousand dollars of GDP per capita. I base the hypotheses of my study on these results.

Johnson and Ali (2000) considered the effects of political systems on success and participation at the Olympics. They specifically accounted for a state of democracy versus other government systems including monarchy, single party or communist and military. They find that political systems have little effect on participation. However, monarchies and “other political systems” (explained as those in transition to democracy or with ill-defined power structures) send fewer athletes than democracies. Conversely, military, single-party and communists send no more than is expected. Surprisingly, single party and communist systems showed a positive and significant effect on success rates by tending to win more medals on average.

Though Olympics has a religious foundation, dating back to the early games when they were dedicated to the Olympian gods in Greece, I did not find a study which shows the influence of religious institutions on the Olympics, particularly on its success and participation. A study by Barro and McCleary (2003), however, not specific to the Olympics provides insight to the influence of religion on economic growth.

Barro and McCleary (2003) used a cross-country panel data set that included country level information on church attendance and religious beliefs. The data was derived from individual information collected in six international surveys between 1981 and 1999. In their study, they find that religiosity tended to decline overall with economic development, but varied depending on the specific dimensions of development. For instance they noted that measures of religiosity are positively related to education, and the presence of children, yet negatively related to urbanization. They further observed that religiosity was negatively related with government regulation of the religion market and with religious oppression that accompanied the presence of a Communist government. While Barro and McCleary’s study is not directly related to the study of success and participation in the Olympics, they provide a framework on how to explore issues of religiosity within the context of economics.

Having considered the economic, political and religious institutions, other factors remain that affect participation and success of countries at the Olympics. One of the major factors considered is population, based on the rationale that talent is randomly distributed. Consequently, a country with a large population has a wider pool of talented, prospective Olympic athletes. Population had positive and significant coefficients. In Bernard and Busse’s (2004) econometric analysis, they showed that if a country were to double its population, it would win an additional 1.5% of available medals.

Host country effects and lagged medal shares were also considered. Host countries were argued to have advantages stemming from a reduced cost of Game attendance, their home crowds ability to sway judges, the tailoring of facilities to meet national athlete’s needs and their being influential in the addition of new sports to the Games. The host country advantage was found to be significant, and increased the likelihood of winning medals.
Lagged medal share, which depicted past medal successes also proved to have a significant effect on the probability of winning a medal in the current Games.

While Johnson and Ali (2000) considered political systems, Bernard and Busse (2004) control for political influences, specifically in the form of large-scale boycotts and the medal premium enjoyed by the former Soviet Union and its satellites. The 1980 and 1984 Olympics were subject to boycotts by a number of countries. Johnson and Ali found that the host country tended to benefit more from the boycotts, receiving as many as 19 percentage points compared to other participating countries. This is to be expected as the boycotts led to a significant decrease in competitors in certain fields, making more medals available to be won by the remaining competitors, which in most cases were the host countries.

The Soviet Union and its satellites were also seen to enjoy a premium in the Olympics. As seen by Bernard and Busse, the Soviet Union and Eastern Bloc countries won a share of medals higher by 3+ percentage points than predicted by their GDP. The increased win shows that the size of a nation’s sports budget, as well as the intensity and collective efforts made to win more medals will undoubtedly affect a country’s performance at the Games (Tcha and Pershin 2002). Other studies like that of Kuper and Sterken (2003), Hoffman, Ging and Ramasamy (2004) have carried out parallel analysis to that of Bernard and Busse and Johnson and Ali, and have reached similar conclusions.

Tcha and Pershin (2002) used an unconventional approach to investigating factors affecting success at the Olympics. Analyzing 66 of the nations which participated in the Seoul (1992), Barcelona (1992) and Atlanta (1996) Games, they highlight the issue of specialization, and hypothesized that each country will concentrate on the sport in which they have comparative advantage, in a bid to increase its medal wins.

In addition to variables such as population and a dummy for the socialist countries used by Johnson and Ali (2000) and Bernard and Busse (2004), Tcha and Pershin (2002) considered environmental, physical and biological attributes. These variables were seen to give countries comparative advantages, thereby leading to increased performance. For instance they hypothesized that performance in beach volley ball would be related to the temperature and the length of the coastline in a country.

Tcha and Pershin’s (2003) study differed from that of Johnson and Ali (2000) and Bernard and Busse (2004) in that they looked at country participation in specific types of sports versus an overview of total participation at the Games by individual countries. Tcha and Pershin (2003) result that GDP and population were not found to be significant explanatory factors across all sports that were considered. They were however significant in at least one group of sports. For instance countries with RCA in athletics were characterized with high GDP per capita in addition to being relatively endowed with landmass, high altitude, but relatively less well endowed coastlines. Total population and former/current socialist countries significantly affected the RCA indexes in ball games and gymnastics. Tcha and Pershin concluded that countries with higher GDP per capita tend to partially specialize i.e. collect medals from a more diversified range of sports.

In the studies of Johnson and Ali (2000), Bernard and Busse (2004) and Tcha and
Pershin (2003), only Johnson and Ali extended their study to analyze the factors affecting participation. However, in all three studies, political factors were measured in terms of governmental structures. None of these studies highlighted the probable cause of political unrest and instability within the countries affecting participation and in turn success at the Games. Religion is also not considered as a potential factor that could affect participation and success at the Games. All three studies were also carried out without analyzing the Sydney 2000 and Athens 2004 Olympics, which my study covers.

As briefly mentioned in section 2.0, several economic theories have been used in the study of Olympic participation and success. These theories will be discussed in the next section.

IV. Economic Theories Used in Past Studies

Fundamental economic theories have been employed by economists in the study of Olympic participation and success. These theories include the concept of utility maximization & uncertainty, neoclassical trade models applied to revealed comparative advantage. Another underling theme proposed is that of opportunity cost.

The Concept of Utility Maximization and Uncertainty

Utility maximization and its risk components is another concept used to measure success. Tcha (2004) explores how a country can maximize its utility from the Olympic Games when uncertainty exists in medal production. Uncertainty in production of medals is here described as the possibility that a country fails to win medals in spite of its investments in time, training and funds. Tcha (2004) also highlights other factors which determine success. These factors include health, demography, location, and the player’s physical and mental condition. Tcha (2004) explains that uncertainty is considered to be very high when the decision maker, that is the participants, does not have sufficient information or knowledge about the game, place and the strength of the competition.

To this end, assuming all players (athletes and nations) are rational thinkers, how can they maximize their utility in the Olympic Games based on the amount of resources they have available and given the choices that constrain them? Decisively, the goal of the Olympic is to participate, and do one’s best, however, it can be assumed that a player will be most satisfied winning a medal.

Tcha concludes that as the overall welfare level and in turn, investment in sports increase, the related risk premium decreases which reduces performance gaps between countries. Furthermore, the study determined that the relative performance of the (eastern bloc) country is expected to improve, given that the representative consumer (eastern bloc) is more risk averse, uncertainty in medal production increases and the cost of medal production in the other (Western bloc) country decreases.

The concept of utility maximization and uncertainty shows how economists, politicians and policy makers determine what events are most realistic for their countries to participate in. It also shows if countries are more risk averse by competing outside their areas of specialization, or more risk seeking by frequently participating in events in which they never maximize their utility by winning a medal. It is important to note that a lot of the
decisions made will be subject to constraints such as economics, demographics, health, and political situations characteristic of the individual countries.

**Revealed Comparative Advantage**

Tcha and Pershin (2002) introduced the concept of revealed comparative advantage (RCA) and specialization, in analyzing performance at the Summer Olympics. They specifically apply neoclassical trade theories which predict that trade leads a country into specialization, to discuss the issues related to RCA. Tcha and Pershin (2002) note that a country will therefore specialize, according to the size of its relative factor endowment and the factor intensity of its commodities. Tcha and Pershin (2002) state that the concept of RCA pertains to the relative trade performances of individual countries in particular commodities, and is based on the assumption that the commodity pattern of trade reflects inter-country differences in relative costs, as well as in non-price factors.

In applying the concept of RCA to the performance of countries at the Olympics, they highlight underlying economic ideas. The first is iterating limitation of resources, specifically that each country participating at the Olympic Games is endowed with a limited amount of resources to invest in sports competition. The second is the issue of choice – each country concentrates on the sport in which it has a comparative advantage. They give the following example to explain this point:

“In a simple two-factor (capital and labor) model, that a developed country with a relatively large supply of capital but a small population would specialize in capital-intensive sports, such as yachting. In contrast, a poor country with a relatively low level of capital would specialize in those sports where capital is relatively less important (or labor-intensive), say marathon running or boxing” (Tcha and Pershin, 2002).

In order to understand the RCA model proposed by Tcha and Pershin (2002), four assumptions adopted in the neoclassical trade models were stated:

i) Medals of the same color provide the same level of utility

ii) Each country uses its resources to produce medals

iii) Countries share homogenous tastes and the production function of medals is the same in all countries

iv) Factor intensity reversal in medal production does not occur

The RCA index which Tcha and Pershin (2002) provide becomes their proxy for specialization, for country i in sport j as calculated as

\[ R_{ij} = \frac{M_{ij}}{T_j} / \frac{M_i}{T} \]

where \( M_{ij} \) is the total number of medals country I wins from sport j, \( M_i \) represents the number of medals won by country i in the sports at the Summer Olympics, \( T_j \) is that assigned to sport j, and \( T \) is the total medals won by country I in all sports at the Summer Olympics. If \( R_{ij} > 1 \), it implies that this country performs relatively better in sport j than it is supposed to. That is, it has RCA in sport j and vice versa. Therefore, the country, which has neither RCA nor revealed comparative disadvantage (RCD) in any sport will have \( R_j = 1 \) for all \( j = 1, 2, \ldots, N \).

The determinants of RCA used were geographical, biological and economic variables. Tcha and Pershin concluded that countries with higher GDP per capita tend to partially
specialize i.e. collect medals from a more diversified range of sports.

The concept of RCA provides insight to understanding success at the Olympics. The theory can be even further developed to explain participation at the Games. As a shortcoming to their model, Tcha and Pershin (2002) identify the need for better proxies to account for unquantifiable factors. For instance, to account for physical and biological attributes, a dummy for Asians and another for Africans was used. Tcha and Pershin (2002) express the need for a better analysis of more detailed variables such as weight, and height. But, information on better physiological variables is scarce. Other theories such as that of opportunity cost and first-mover advantages also cause RCA to be insufficient in completely explaining performance.

**The Concept of Opportunity Cost**

A fundamental theory of economics is that of opportunity cost. Individuals and nations are faced with choices about how many resources they should allocate to the Games. They have to make a choice. Assuming then that the medals available to be won are the goods, and the Games are the market place, how many medals, if any, should a participant (the consumer) be willing to purchase at any given price. The decision then to participate and win in the Olympics, given a limited amount of resources, depends on how much the player values the goods.

Decidedly as inherent in the concept of opportunity cost is the idea of choice. Assuming then a two product economy, where all that is produced are medals and other goods, we can expect that a country with a higher GDP per capita will invest more resources to the manufacturing of medals than a country with a lower GDP per capita.

**V. Data**

This paper analyzes 14 summer Olympic Games from 1952 Olympics held in Helsinki to that in Athens in 2004. Analysis on all Summer Olympics from that held in Athens in 1896 until that in 2004 was not done due to constraints posed by data availability and changes in the world economy between the start of the modern day Olympics and now. The summer Games were chosen over the winter Games due to greater participation by countries in the Summer Olympics. In 2002, only 77 countries (2399 athletes - 886 women, 1513 men) were represented at the Winter Olympic Games held in Salt Lake City, U.S.A, while 199 NOCs participated in the 2000 Sydney Olympics (10651 athletes - 4069 women, 6582 men). The Summer Olympics are also less subjected to advantages of climate, where many developing countries do not have weather conditions that would be favorable to train for winter specific sports.

Data on Olympic specific information including medals won, total number of participants, host countries, was obtained from the Olympic Games’ official website and other secondary sources on Olympic history.

For GDP per capita, Angus Maddison’s 1990 international Geary-Khamis dollars was used for statistics between 1952 and 2000. Information on missing data was retrieved from the population division of the U.S. Census Bureau’s International Program Center, where the figures were given in constant 2000 dollars. GDP per capita is measured in unitary
Data measuring political factors were obtained from two main data bases. For polity, information was gotten from the Polity IV project, a Polity research which codes the authority characteristics of states in the world system for purposes of comparative, quantitative analysis. Information on social unrest and political instability were gotten from the State Failure Problem Set dataset that accesses internal war and failures of governance between 1955 and 2000. Soviet Union and planned economies were classified based on the identifications given by Bernard and Busse (2004).

The data on religion was obtained from the Word Fact Book, also known as the CIA Fact Book. Here, the percentage of each nation’s population that represented a certain religion was given. Additional information was filled in from information on nation’s websites, where the percentages were not specific. The Muslim and Dharmist religious variables were added in the regressions, resulting in the MUSHDA variable. The Dharmist religions included Buddhism, Taoism, Confucian and Hinduism. Other religions included all other specified religions like Christianity, indigenous and traditional beliefs, and spiritualists, atheism, unspecified and unaffiliated religions.

The main problem which had to be reconciled with the data was dealing with countries which had been joined and then separated between 1952 and 2004. These countries were Germany, Czechoslovakia, USSR and Yugoslavia. With Germany it was difficult to find GDP per capita information for all the years when the country was divided into West and East Germany. Angus Maddison’s (2001) however provided information for some two of the years. For the other countries, information on participants and medals were non-existent for the years when the countries were unified; therefore it was easy to eliminate the observations in those years. For instance, no participation or medals were available for Russia between 1952 and 1988. All medals and participants were thus recorded under the Soviet Union. From the 1992 Games, the medals and participation were then allocated to the respective countries, including Russia. The dummies on the Soviet Union and planned economies also help to prevent the inflation of results due to the presence of these joint participants.

VI. Empirical Analysis

Though theories of utility maximization, comparative advantage and opportunity cost apply in different ways to this study, it is difficult to identify one of the three as the overarching theory of this study. This stems from the multi-faceted aspects of the explanatory variables that would be used. To this end, the rest of the paper identifies the empirical analysis and models that lead to the conclusion of this paper.

Analytical Model

To test the impact of economic, political and religious institutions on participation and success in the Olympics, Ordinary Least Squares (OLS) regressions will be used. This OLS method is similar to that employed by Johnson and Ali (2000) and Bernard and Busse (2004). This method was chosen, as it would show the magnitude and significance of the impact that each of the explanatory variables have on participation and success. The
regression results would also show the difference that these effects have on male versus female participants at the Olympics. Three hypotheses will be tested.

Hypothesis 1: Economic wealth reflected in a nation’s income level, stage of development, and availability of resources, have a significant influence on participation by nations and national success at the Olympics.

Hypothesis 2: Political institutions as reflective of the system of government, the control of the national resources and level of unrest and instability within the country influence participation by nations and national success at the Olympics.

Hypothesis 3: Religious inclinations as reflective of the culture of a country influences participation by nations and national success at the Olympics; and have a greater influence on female than male participation.

The Cobb-Douglas form, which assumes a non-linear relationship between the dependent and explanatory variables, was considered over the OLS but decided against. Bian (2004) used the Cobb-Douglas to estimate the diminishing marginal return of both population size and per capita GDP on success at the Olympics. The existence of multicollinearity between GDP and population required the Cobb-Douglas function to be converted into one, which assumes constant returns to scale. This is specific to Bian’s (2004) study, as Bernard and Busse (2004) also used the Cobb-Douglas form without reporting the need for the imposition of restrictions. However, after using their Cobb-Douglas based model to predict future Olympic medals, they conclude that:

“The model underpredicts medal shares at both the low and high ends of the range and overpredicts in the middle. Although the additions of log GDP per capita and several dummies have improved the fit substantially, the current model is lacking in overall predictive power” (Bernard and Busse, 415).

OLS thus presents a less-restrictive model that is better for the analysis of this study. OLS will also provide a good comparative basis for my study, compared with previous ones, as it is the functional form most used in studies measuring factors affecting Olympic participation and success. That is, I can better compare the results I find of variables used in other studies and observe the effects of the introduction of the political and religious variables, which have previously not been measured.

To test hypotheses 1-3 above, four main regressions are carried out. The dependent variables are total number of participants, total medals won, total number of men athletes, and total number of female participants per nation. Participation as a dependent variable is measured as the sum total of all athletes sent to the Olympic Games by each nation. In order to capture the difference in male and female participation in the Olympics, I will use the number of men and that of women sent to the Games as the dependent variables. Success is also measured in absolute terms as the sum total of all the medals won by each nation.

The final empirical analysis therefore takes the following form:
Regression 1

$$PTOTAL = \beta_0 + \beta_1 GDP + \sum \beta_2 Popula + \sum \beta_3 Mushda + \beta_4 Polity + \sum \beta_5 Soviet + \sum \beta_6 Planned + \sum \beta_7 Ethnicwar + \beta_8 Regimech + \sum \beta_9 Poliwar + \sum \beta_{10} Genocide + \beta_{11} Host + \sum \beta_{12} Boy1976 + \sum \beta_{13} Boy1980 + \sum \beta_{14} Boy1984 + \beta_{15} Time + \mu$$

All four regressions follow this model, as the explanatory variables remain the same. The difference exists among the dependent variables, which are summarized in Table 1 below:

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Definition</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Participants</td>
<td>Total number of athletes sent by country i, during a particular Olympic year</td>
<td>IOC</td>
</tr>
<tr>
<td>Total Medals</td>
<td>The number of medals won by country i during a particular Olympic year</td>
<td>IOC &amp; Wallechinsky</td>
</tr>
<tr>
<td>Female Participants</td>
<td>Total number of female athletes sent by country i, during a particular Olympic year</td>
<td>IOC</td>
</tr>
<tr>
<td>(Ptotal)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(TMEDAL)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(PMEN)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(PWOMEN)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Description of Explanatory Variables**

**Economic Factors**

The Gross Domestic Product per Capita (GDP per capita) will be used as a measure of the extent to which a country’s economic resources and income affect success and participation in the Olympics. GDP per capita measures the productive capacity of the nation, the wealth of its people, its level of development, and specifically reflects the average individual’s lifestyle and standard of living.

$\beta_1$, the coefficient on GDP per capita is expected to be positive for all regressions. This will indicate that countries with higher incomes tend to send more athletes to the Games and win more medals at the Olympics than those with lower incomes. The rationale is that a higher income country can invest more time, resources and training prior to sending their athletes to the Olympics.

Angus Maddison’s 1990 international Geary-Khamis dollars is used for statistics between 1952 and 2000. Missing data were retrieved from the population division of the U.S. Census Bureau’s International Program Center, where the figures were given in constant 2000 dollars. GDP per capita is measured in unitary terms.

**Political Factors**

To capture the extent to which political instability and transitions affect success and participation in the Olympics, three sets of variables are used. The first looks at the effect of centrally planned governments, specifically the Soviet Union and other socialist governments. To capture this effect, one dummy variable is created for countries that are directly under the sphere of the old Soviet Union. These countries include Bulgaria, Czechoslovakia, Poland, the USSR, East Germany, Hungary and Romania from 1960 to
1988; the Unified Team in 1992; and Cuba from 1952 to 2004. Another dummy is created for other socialist, typically communist countries and these include China, Albania, Yugoslavia from 1952 to 1988 and North Korea. For the Soviet and planned countries, the dummy variables assigned take the following form:

- \( \text{Soviet}_i = 1 \) if country \( i \) is a country in the Soviet Union
- \( \text{Soviet}_i = 0 \) otherwise, and
- \( \text{Planned}_i = 1 \) if country \( i \) has a planned economy
- \( \text{Planned}_i = 0 \) otherwise.

These dummies provide insight into the basic economic question of “who to produce?” I expect countries in these centrally planned governments to send more athletes and win more medals than the democratic nations. It is however unclear what the sign on these coefficients would be with respect to male and female participation.

The second political measurement uses “polity” as the unit of analysis. “Webster’s New World College Dictionary defines a "polity" as a "political or governmental organization; a society or institution with an organized government; state; body politic" (Polity IV Project, 1). Subtracting an indicator of autocracy from that of democracy gets the weight of polity. The resulting unified polity score ranges on a scale from +10 (strongly democratic) to -10 (strongly autocratic). The Polity IV project, a Polity research that codes the authority characteristics of states in the world system for purposes of comparative, quantitative analysis, is the basis of polity measurement used. The polity unit for the particular year of the Olympics is what is accounted for. The resulting coefficient on polity is ambiguous. It is uncertain whether the more democratic governments will send more participants to the Games or win more medals. Also, whether or not governments contribute to the differential participation of male and female athletes in international level sports is uncertain.

The third measure of political institutions assesses social unrest and political instability. The effects of ethnic wars, political wars, adverse regime changes and genocide/politicide are considered to measure the extent to which political instability affects Olympic participation by nations and gender, and also Olympic success. Ethnic wars, political wars and adverse regime changes are estimated to affect Olympic participation and success if they occur in the year of the Olympics or the year before and are measured on a scale of 1-4. Genocide/politicide are estimated to affect Olympic participation and success if they occur two years or a year before the Olympics and the year of the Olympics, and are measured on a scale of 0-5. A higher value on the scales of either 1-4 or 0-5 depicts a worse situation. Consequently, the coefficients on ethnic wars, political wars, adverse regime changes and genocides are expected to be negative.

As questions may arise as to what constitutes each of these measures of instability, the definitions are therefore given as:

i) Ethnic wars: Episodes of violent conflict between governments and national, ethnic, religious, or other communal minorities (ethnic challengers) in which the challengers seek major changes in their status.

ii) Revolutionary wars: Episodes of violent conflict between governments and politi-
cally organized groups (political challengers) that seek to overthrow the central government, to replace the leaders, or to seize power in one region.

iii) Adverse Regime Changes: Major, adverse shifts in patterns of governance.

iv) Genocide: These involve the promotion, execution, and/or implied consent of sustained policies by governing elites or their agents --- or in the case of civil war, either of the contending authorities -- that results in the deaths of a substantial portion of a communal group or politicized non-communal group. (Marshall et. al, 5, 6, 10, 15)

**Religious Factors**

The effect of religiosity on success and participation in the Olympics will be studied in this paper. The CIA fact book provides information on the percentage of each population that is of a particular religion. The end result will be in the form – country i has 40% Christians, 30% Muslims, 20% indigenous and 10% indigenous beliefs, 10% undeclared; national religion- Islam.

As yet there are no known studies on the correlation between religion, participation and success at the Olympics. The end result is therefore uncertain. The hypothesis to be tested in this study is that religions with conservative cultures show less propensity to participate, and consequently record less success at the Olympics. Also, the impact of religion is expected to influence female athletes more negatively than their male counterparts.

**Other Factors**

A significant influence on participation and success in the Olympics has been population. It is assumed that talent is normally distributed among countries. Thus a country with a larger population will tend to have a wider pool of competent athletes to represent it at the Olympics. \( â_4 \), the coefficient on population, is expected to be positive and significant. In essence, the bigger the population, the larger the number of athletes the country will send to the Olympics, and vice versa. Also, it is expected that the larger the country’s population, the more likely it is to participate in team sports, thus increasing the number of medals it could win. For example, India with a population of 1,080,264,388 (July 2005 est.) has won a total of 175 medals. However, only 6 of these are attributed to individual accomplishments, while 169 medals are received from the 11 times that its men’s hockey team has received a medal.

The host country advantage cannot be underestimated. For the host country, there is reduced cost of travel to and from the Olympic Games. Again, the athletes play in front of a home crowd, which motivates them to compete better, making them more likely to succeed. Also host countries with the approval of the International Organization Committee are allowed to introduce new events to the Olympics in which they are more likely to succeed. For instance in the Games held in Tokyo in 1964, judo was introduced while mountain biking, beach volley-ball and softball were introduced at the Atlanta Games in 1996. These new events are largely reflective of the countries where the Games were being held.

A dummy variable is also included for all 14 host nations that would be considered in this study. For instance for 1952, the variable would be 1 for Finland for the Helsinki games, 0 if not, 1 for Australia for the Melbourne/Stockholm games in 1956, 0 if not. The coefficient on this variable is expected to be positive and significant. This result will be
consistent with past studies that have been done. While certain studies regard host coun-
try advantage merely as a factor to explain increased successes, others used them as a sig-
nal of the sports culture within the nations. Thus, a country, which put up a bid to host the
Olympics, indicates that it is more likely to value sports more than those who do not make
attempts to host it.

In this model, time was also accounted for. Each year of the Olympic had a number
allocated to it. For instance 1 for 1952, 2 for 1956 and so on. The coefficient on time is
expected to be positive and to show the changes that have occurred in success and partic-
ipation over the 14 years.

Finally, a disturbance term, µ, will be included to complete the econometric model.
The regressions for participation and success are very similar although they will each be
symbolic in different ways. The table below

<table>
<thead>
<tr>
<th>Table 2: Descriptive Statistics of Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>PTOTAL</td>
</tr>
<tr>
<td>TMEDAL</td>
</tr>
<tr>
<td>PMEN</td>
</tr>
<tr>
<td>PWOMEN</td>
</tr>
<tr>
<td>GDPCAP</td>
</tr>
<tr>
<td>POPULA</td>
</tr>
<tr>
<td>MUSHDA</td>
</tr>
<tr>
<td>POLITY</td>
</tr>
<tr>
<td>ETHNICWAR</td>
</tr>
<tr>
<td>REGIMECH</td>
</tr>
<tr>
<td>POLIWAR</td>
</tr>
<tr>
<td>GENOCIDE</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
</tr>
</tbody>
</table>

Critique of the Model
The regression models in this study exempt many potential explanatory factors of suc-
cess and participation in the Olympics. In light of the economic theory of utility maximiza-
tion, it is important to note that different players are satisfied at varying levels. That is, cer-
tain nations are satisfied with participating and not winning, while others seek to win gold.
Regardless of factors such as GDP per capita, population, and host countries, certain vari-
ables would better reflect the efforts made by a country to attain ultimate satisfaction, that
is, medals. These include percentage of national GDP allocated to sports by a country. This
variable will signify local and national level efforts made to be better sports competitors.
It would show how much is spent in local gyms, how many training facilities exist in the
nation and ultimately how much is spent on recruiting, training and sending the athletes to
the Olympics. However, this information will prove quite difficult to get, especially for
many of the developing countries.

Again, the effect of player transference at the international level may affect the partic-
ipation and success by nations in the Olympics. This situation arises when athletes com-
pete for countries other than that of their origin. The Olympic Games has seen this increase in player transference in recent years. While certain countries are reaping this benefit, others like the Cuban government view it as “the trafficking of people… the theft of athletes,” where rich countries are robbing the best athletes from the poor ones (The Guardian, 2000). However, this transfer of athletes due to larger incentives, though interesting, is outside the scope of this paper.

Other potential factors that could have been considered are health factors. Given that a nation with a healthy population is more likely to have more people to send to the games. Again, massive natural disasters like political instability and unrest could act as a deterrent to a country’s participation and success at the Games.

VII. Results and Discussion of Results

One of the criticisms mentioned by Tcha and Pershin (2002) was the fact that past studies analyzing success and participation at the Olympics failed to account for the possibility of heteroscedasticity. To this end, a white test was performed before the final regressions for this study were run, and heteroscedasticity was found to be present. It was treated by running weighted least squares regressions. The reported results are therefore those which have been corrected for heteroscedasticity.

Also F-tests were run to decide whether or not to add the Muslim and Dharmist explanatory variables. Based on the conclusions of the test, the two variables were included resulting in the MUSHDA variable. The religious analysis are therefore made comparing Muslim and Dharmist religions to others including Christianity and indigenous.

Results for Regressions 1 & 2

Table 3 below summarizes the results of regressing total participants and medals won at the Olympics on the explanatory variables.

As expected, the coefficients on GDP Per Capita and population are positive and highly significant at a 99.9% level for both participation and medals. For every $1000 increase in a country’s GDP Per Capita, a country can send five more participants to the Olympic Games and win 3/5 more of a medal. Also, a country sends 22.5 participants to the Games and wins 3.62 additional medals with every 100 million increase in its population. The host country coefficients, depicting the home field advantage, is also positive and highly significant, showing that countries which hosted the Olympics send approximately 291 more athletes to the Games, and win 35 more medals than their non-hosting fellow competitors.

The coefficients on the boycott dummies are negative for both participation and medals won. For total participation, only the 1980 Olympics boycott is significant, and that at a 95% level for total participation, showing that 10 less participants were sent to the Games as a result of the boycott. On the other hand, the 1976 and 1984 boycott years were both significant for total medals won, at a 90% level. Countries tended to win fewer medals in these years than in the non-boycott years.

The Soviet bloc coefficient is also positive and highly significant for both participation and total medals at a 99.9% level. Countries within the Soviet bloc send up to 97 more par-
participants, winning 13 more medals than the non-Soviet countries. On the other hand, the planned economy dummy coefficient is negative but also highly significant. Countries with planned economies tended to send approximately 37 fewer participants to the Games, winning 4 less medals than other non-planned economies. The positive value for the Soviet nations is consistent with that of other nations, but a negative coefficient on the planned nations dummy was unexpected.

The coefficient on the Muslim/Dharmist combined variable is negative and highly significant at a 99.9% level. For every 10% increase of Muslims or Dharmists in a country, 2 fewer participants are sent to the Games. Also with a 10% increase of Muslims and Dharmists, the nation reduces wins by .28 fewer medals compared to Christian and indigenous religions.

<table>
<thead>
<tr>
<th>Table 3: Summary table of results from regressions 1 and 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variable</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>CONS</td>
</tr>
<tr>
<td>GDPCAP</td>
</tr>
<tr>
<td>POPULA</td>
</tr>
<tr>
<td>MUSHDA</td>
</tr>
<tr>
<td>POLITY</td>
</tr>
<tr>
<td>SOVIET</td>
</tr>
<tr>
<td>PLANNED</td>
</tr>
<tr>
<td>ETHNICWA</td>
</tr>
<tr>
<td>REGIMECH</td>
</tr>
<tr>
<td>POLIWAR</td>
</tr>
<tr>
<td>GENOCIDE</td>
</tr>
<tr>
<td>HOST</td>
</tr>
<tr>
<td>BOY1976</td>
</tr>
<tr>
<td>BOY1980</td>
</tr>
<tr>
<td>BOY1984</td>
</tr>
<tr>
<td>TIME</td>
</tr>
<tr>
<td>R^2</td>
</tr>
<tr>
<td>Number of observations</td>
</tr>
</tbody>
</table>

Notes: Significance indicated as * for 90 percent, ** for 95% and *** for 99.9%

The coefficient on polity is also positive for participation at a 99.9% level. For medals, the polity coefficient though negative is insignificant. A one-unit increase towards democracy in a country increases the number of participants it would send by .8 but does not suffice in determining how many medals a nation will win. The insignificance of polity as a measure for determining medal success could result from the fact that success at the Games is better explained by individual-specific variables.

The state failure coefficients are all negative. However, only those of ethnic war and
genocide are significant at a 90% and 95% level consecutively for participation. For medals won, genocide is the only significant variable at a 99.9% level. A country going through ethnic war during the year of the Olympic Games or the year before hand will send about 4 less participants to the Games. Also, a country experiencing genocide in the year of the Games or for two years before the Games will send approximately 7 less participants to the Games and win 1.5 less medals.

Finally, the coefficient on time is positive and significant at a 95% level for participation, showing that each Olympic year, there is a .692 increase in the participants at the Olympic Games. The time variable, though positive, is not significant in explaining total medals won.

Results for Regressions 3 & 4

Table 4: Summary table of results from regressions 3 and 4

<table>
<thead>
<tr>
<th>Variable</th>
<th>PMEN</th>
<th>PWOMEN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>t-stat.</td>
</tr>
<tr>
<td>CONS</td>
<td>17.53795</td>
<td>7.598</td>
</tr>
<tr>
<td>GDP CAP</td>
<td>0.00364</td>
<td>10.302</td>
</tr>
<tr>
<td>POPULA</td>
<td>1.34E-07</td>
<td>8.094</td>
</tr>
<tr>
<td>MUSHDA</td>
<td>-0.14307</td>
<td>-6.704</td>
</tr>
<tr>
<td>POLITY</td>
<td>0.794702</td>
<td>4.062</td>
</tr>
<tr>
<td>SOVIET</td>
<td>72.26571</td>
<td>10.213</td>
</tr>
<tr>
<td>PLANNED</td>
<td>-31.4597</td>
<td>-4.122</td>
</tr>
<tr>
<td>ETHNICWA</td>
<td>-3.19569</td>
<td>-2.138</td>
</tr>
<tr>
<td>REGIMECH</td>
<td>-0.94528</td>
<td>-0.394</td>
</tr>
<tr>
<td>POLIWAR</td>
<td>-0.50528</td>
<td>-0.403</td>
</tr>
<tr>
<td>GENOCIDE</td>
<td>-3.92029</td>
<td>-2.003</td>
</tr>
<tr>
<td>HOST</td>
<td>196.3589</td>
<td>13.024</td>
</tr>
<tr>
<td>BOY 1976</td>
<td>-3.69793</td>
<td>-1.184</td>
</tr>
<tr>
<td>BOY 1980</td>
<td>-4.79465</td>
<td>-1.918</td>
</tr>
<tr>
<td>BOY 1984</td>
<td>-0.68617</td>
<td>-0.206</td>
</tr>
<tr>
<td>TIME</td>
<td>-0.41899</td>
<td>-1.838</td>
</tr>
<tr>
<td>R^2</td>
<td>0.4228</td>
<td>0.4441</td>
</tr>
<tr>
<td>Number of observations</td>
<td>2189</td>
<td>2189</td>
</tr>
</tbody>
</table>

Notes: Significance indicated as * for 90 percent, ** for 95% and *** for 99.9%

The coefficients on GDP per Capita, population and host are all positive and significant at a 99.9% level. However, for a $1,000 increase in GDP per capita, a country sends 3.6 more male athletes, and 1.9 extra female athletes. Also, a 10 million increase in total population would ensure 1.34 male athletes are sent to the Games while .911 female athletes will be sent. The host nations also send 196 more male athletes, which is more than twice the extra 95 female athletes that are sent to the Games.
The boycott coefficients are also negative and are all highly significant at a greater than 95% level for women athletes, while only the 1980 boycott is significant for the male athletes, and that at a 90% level. Thus in 1976, 1980 and 1984, an average of 4 less female athletes were sent to the Games. For male athletes in 1980, 4 less athletes were sent to the Olympics by countries than in other years.

The coefficients on the Soviet bloc are also positive and highly significant. Countries under the direct influence of the Soviet bloc tended to send 72 more male athletes and 25 more female athletes to the Games than those who were not in the Soviet bloc. The planned economies coefficient is also negative and significant at a 90% level for men, but insignificant for female athletes. The planned economies sent about 31 less male athletes to the Games than non-planned economies; the governmental structure therefore played no role in determining how many female athletes were sent by them compared to the average and Soviet nations.

The coefficient on the Muslim, Dharma variable is also negative and highly significant at a 99.9% level. The coefficient shows that a 100% increase in Muslim and Dharmists in a country, leads to 14 fewer male participants and 7 fewer female participants being sent by that country than other religions would to the Olympic Games. Surprisingly, the coefficient on MUSHDA has a greater influence on male athletes than female athletes, contrary to our hypothesis and generally held notions. The male athletes therefore tended to be more negatively affected by culture.

Polity is also positive for both male and female athletes. It is not significant for women but is highly significant at the 99.9% for male athletes. A one unit increase in a nation’s polity increases male athletes sent to the Games by .795. The state failure coefficients are all negative for male athletes and that on ethnic war and genocide are significant at a 95% level. For the female athletes, the coefficients on ethnic war, adverse regime change, and genocide are all negative, with only genocide being significant at a 95% level. A nation which has a genocide in the year of the Games or two years prior to the Games tends to send approximately 4 less male athletes and 3 less female athletes to the Games. While ethnic wars are not significant in reducing the number of female athletes sent to the Games, a nation which has an ethnic war in the year of the Games or the year before it sends 3 less male athletes than countries who do not have an ethnic war. Worth noting is the coefficient on political wars, which, although not significant, is positive for female athletes; showing political unrest favors the participation of female athletes at the Olympics.

Finally, the time variable, which is negative and significant at a 90% level for male athletes, is positive and highly significant at a 99.9% level for female athletes’ participation at the Olympics. This shows a decrease in male athletes going to the Games by .419 over time but a 1.1 increase in female athletes participating at the Games.

VIII. Suggestions for future research

Although I have not succeeded in providing a full understanding of the economic factors determining participation and success, this study has gone further in explaining other complex socio-cultural factors. It will be useful to follow up on the non-economic factors
and try to better comprehend the underlying dynamics of why we have the kinds of results we have and direction of causality. Are nations underrepresented because they are relatively poor or are there more significant reasons? Why do certain countries win medals in the particular sports and not in others? The role of women in sports could be a signal of deeper national and regional propensities, more work needs to be done in understanding the issues.

Also, it would be effective to run these regressions on specific sports, especially to see how political and religious factors may affect participation and success in them.

IX. Conclusion

The Olympic Games have always been a unifier of nations, a symbol of endurance of the human spirit, but also a platform for making political and cultural statements. In this study I have employed socio-economic and political proxies to test the determinants of participation and success at the Olympics. Three hypotheses along these lines were formulated and our propositions were largely borne out, except in one respect. Cultural factors did not seem to affect women more than men but this is a conjecture that cannot be taken as conclusive. It however points out the need for us to be cautious in the way we treat both anecdotes and the results of this study.

In another important respect, this study breaks new ground by testing for religiosity in sports, specifically in the Olympics. I have not come across any study that tests a direct relationship of the two variables. This might well be the most significant contribution of this paper.

In broad terms, poor countries are under-represented based on the findings of the correlation of GDP and participation. A cursory glance at the maps showing participation and success at the Olympics shows that there are still countries, which have not yet participated in the Games and many others that have never won a medal. After studying factors affecting participation and success rates, it is quite clear that participation and success are influenced by a multiplicity of factors that is far from being described by rational choice. Uncertainty level is high because individuals making decisions do not have perfect information about the game, on other competitors, and therefore, on the chances of success.

While economic theory might succeed in explaining some of the issues determining participation and success, the reality is far more complex, uncertain and irrational. How for example can one explain issues of individual and national choices to keep healthy and perfectly competent individuals from the games on account of “faith”? How is it possible to understand in an econometric model why women are kept from participating on account of being a woman?

However, the models have shown us certain tendencies that point us in directions for further studies of the determinants of participation and performance.

References


Andreff, W. (2005), A Coubertobin Tax Against Muscle Drain, 4th Play the Game Conference: Governance in Sport: The Good, the Bad & the Ugly, Copenhagen, 6-10 November.


Olympic Official Website. www.olympic.org

http://www.guardian.co.uk/cuba/story/0,11983,712796,00.html


Appendices

1) Regression models of dependent variables
Regression 2
TMEDAL = β₀ + β₁GDP + Σ β₂Popula + Σ β₃Mushda + β₄Polity + Σ β₅Soviet + Σ β₆Planned + Σ β₇Ethnicwar + β₈Regimech + Σ β₉Poliwar + Σ β₁₀Genocide + β₁₁Host + Σ β₁₂Boy1976 + Σ β₁₃Boy1980 + Σ β₁₄Boy1984 + β₁₅Time + µ

Regression 3
PMEN = β₀ + β₁GDP + Σ β₂Popula + Σ β₃Mushda + β₄Polity + Σ β₅Soviet + Σ β₆Planned + Σ β₇Ethnicwar + β₈Regimech + Σ β₉Poliwar + Σ β₁₀Genocide + β₁₁Host + Σ β₁₂Boy1976 + Σ β₁₃Boy1980 + Σ β₁₄Boy1984 + β₁₅Time + µ

Regression 4
PWOMEN = β₀ + β₁GDP + Σ β₂Popula + Σ β₃Mushda + β₄Polity + Σ β₅Soviet + Σ β₆Planned + Σ β₇Ethnicwar + β₈Regimech + Σ β₉Poliwar + Σ β₁₀Genocide + β₁₁Host + Σ β₁₂Boy1976 + Σ β₁₃Boy1980 + Σ β₁₄Boy1984 + β₁₅Time + µ
THE IMPACT OF IMPLICIT PENSION DEBT ON SOVEREIGN CREDIT RATINGS

David G. Wiczer
Carleton College

ABSTRACT
As sovereign bonds gain importance for national budgets and international credit markets, rating agencies play an especially important role in these credit markets. The raters must anticipate all the risks that could cause a default, but there seems to be a giant risk looming over the modern welfare state that is getting coverage from the popular media but scant in the raters’ reports. Public pensions will pose a tremendous drain that may destabilize budgets in coming years. I investigate whether ratings agencies are systematically considering these payments in their assessments. Though I am able to predict ratings reasonably well, I do not find conclusive evidence that the raters give a negative weight to the size of the pension burdens, nor even to explicit debt stock. I suspect that these results are distorted by an unaccounted structural relationship that determines the stock of public liabilities. Still, generalizing from a limited sample, I am able to project that ratings might be roundly lower if pension debt were considered analogous to the stock of explicit debt. This has wide-ranging implications for government budgets and lenders alike.

Italy’s public debt stood at 107% of its GDP last year and Japan’s was 173%, yet both continue to float bonds rated at AA- by Standard and Poor’s. Presumably, such a debt load increases the risk of default on future issues of bonds when the payment schedule simply becomes too large for the borrower. Yet, the major credit ratings agencies (e.g. Standard & Poor’s, Moody’s, Fitch) endorse these states as investment-grade issuers. If a private borrower were carrying such loads they would be considered a high risk, but in the sovereign market what explains the raters’ tolerance? Sovereigns have an advantage in credit markets because they underwrite their own debt with the great power of taxation. But perhaps Uncle Ben from Spider Man ought to remind markets, that along “with great power comes great responsibility.” Along with the power to generate revenues from their citizens, governments have corresponding responsibilities to provide services. This arrangement is
even so rigid that the responsibilities are labeled “entitlements.” While raters copiously document the explicit liabilities of a state, they make little mention of these implicit liabilities, which are fast becoming the largest burden on state coffers. If a sovereign issuer’s rating is bolstered by the understood power to tax, perhaps the special liabilities of modern welfare states should also be acknowledged.

If raters are to assess implicit liabilities, they might begin with the most prominent of the entitlements, public pensions. A recent article in The Economist chronicles the great progress being made at accurately assessing the corporate books. The process now includes pension liabilities in the “company's balance sheet and risk profile.” For justification, it likens these future payments to explicit lines of bond debt. In the public sphere, the implicit pension debt (IPD) is also analogous to a bond: plan participants ‘lend’ their taxes today, with the assumption that they will receive an annuity in the future. According to this analogy, governments are constantly floating bonds to finance current pension payments and IPD is a vast liability, greater even than the explicit state debt. If sovereign debt markets consider only formal debt, then they systematically undervalue the potential future budgetary drain and an additional risk for would-be lenders to consider. Governments benefit from the current tax inflows, but it is not obvious if they are being held to task for their reciprocal promises.

Though ratings agencies (e.g. Standard & Poor’s, Moody’s, Fitch) list some of the factors they consider (Fitch 2002; S&P 2005), they do not provide any weighting scheme nor much word on how they treat the unique circumstances of a welfare state. Prior studies have predicted sovereign credit ratings with major macroeconomic statistics but include only explicit liabilities. I add IPD as an explanatory factor to test the manner, if any, that it accounts for some of the variance in ratings. Because the ratings are meant to analyze states qua debt issuers, IPD ought to figure in as it is relevant to their debt repayment ability. It is here treated as comparable to the stock of explicit debt, which clearly ought to add to the risk associated with additional debt. To mirror the manner that lenders now worry about the accumulated national debt, I study the present-value of the expected future stream of pension payments analogous to the value of outstanding government bonds.

Were agencies considering IPD, this regressor would be negatively correlated with the rating, all else equal. However, the tests that follow generally do not find a significant systematic relationship between IPD and rating. Unexpectedly, even the relationship between rating and explicit debt was dubious. Certainly, this empirical study is raked with estimation problems, and the estimated coefficients may have been upwardly biased by under-identified endogenous relationships. One way to abrogate these estimation difficulties has been to limit the sample to only non-OECD, low-to-middle income countries, in which the relationships between credit rating and debt stocks is more pronounced. However, the intent of this study is to investigate the general behavior of credit raters, so this fix is less than ideal.

To depart from the positive analysis for a moment, ratings agencies should be considering IPD and consistent, comparable estimates of IPD should be public. Ostrich-like behavior is always an attractive option, but IPD could be disastrous if ignored, while full
disclosure can help market mechanisms mitigate the pain. Credit markets rely upon ratings agencies to expose future risks, identify emerging issues, and generally alert potential investors of the full level of risk; to fulfill these function raters cannot ignore IPD. If IPD goes undetected, governments have an incentive to trade off their explicit debt and build implicit liabilities. Borrowing on credit markets costs a market rate of interest and has some incremental effect of increasing their risk premium by increasing their overall debt level. If markets do not include IPD in their risk appraisals, then a country might manipulate its budget to borrow less through traditional bond offerings but underfund the promised payments it makes for the future. The US is now flirting with this scenario, as the social security tax exceeds current payments, and this fund is tapped by the government general budget. At present, the government explicitly borrows from the fund, and though this vast supply of loanable funds at the Treasury’s disposal may well exert downward pressure on the interest rate, it does not hide the borrowing. Were the government less transparent, or merely to run its public pension from a single general budget, current surpluses could pay for general government expenditures instead of a bond issue. Pension taxes are happily paid by citizens who expect compensation through an old-age annuity. In this metaphor, the tax takes the place of a bond issue and the IPD is the face value of such debt.

Unchecked, IPD can easily grow out of hand as legislators spend their successors’ budgets. This continues until the government has amassed too many promises to cover. At that point, they must choose what liability they will default. Even in recent history, governments are not consistent with whose claims they prioritize. In their 1998 crisis, Russia decided to honor their foreign bonds but scrapped its domestic debt and allowed pension and welfare payments to build as arrears. In 2001, Argentina scoffed at foreign debt collectors, but largely maintained domestic benefits. Hence, foreign creditors ought to be just as interested in the solvency of IPD as the potential domestic beneficiaries. Even without outright default, a debt proto-crisis could be precipitated by large IPD. As in so-called “first generation models” for currency crises, exorbitant debt loads will begin to scare away investors, and, as capital flees, the value of remaining assets plummets. Credit raters could avert an asymmetric response (too lax, and then too harsh) by alerting the markets to the real risk posed by IPD. Markets would then price in the risk premium, chastening governments that are loose with their IPD and increasing their borrowing costs until the quantity decreases. Conversely, ratings agencies could join the rest with their head in the sands and exacerbate the extremes by encouraging unfounded optimism until the moment of crisis.

Presumably, the present value of IPD incorporates the full cost of pensions ceteris paribus, but just maintaining constant payout schemes will cost more as a so-called demographic time bomb increases the number and longevity of recipients. The prior “doomsday” scenario likened an IPD crisis to the first-generation model of debt crisis, fueled by government policy choices. Now, the aging population will autonomously increase the government’s load without any policy changes. The alarm is beginning to sound in a host of studies from IGOs, the popular press and academe. To describe the crisis, they key upon
the dependency ratio, which describes how many workers support each pensioner. The limit for an unfunded scheme seems to be around two supporters for every dependent (50%). After this point, the burden on current labor becomes overwhelming and taxes cannot be raised any further. Providers must either borrow or tap savings. In accordance with the bond analogy, this is the point at which tax revenues become insufficient to meet the repayment schedule and some manner of default becomes a very likely possibility. Because demographic trends are usually well set, reliable cassandras see this point is rapidly approaching for IPD (Disney 2000; Fenge et al. 2002). The situation will deteriorate most rapidly in Japan and Italy, where dependency ratios are expected to rise 190% and 170%, respectively, by 2050 and they will surpass the 2-1 ratio around 2020. Others look slightly less bleak, but many such as France and Canada will still face a 40% ratio by 2030 (Hviding and Mérette 1998).

In addition to the demographic effect, incentives created by the scheme design may have one-off effects (see Kohl et al. 1998 for a survey of these topics). Long-term studies in developing countries even find that pension plans themselves reduce fertility rates because children are not needed as a substitute for old-age insurance. In correcting a market failure, the national pension scheme undermines its tax and labor base and makes its own promises more uncertain (Yoon 2001). Further, if a generous pension depresses savings rates, it saps the very credit markets that may be needed to cover benefits in the future. In this case, interest rates would rise, but ironically, the higher cost of capital would imply a higher discount rate of the liability and make it seem less daunting in present-value terms. This is one reason that this study uses a single, consistent discount rate. Other scheme aspects, like incentives to continue working, can reduce the number of payments that must be made but also have myriad effects in the labor market that confuses any attempt to project future pension costs. If participants work longer in their later, higher
salaried years then this may increase their benefits under certain schemes and therefore has an ambiguous effect on IPD and the macroeconomy. Certainly, many of the cascading effects cannot be predicted, yet even a rudimentary valuation of the future IPD burden would increase understanding of a country’s balance sheet. Anything to clear through the smoke screen would improve the markets’ efficiency. Credit rating agencies are in the business of disseminating relevant information and therefore ought to include IPD in their scores.

**Institutional Background**

Three major institutions relate to my study: public pension systems, sovereign debt markets, and credit ratings agencies. My study explores the interaction of the three together, though most prior work has considered each in isolation and missed the feedback mechanisms amongst them.

A public pension system is essentially old-age insurance plans managed and financed by the government. It fulfills a market role, but is created as a complicated legal edifice. The most common arrangement of public pension is a “pay-as-you-go” (PAYG) defined benefit plan. PAYG is something of a political euphemism for what corporate pension argot would deem “unfunded.” Current tax revenues cover current pension payments in a form of direct transfer from workers to beneficiaries. An individual’s benefits are determined by the “defined benefit” (DB) program, which sets them by the recipients’ final wage level and/or years of contribution. The payments are independent, however, of the funds available to cover them and therefore it is something of a coincidence when inflows of the PAYG are exactly the same as the outflows determined by the DB. From the perspective of the participating laborer, he is “buying” old-age insurance by contributing some percentage of his wage for his entire working life in return for the promise to receive some fixed benefit level ad infinitum. For the government, it is covering current liabilities with receipts that carry with them future liabilities. In essence, the government is continually floating bonds, underwritten only by the power to tax the next generations. The macroeconomic position of the plan is fairly straightforward. The liabilities are the average yearly benefit level, regardless of variation on the individual level, times the number of years the participant collects, times the number of recipients. These are balanced by assets, which are contributions that exceeded the payment obligations and have been put away in savings. For most public pensions, these assets are relatively minor, so the liability is covered only by the anticipated tax revenues.

Alternative schemes either manipulate the funding mechanism or the payment schedule. A funded plan puts the receipts in savings to cover the future liability. This amounts to a government-run savings plan, by which the present-value of liabilities is equal to the assets of the plan. Benefits must be constrained to the amount of available funds. Even without creating a funded system, future payments can be brought more inline with income through the “notional direct contribution” (NDC) systems being tried in Sweden, Poland and a few other states. The funding is still PAYG, but future benefits are determined by how much the participant contributed. This is designed to encourage greater participation today, but future payments will be just as large and still are not funded by cur-
rent savings. Tomorrows benefits cannot exceed today’s contributions, but a payment
deficit might still occur if future receipts flag. The scheme design itself may even exacer-
bate IPD because it explicitly gives participants a benefit level they can claim, where as
the DB may be reduced more stealthily. For example, DBs are currently adjusted for inflation
by indexing the payments to wage growth, but rebasing them to a general consumer
price index would allow the liability to grow more slowly than the economy and thereby
reduce the IPD level. NDCs, by contrast, pay according to contributions, so a reduction
of payments would have to explicitly admit that the participants are receiving a lower per-
cecentage from their notional savings package—a lower return on their “investment.” The
macroeconomic impact of any reform might be the same reduction in IPD, but under the
NDC the microeconomic consequences are more transparent to those being hurt. NDCs
may raise the resistance to such an escape valve mechanism, and their unanticipated con-
sequence may be to force an outright default, and thus make the entire government prom-
ise less tenable.

Governments finance their activities not only through taxation; they are also some of
the largest participants in the global bond markets. “Sovereign” issued bonds have a
mature, deep, and stable market by which these entities can raise funds at interest rates sig-
nificantly lower than most private firms. As earlier alluded, their perceived security is not
solely because of their current budget size, which some companies already exceed, but
also because their revenues can be expanded by fiat. In these markets, a government may
float bonds in any number of flavors. Short-term bonds, usually called bills, are issued fre-
quently, traded regularly and carry a lower interest rate. As the bonds become less liquid,
their interest rates tend to rise and this term structure is shown by an upward-sloping yield
curve. The interest premium on longer-term bonds partly reflects the increased risk asso-
ciated with holding debt that may become less secure in the future. In addition to gener-
al risk of default, debt denominated in a so-called “soft currency” is subject to exchange
rate risk that increases the volatility of their returns in terms of the home, stable currency.
This effect reduces the liquidity of an investment by limiting when returns can be repatri-
ated. To mitigate this problem, countries will issue bonds on foreign markets denominat-
ed in foreign currencies so a lender may use his own currency. Such instruments are not
unequivocally less risky, however. The government borrower may be more likely to
default because it must obtain foreign currency to repay. A distressed borrower cannot just
tap the local tax base, because foreign exchange resources are contingent upon the state
reserves and the external balance of the economy. Of course, any currency denomination
 carries its own risks, as do different term structures. Each risk assessment considers the
idiosyncrasies of particular issuances, which will have bearing on the choice of model and
will be further discussed later in this paper.

Credit ratings agencies estimate the risk level of debt instruments issued by private and
public institutions and publish their findings as an indicator on a 16 point scale. The three
most prominent firms, Standard and Poor’s, Moody’s and Fitch, are concerned mostly with
the risk of default, which includes any disturbance in the debt repayment schedule. Their
private sector assessments are almost exclusively determined by bottom-line indicators,
but sovereign issuers pose more of a challenge because various qualitative factors of the political economy must be considered. A sovereign first requests a rating. The issuer then pays the agencies to send experts to assess all of the relevant risks that might interfere with timely repayment of proposed debt. Along with a sixteen-point rating scale, denoted by a letter scheme, the agencies publish relevant numerical indicators, and a qualitative explanation. The ratings are periodically updated, and often an “outlook” forecasts an expected change.

Only recently have sovereign ratings become widespread and extended beyond the major economies. For most of their history, the ratings have been very stable—S&P’s were 87% consistent year-to-year (S&P 2005). But whereas twenty years ago the ratings process may have been mostly a self-congratulatory occasion for rich-country financiers to announce a new bond issue, countries with serious debt concerns are now being rated. Even today, the bonds of only about 30 countries are actively traded. Another 60-75 have been rated but very few have long histories. Most of the outstanding debt has a long maturity and new issues are relatively infrequent. However, as developing countries enter to the field, high-profile debt crises spur investors to demand comprehensive and reliable ratings. This puts a much greater onus on the raters who must untangle a morass of complicating risk factors in markets that can be quite unstable. For example, with local currency bonds, they have to understand local financial markets and the government’s ability to leverage debts against each other, should the need arise. Foreign currency bonds study macroeconomic factors and especially external sector dynamics. For any sovereign instrument the raters critique the government budget and the risk of political or economic shocks.

As is becoming clear, there seems to be some division of sovereign issuers between those in wealthy OECD countries and all the rest. The low-to-middle income countries that have been issuing new debt might be more analogous to the mid-cap corporate debt or even the new mezzanine instruments that have become popular of late. In the corporate world, these ratings are not just smaller, riskier versions of their older, stable, better-understood large-cap cousins. Debt on middle-sized growing companies has a whole new set of concerns, beyond even the fast growing and changing nature of the business. The total stock is more closely scrutinized because they have not built a record of debt repayments and management is presumed to be inexperienced with issues of leverage. These special borrowing constraints suggest that a similar stock of debt as a percent of revenue level might be interpreted more critically in the market for larger issuers (Shuettrim et al. 1993). This same metaphor might hold for comparing the new, middle income countries to more established issuers. Mezzanine debt, which combines some characteristics of equity, also might be instructive, because there are more considerations of governance tied to the debt of developing countries (Fitch 2003). These observations presage the following tests, whereby I unsuccessfully try to pool older, better-developed issuers with the larger cohort of emerging-market sovereigns. Just the ratings criteria are not completely standardized across all breeds of corporate issuer, so too for sovereigns. My tests seem to corroborate other evidence that the non-OECD sample tends to be under heavier scrutiny, and that the stocks of debt are carefully considered.
The role of raters is only abstract unless a default actually occurs. This is no recent phenomena, and even though the determinants constantly change, history has well established that default is a real possibility. In these circumstances, the ratings agencies come very much to the fore. Even before their advent, the ancien régime was a particularly troublesome borrower that frequently cancelled or rescheduled its debt payments. In the post-world war II era, only a handful of countries have been unable to meet their payments, but debt distress is widespread and many have openly flirted with bankruptcy. A spate of recent episodes in East Asia, Eastern Europe and South America have caused great financial loss and kept headlines focused on the risks associated with international capital.

![Sovereigns With Foreign Currency Bond Defaults](image)

The ratings agencies are interested in the dynamics of such crises. Not only are they concerned with predicting the acute and long term causes, they also consider how the fall-out will affect different parties. They are not only predicting the likelihood that a debtor will become distressed, but also how debt distress can change a nation’s priorities. When an issuer cannot pay one liability, it affects the entire balance sheet. In a government budget, there is a set amount of revenue, and distributing it amongst the various responsibilities often involves a conscious calculation about which creditor might be more forgiving. Thus, although IPD is most directly analogous to debt held in local currency and by domestic creditors, it poses a threat to anyone with a claim on government coffers. One theory posits that domestically held debt is mostly considered a redistributive concern. Because the government has power of taxation over these creditors, any payments from the government becomes fair game for an income tax—these parties are both assets and liabilities. However, IPD does not follow this logic: its sheer magnitude means that a tax increase of this size would be untenable, and also because pension income is not just gravy for an institutional investor, it is the livelihood of a country’s most vulnerable. To simply tax it away is akin to default and abandoning the entitlement—IPD is no more fungible a budget line than any other of the explicit liabilities.
Review of Literature

Three independent bodies of literature inform the relationship of pension liability to sovereign credit rating: financial economists have studied the (a) determinants of sovereign ratings, and another group has projected (b) the impact of pension plans for private issuer debt. A huge amount of work has focused on (c) the future of public pension systems and these efforts have created several methodologies to measure IPD. Like any effort in economics, this study leans heavily on these prior works; however each seems to tell a separate story. Pension work has not considered its sovereign finance implications and sovereign finance has not systematically looked at the unique challenges associated with the welfare state. The most difficult take is not so much to understand these arguments but to relate them.

a) Sovereign credit ratings and crisis risk

Many studies examine the meaning and determinants of sovereign credit rating. A rating ought to consolidate the indicators of credit risk, add some expert discretion and then exert some sway over the premium that a country must pay on its debt. Before credit ratings were widespread, country risk rankings like Euromoney of Institutional Investor were predicted using a host of macroeconomic and political variables. These figures are more relevant to investment within a country than the state’s likelihood to service its own sovereign debt, but are still highly correlated with interest rate spreads within the country (Aboulamer 2004). These have the advantage of a wide country sample while samples in prior examinations of credit ratings have been weighted against developing countries. These large samples are relevant to my own efforts because I argue the debt is a concern both for the indebted poor and seemingly healthy rich countries. Cosset and Roy (1991) use several macroeconomic variables to describe around 80% of the variance in risk rating. However, among the trends picked out by these studies, external debt does not have a significant, systematic relationship. This runs counter to my own theory whereby larger stocks of debt increase the issuer’s risk. Perhaps risk rankings are concerned more with short-term risks, as they cater to international portfolio investors. IPD and other long-term debt may be less relevant to these parties, but long-term bonds cannot pooh off the long-term risks involved from debt stock variables and IPD.

Cantor and Packer (1996) applied Cosset and Roy’s technique to determine long-term foreign currency sovereign ratings. They constructed a model of eight macroeconomic variables that theoretically play a role in a country’s capacity for debt repayment. The model predicted 90% of variation in rating across countries. Still, the rating itself still predicted the rate premium better than the component elements. This suggests that ratings are based largely upon a country’s vital statistics but they also add meaningful information beyond merely summarizing a few numerical indicators. Surprisingly, among Cantor and Packer’s eight macroeconomic ratings determinants, neither external nor fiscal deficit is significant in their regression. The authors suggest that this counterintuitive finding may be because heavily indebted countries face high borrowing costs that prevent them from running further deficits: market discipline preempts the ratings agencies. They suffer the same endogeneity that will plague my own study. Debt is cheaper for countries with good
macroeconomic profiles, so the partial slope coefficients are upwardly biased, though, in theory, high levels of debt hampers future borrowing. Even if the study implies that markets themselves limit debt in high-risk countries, however, with traditional lines of credit cut, perhaps these states will turn to covert forms of borrowing. IPD is one such off-budget method to avert credit limitations and continue to borrow.

While the response to debt can be draconian in some cases, others are escaping needed discipline. The vagary of the sovereign debt market itself deserves some analysis. The current, near-sighted paradigm overemphasizes the flow of debt and output to assess risk. Whether because of endemic myopia or just a short memory, the contemporaneous budget figures much more prominently than past accumulations when a county floats a bond. As a result, Alesina et al. (1992) find that high-debt wealthy countries can still borrow cheaply, perhaps because their current flows seem sustainable. Studying countries entering the EMU, they find that as a whole, the debt to GDP ratio is not a statistically significant determinant of the interest rate risk premium. In an interesting twist, splitting their sample by debt level, debt is significant for the high-debt countries. The market response seems incongruent, tolerating debt with no marginal increase in interest rate until it “wakes up” at a certain cusp and responds. Might markets respond the same to IPD—complacent until a certain point? The implications in my own study are important. While I attempt to combine a sample from across wealth levels, these studies herald the difficulty in this broad sample and the greater success when it is limited to middle income countries.

If it is true that a broad swath of countries are taking advantage of blithe market confidence, perhaps a groggy awakening by investors may bring a crisis. If markets are efficient and consider all information, then they will already price-in risks and marginally adjust to new risks as they emerge. A rating change would not have a significant or sustained impact upon the rate a borrower faces. However, this is not observed empirically. Packer and Cantor (1996) find that rating change events have a significant impact on borrowing rates for countries. This suggests that markets do not automatically integrate all available information by themselves. Rather, they trust ratings to summarize and interpret macroeconomic data that is usually freely available and timelier. With central role of agencies in mind, Reisen and Maltzan (2004) call on them to check market “euphoria” and thereby mitigate the long and pronounced boom-bust cycles in sovereign debt. Again with event study analysis they conclude that the market responds strongly to a rating event, even when it is already anticipated. But they also believe that the agencies are not using this power to intervene in an overzealous boom in a way that might mitigate the inevitable bust. If IPD follows this pattern, it could cause great turmoil as investors re-evaluate their positions in state bonds once thought risk-free.

To counter the prevailing, over-buoyant moda, the IMF has begun to endorse a “balance sheet approach,” which considers the risks arising from stocks, rather than just flows (Allen et al. 2002). This new method goes beyond explicit financial arrangements to include liabilities to households such as IPD. In fact, this philosophy is even more inclusive than only the public debt that my study keys upon. The IMF’s stance focuses on a country’s total external debt stock which also incorporates the way private entities carry
debt as part of the risk of an all-encompassing debt crisis. Their recommendation is informed by more sophisticated models of debt crisis which emphasize that crises are set in motion by untenably leveraged financial institutions. An IPD crisis would follow this model well as huge liabilities are only waiting to be exposed and threaten the whole system’s solvency. The IMF calls for more complete accounting that would reveal impending risks and impose more justified credit constraints. Though IPD is not the only source of underreported unfunded liability (e.g. public healthcare), it is a large and calculable element whose full consideration would be a significant step towards better accountability and risk management.

b) Assessing the value of public pensions

Though IPD may seem concrete in abstraction, the valuation of such liabilities is controversial and has inspired a whole body of literature. My notion of IPD as a quasi-government liability and tax receipts as an implicit loan derives from a framework developed by Kotlikoff and Raffelhuschen (1996). They developed a theory of “intergenerational accounting,” to convert deferred costs into a present value balance sheet. The authors have laid the foundation for this perspective of budgetary analysis, which closely resembles Ricardian equivalence. They theorize that governments endemically defer costs and move asset values into the present, a constant balancing game of mortgaging against the future without explicitly creating liabilities. They illustrate this point with simple to period example: a government could levy a $100 of tax on N or take an $800 loan from N, less a $700 transfer. Then levy a tax next period on N of $800(1+r) and repay N the $800(1+r). The real impact on N and the government is the same, but it seems that the government actually is paying N $700 in this first period. IPD echoes the scheme with N, but more opaquely and on a much larger scale.

Fenge and Werding (2003) key on the Kotlikoff method and characterize the full pension burden as the amount required to compensate the present participants in the plan. To put this estimate in terms of a government’s debt liability, they perform a thought experiment that asks how much additional tax is required to fully settle the plan as it stands. An alternative assumption would allow new claimants to enter the plan even as existing liabilities stand. This dynamic model might be more realistic because future governments will probably have to cover new entrants in addition to current participants. However, this is inconsistent with my purpose, which sees IPD as debt, rather than a means to estimate the future fiscal position of the government. Applying the bond analogy; the stock of debt does not include potential additional borrowing even if this is likely before the existing liabilities are covered.

Another group of studies aims to estimate future budgetary flows due to pension plans. They must project retirement and demographic patterns and, because defined benefits are indexed to rise with wages, these must be forecast for the life of the IPD. A model by Kohl and O’Brien (1998) projects the fiscal and macroeconomic picture for OECD countries. However, it goes beyond the scope of IPD, considering such liabilities as increased health care. Unfortunately, it cannot be converted to a present value estimate of IPD because it only forecasts flows in a few years. Roseveare et al. (1996) and Holzmann et al. (2004)
make present value calculations for IPD in about 50 countries, but use slightly different macroeconomic and methodological assumptions—a point that will prove crucial.

c) The influence of pension deficits on private sector rating

Though studies have not directly linked IPD to sovereign debt ratings, a third body of literature draws parallel observations in the private sector. Carroll and Niehaus (1998) analyze the impact of corporate unfunded liabilities on credit ratings. The analogy could be informative; the market is asymmetric by scarcely rewarding adequate funding and insufficiently chastising under funding. Extrapolated to the state level, where plans are all essentially unfunded, this asymmetry removes a potential incentive to increase the funded coverage. Dhaliwal (1986) performed a similar analysis for the impact of unfunded pension liability on risk premium and found a statistical connection. Further, he could not reject the hypothesis that pension debt has the same effect as other debt. In the private sector, it seems that markets do account for unfunded pension debt. But while corporate pension liability is fully disclosed, markets may misstate IPD because it is not explicitly budgeted.

Research Design

My model tests the impact of IPD on long-term foreign-currency-denominated credit ratings for 50 sovereign issuers. The case selection is mostly motivated by the availability of data because only a limited number of sovereigns have both estimable IPD and a credit rating. On theoretical grounds, this selection is defensible because it is a more representative sample of countries than merely including the entire population of rated countries. If I were only trying to understand the determination of sovereign ratings, using the entire set would make sense. However, I also want to create an accurate representation of the impact of IPD on public finance. Only about half of recognized states are rated, and these are clustered in the well-rated, wealthy countries. As much of the literature has already established, the market seems to respond differently to rich and poor countries. An accurate and generalized relationship must select cases that are more evenly distributed so that the model does not over-emphasize the rich country response. The countries in my sample have a wide range of IPD and credit ratings. Within my whole dataset, Argentina has a rating of DDD, while the US is rated AAA, and IPD values range from under 50% to 300% of GDP. Though twenty of the observations are from the OECD and all have ratings of A or higher, pooling all of my data should mitigate heteroscedasticity.

The dearth of comparable IPD estimates is the most significant shortcoming in the cross-section of data. In fact, these estimates may prove even less comparable than thought at the outset of this study. This was certainly expected, as this study primarily seeks to explore the possibility that IPD may catch the world unaware and precipitate a crisis. Though some components of government budgetary data are available and the ILO even publishes a methodology for estimating IPD (2001), the most crucial projections for population and program growth either do not exist or are proprietary. My solution has been to rely upon two studies, one commissioned by the OECD by Roseveare in 1996 and the other in 2004 by Holzmann. I pool the data and use something like a truncated panel in which all of the data for a given country corresponds to the year of the IPD estimates.
Data for the OECD countries is a snapshot in 1996, and the Holzmann set of middle income countries all use 2004 so that all the rest of my data correspond to the same year as the IPD data. The basic model structure employed is as follows:

Average Foreign Currency Credit Rating = $\beta_0 + \beta_1$*IPD estimate + $\beta_2$*Dummy Direct Contribution + $\beta_3$*GDP/Capita + $\beta_4$*real GDP growth + $\beta_5$*Inflation + $\beta_6$*External Balance + $\beta_7$* Public Debt + $\beta_8$*Dummy Default History + $\beta_9$*Dummy Developed

Within this reduced-form function, there is a deeper, structural relationship. The level of public debt is determined by the ability to pay and the cost of borrowing—both correlated with the credit rating—which introduces a relationship between public debt and the disturbance term. One would suspect that this endogeneity, a positive correlation between debt stocks and average credit rating, will upwardly bias the estimated coefficient $\beta_7$.

Unfortunately, there are no obvious instruments to be found in the literature or theory. One might suggest interest rate to proxy for cost and GDP per capita to represent demand, but both are highly correlated with credit ratings so they are not the optimal, exogenous instruments. Rather, this formulation would turn the under-identified reduced form equation above into an over-identified relationship. To explore potential in this format, I attempt a two-staged approach with average government bond yield as the instrument. This is admittedly expected to fail because it is not exogenous nor does it actually predict debt well. This may be because the yield is calculated for a relatively small time window, while the debt stock has been built over many years during which yields have not been steady. Empically, this year’s interest rate is not greatly covariant with public debt, and its own endogeneity suggest the need for a third stage which would further reduce the efficiency of the estimates. I press ahead with the two-staged approach anyway for lack of a better alternative—yet.

To grasp the market for public debt, one considers the interest rate on that debt and the demand and supply functions concerning it. At any given time, the quantity that the government demands to borrow is inversely related to the interest rate which it must pay on the service, while the supply of funds will increase with a higher yield. Over time, the debt stock is the accumulated total of each of these market-clearing moments. Because debt issued at different dates for different durations have different nominal interest rates, it would be arbitrary to choose a single, benchmark. Moreover, it is difficult to find a single, cross-country comparable bond. In light of these concerns, the average yield is computed as the implied interest rate from the average amount of public debt and the average interest payments on that debt over a data window of four years. A larger window over which to average the rate would help improve its fit on the data, but this information is often controlled in private databases from which this research is bared. Another instrument would have to be found that might predict the demand for debt, but independent of the credit rating. This is particularly difficult because the rating considers so many macro-economic variables, which may also motivate the debt stock. Even a political variable, perhaps some leaning is associated with chronic profligacy, would also presumably be associated with a low credit rating. With so many entangled variables, the “interconnectedness of everything” makes suitable independents difficult to identify, but a rough estimation
with average interest rate will have to suffice. If they perform their function, other variables will be uncorrelated, and predicted debt will contain only the fitted values. The second stage has “filtered” out the cross-effects that interrelated it to the disturbance term. Unfortunately, the results show that the instrument does not have sufficient explanatory power over public debt, nor does it sift out the residual’s correlation nor the multicollinearity.

All of my models use some form of the average rating on long-term, foreign currency-denominated bonds as a dependent. There are many nuances layered into my choice of dependent, AVR. My study intends to test the breadth of the raters’ outlook and to question whether they are systematically using all of the information that is available to them. Ratings are divided into three bond categories: short-term, long-term local currency and long-term foreign currency. Though IPD is repaid in the country’s own currency, it is not so obvious that it ought to impact the risk on local currency bonds more than foreign-denominated ones. The “balance sheet approach” suggests that the risk to the external standing is based upon the whole stock of liabilities. S&P notes that “in a stress scenario such [domestic, local currency] debt behaves much like foreign currency debt,” (S&P 2005) because a country generally may not have access to enough liquidity or bond holders may convert their asset and sap foreign exchange markets. From a theoretical perspective, local currency bonds are generally thought less susceptible to default because governments can monetize the debt and still keep their repayment schedule. Presumably, they also have greater access to their own currency through the local tax base while foreign currency must be obtained from current account receipts. IPD debt, however, differs from other local currency-denominated debt because excess inflation will not diminish its onus. Benefits are linked to the nominal growth in the economy, either wages or GDP and therefore payments maintain a constant real value. Moreover, defined benefits generally guarantee real purchasing power for their recipients because they are also designed as poverty insurance.

IPD does not enjoy the advantage of easily available currency like standard local currency instruments. Though the government still has access to the domestic tax base to draw currency for its pension liability, in reality it is not paid more readily than the externally held debt. The track record for defaults suggests that often sovereigns prefer to default than to raise taxes (Packer 2003) Recent experience shows that even middle income coun-
tries will falter in their pension promises for lack of will or ability to service it (see literature on Ukraine, Aslund 1995). Despite their domestic default, these countries may still have external reserves to service their foreign debt. As argued earlier, the scale of IPD can so overburden a country’s balance sheet that sufficient taxes cannot be raised without disastrous economic impact. Therefore in a crisis, the country must allocate its resource pool either to pay local or foreign creditors, and therefore both groups of lenders incur risk from IPD. Local currency ratings are not widely studied, but they seem to vary due to the structure of domestic financial institutions rather than broad macroeconomic indicators. Imperfect domestic financial markets might play a significant role in the risk associated with a country’s explicit domestic liabilities. IPD, on the other hand, circumvents the capital exchange. Though IPD is denominated in the local currency, it is not necessarily a significant determinant of local debt ratings. There is no systematic way to control for these domestic idiosyncrasies to relate IPD to these ratings, rather, they seem to exist in different planes. Foreign currency ratings are more broadly studied and more directly relate to the macroeconomics factors of a country—presumably including IPD.

Though the foreign currency debt rating is published in code form (e.g. BB- or Ba1), it must be transformed into a 1-16 scale for a quantitative analysis. Still, this variable is categorical, but in the first OLS model it will behave as a continuous distribution. The theoretical reasoning behind this move is described well in Ibara-Salazar (2005), whereby the observed integers are overly a continuous ‘ratings index.’ The model predicts this dependent index from -? to ? and then a confidence interval may be drawn to interpret this on a discrete 1-16 scale. Aboulamer (2003) suggests a method to interpret his own results by which the actual prediction is taken along a normal distribution. Then, with the standard error of the estimate, a rating category is given a probability. For example, if the predicted value is 14.5 then the standard error allows one to build a confidence interval that the rating would be 14 or 15. This estimation may be forgivable. The difference between a 14 (AA) and 15 (AA+) may have little functional difference to credit markets, who understand that ratings ought to be interpreted with an eye to future revisions. In my model, Croatia has a forecasted rating of 7.66. With an 80% probability, I cannot reject the hypothesis that the rating is 7, its true value. Within this same interval is the other possible rating of 8, which Fitch had assigned to Croatia’s short-term liabilities. Without detailing the available Croatian government securities, they have a range of maturities from short to long, and therefore most securities may have a true risk profile somewhere in between 7 and 8.

However, the model cannot always be finessed in this way, as a non-asymptotic fit line can make spurious predictions beyond 16 or below zero. These could be interpreted as a one-tailed probability, but this is dubious. If the model predicted a 17, a one-tailed interval would make a true 16 value seem more likely than if one abstractly drew an interval between 16 and 18. There is, however, no good reason that a true 16 is more likely from the forecast 17 than if the forecast had been within the acceptable range and we had used the two-tailed approach to predict two possible and less likely true values. This is not merely a theoretical shortcoming, as Japan has such good macro economics that its fore-
cast value is 18.7. With a single-tailed interval, the true value of 16 is within an acceptable 90%, however, the prediction actually grossly overstated Japanese risk and should not seem a better response than if the model predicted 10.7 on a true 8. Further, while it does seem that in some cases the static difference between a 14 and 15 is not material, event studies conclude that moving from 15 to 14 significantly influences markets, and the even a static difference between 11 and 10 is likely significant, where countries are already under tighter scrutiny.

To address some of these concerns, I use two types of binary choice models, which convert the average rating into a categorical variable, either one or zero. This simplification is justifiable as the ratings agencies make a distinction between “investment” or “speculative” grade. A rating above a 9 qualifies as investment grade and there is not much risk premium assessed to any country with this label. Placing the “cusp” between the grades is quite important, as numerous event studies show a downgrade can significantly disturb a country’s borrowing ability, and lower-rated debt is generally more cautiously assessed. The distribution of dependent in this sample even seems to lend itself to dichotomous categories, as it has a fairly clear bimodal distribution. A happy coincidence, the “9” rule also divides the 50 observations right down the middle, with 25 investment and 25 speculative grade countries. Still the division is somewhat arbitrary and so models were attempted with the line at both 8 and 10 but were rejected because the results significantly weaker.

The linear probability model (LPM) introduces the concept of a probability model, whose dependent is the probability of one of two results and whose coefficients are conditional changes in probability. The LPM uses the same estimation technique as the least squares regression, and maintains the assumption of linearity in the dependent. The forecast range is not limited between zero and one, so the model still falls prey to many of the conceptual difficulties that were found in the simple OLS model—fitted values may be outside the theoretically permissible range. This may be handled by omitting the errant cases or by setting all those below zero to zero and those above one to one.

The LPM also requires a further transformation to correct for heteroscedasticity. As this is now a probability model, its error term follows a Bernoulli distribution and therefore the variance is dependent upon the mean expected value in each case—there is no constant σ because each will vary based upon the probability of the dependent such that the variance is the probability of one times the probability of the other. Like a weighted
least squares regression, the heteroscedasticity can be corrected by dividing the function by the individual \( \hat{\sigma}_i \) terms, which can be estimated in this case as \( \hat{\sigma}_i Y(Y-1) \), where \( Y \) is the estimated probability of binomial AVR. Though the transformation makes the estimates unbiased, the terms are not distributed normally, and are therefore not efficient.

The LPM’s preferred alternative is a logit model, the favorite amongst binary choice models. This model avoids many of the shortcomings of a linear distribution because it instead assumes that the dependent has an asymptotic distribution between 0 and 1. The results fall along a log distributed cumulative distribution function (CDF), while the alternative probit assumes normal distribution. One could speculate over the best functional form of a CDF, this is probably not going to yield a firm conclusion. This study opts for the logit over probit mostly based upon precedent from the only similar discrete choice model for credit rating by Ibarra-Salazar. The logit’s fitted estimate along the CDF is the log of the odds ratio, \( \ln \left( \frac{P}{1-P} \right) = \hat{\lambda}_v \). The antilog of the logit is the probability of being investment grade over that it is not— \( \ln \left( \frac{P}{1-P} \right) = 4 \) means that \( P_i = 0.8 \). The function’s coefficients are not the marginal conditional probabilities for the dependent as they were in linear models. In the logit, the distribution function is nonlinear, so marginal effects must be calculated from second derivatives. Given the value of an independent and the fitted probability an estimated probability and holding all else constant this might be denoted \( \delta E (Avri | Xi) \). This calculation is done by \( \hat{\alpha} \times \hat{P}_i (1-\hat{P}_i) \), the reported coefficient times the estimated variance of each observation. By using the estimated probability, the marginal impact of a given coefficient also involves the other coefficients in the model.

The logit estimation is slightly more complicated because the logit index that is calculated is a ratio of \( P/1-P \). Using individual data means that the probability, \( P \), is either 1 or zero. This raises the possibility that it could be \( \ln \left( \frac{1}{0} \right) \), which would make for an undefined entry in the regressand vector. This precludes the standard least-squares approach; instead, the statistical package uses maximum likelihood estimation with individual-level data. This takes advantage of the explosion in computing power to maximize a likelihood function using a convergence algorithm. This asymptotically tends to efficient estimates in large samples, but even if the sample here may be small, either 30 or 50, one must make an assumption about the data. Accordingly, I maintain that it somehow follows the log distribution and that can be maximized under these conditions (Horowitz and Savin 2001). Though the model seems to work well ex post, it is difficult to justify ex ante as to why this data necessarily ought to fit this logistic function. The estimates must be taken with the necessary caveat that it may not actually be the optimal form by which the probability of an investment grade rating is determined.

Because of the initial uncertainty about what format will work best, it is informative that all three models deliver slightly different results. The OLS model seems to gives a more precise estimate for ratings. Because ratings truly are on a 1-16 point gradient, this additional explanatory power is very attractive. The true aim of a model is to model reality, not just the simplified categories of investment or speculate grade. Still, the probability models are more generalizable and robust: their output does not need to be interpreted on a case-wise basis because the dependent will always be meaningful.
The overall drawback of these models is that they are restricted to binomial variable, investment or not. The fitted values are less descriptive and the estimated coefficients only describe how the determinant can cause large-scale changes. Though there are multinomial probability models that allow for several discrete options, these are not suitable in this study because the dependent is itself so abstract. The multinomial model has a well-studied issue with the “independence from irrelevant alternatives” (Horowitz and Savin 2001). McFadden illustrates this property with a travel-choice example that separates riders of a red bus or a blue bus. The choices are functionally the same, but might be separated in a multinomial probability model anyway. In the presence of such irrelevant alternatives, the results do not actually predict the material considerations of the model. While more density could be added to a discrete dependent model, there could be no way of being sure that further divisions were really material or simply “irrelevant alternatives.” Ibara-Salazar uses an ordinal logit model that avoids this problem by assuming that cardinal differences between discrete dependents may be irrelevant but they can still be ranked. Thus, 1 would be inferior to 2, 11 superior to 10, but without speculating as to the relative values. While this is an attractive framework that gives Ibara-Salazar strong results in assessing ratings on Mexican states, the estimation requires more data than is available for this study. There must still be sufficient observations to calculate probabilities at each level, which is unavailable in this set. As no model is entirely without its flaws, this study considers several in an effort to grasp the whole picture.

In all of the tests, IPD is the crucial independent. It ought to have a negative and significant coefficient if it is systematically being considered in credit ratings. Judging from the coefficients on explicit debt stock that other studies found, a good level of significance might be the 90% level. The actual statistics rely upon estimates by Holzmann et al. (2004) for 35 middle income countries and Roseveare et al. (1996) for 20 in the OECD. Countries estimated by the former study will be known as the Holzmann sample, and the latter as the Roseveare sample. Holzmann et al. use simplifying assumptions to hold constant the length of pension service, and constrain benefits to GDP growth. These simplifying assumptions, however, make the calculations feasible and allow for cross-country comparisons. Roseveare is able to use slightly less restrictive assumptions because governments themselves project future liabilities projections. In both Holzmann and Roseveare, the discount rate is the most important assumption behind the present value calculation. This is supposed to reflect the cost of capital, suggesting that poorer countries ought to have a greater discount. Indeed, Holzmann, working with a set of countries with diverse levels of economic development, discounts the estimates at several interest rates. Though the comparison is interesting, a cross-country quantitative study is better served by a consistent 3% rate. This decision is further justified because one cannot be sure of the source of a high interest rate. A high rate of growth and demand for money could justify a higher discount. Another possibility, inadequate savings, potentially endogenous to the pension policy, would also lift the observed interest rate, though this should not make the IPD seem a lighter burden. Still, the two estimates are not perfectly comparable, most basically because Roseveare work in 1996 and Holzmann in 2004. To account, the other
independents are contemporaneous with the IPD estimate, so that each is an individual
year’s snapshot. Still, inconsistencies in the two estimates motivate the limited-sample
test which look at countries measured by either one or the other. Especially within the
Holzmann set, results seem much more consistent with theory.

Systematic differences between studies may have conflated systematic influences of
the IPD. A brief look at the Portuguese system can be an instructive way to compare them
because both estimate it at different times. Roseveare’s 1996 snapshot estimated 277% of
GDP, while Holzmann figured it at 193% in 2004. Part of this difference is due to reforms
that have taken place in the interim. A series of reforms were undertaken in 2000 and 2002
have reduced the burden by emphasizing means-tested payments and an overall reduction
in the replacement ratio (OECD 2004). In the decade that separates the two studies, the
government has significantly reduced its implicit liabilities independent of other macro-
economic factors. Aside from real changes over time, the two studies also make different
assumptions. By estimating the gross debt of a ‘closed’ program, they try to limit such
assumptions: they do not concern themselves with changes to income flows or new
entrants. Still, their demographic models have a particularly pronounced impact on their
final estimates. Holzmann assumes that longevity remains constant, whereas Roseveare
uses national demographic projections, which may include increasing life-spans, especially
in a country that has developed quickly in the past generation such as Portugal.
Portugal, like many countries, indexes part of its liability to cost of living and other parts
to wage growth in a way that some of the effect of this may actually be regressive. Many
public-sector workers have very generous replacement ratios which are indexed to wage
growth. Roseveare uses this information to project costs while Holzmann uses the median,
which can tend to underestimate the growth. Holzmann does not incorporate the somewhat unexpected political economy effect that, in Portugal, some of the larger pensions
will actually grow faster than the smaller and IPD is actually greater than if both grow in
tandem. Still, this odd arrangement is a political economic peculiarity that may not persist if the politics change. Ultimately, the distinctions become discretionary choices, but
the model’s unexpected results will show that consistency is of utmost importance.

A dummy variable demarcates the presence of a direct contribution alternative to the
standard, DB plan. If pensions are a part of credit liabilities of a country, its form might
also be a material consideration, though it is somewhat presumptuous to predict a sign.
These schemes may be in the form of a PAYG notional direct contribution or publicly
sponsored funded program. This will capture the perceptual differences between plan
designs, though they only reorient the manner by which the liability must be paid and
therefore will not necessarily affect IPD. DB plans, as mentioned, may be falling out of
vogue with financial analysts and are thought to depress savings. However, they also may
be a more flexible liability for the government, and therefore the same level of IPD with a
DB may be seen as less threatening to a government’s other liabilities. Newer, NDCs, cre-
ate “notional accounts,” which have been touted for increasing current payments, but do
not necessarily save this excess funding to cover the future liability. More participation
may even increase the size of IPD and, because NDCs are more closely analogous to bond
liability, a government may discover more resistance if it tries to reduce or relinquish its liability. However, NDCs may also lead to secondary effects on growth and reduce overall payments by altering savings behavior (Hviding and Mérette 1998) and reducing budgetary risk. Though actual performance differences due to plan design may not actually materialize, the dummies will capture perceptional differences.

Led by the rating predictions made by Cantor and Packer and Alfonso, the model controls for other determinants commonly associated with foreign currency ratings. These include: per capita income, GDP growth, inflation, external balance, a measure of total or external government debt, and an indicator of default since 1970. The prior studies of sovereign ratings find per capita income to be the most influential determinant, as richer countries roundly have better ratings. Sovereigns of wealthier countries have more resources to service their debt. Frequently have better developed financial systems from which to draw and would incur more costly damage from a default. Along these same lines, a dummy to denote that the IMF considers the country “developed” is often helpful for a sovereign’s rating. The coefficient on GDP growth is also expected to be positive, as it offers a glimpse at the country’s economic trajectory and speedy growth makes debt easier to cover. A positive external balance helps a country to service its foreign-denominated debt by setting the exchange rate at which the liability is domestically valued. The more relevant consideration is often when an indebted country persistently runs deficits and increases its debt burden as the local currency depreciates on world markets. Larger stocks of public debt, whether domestically or externally held, increase the risk on future debt and should therefore have a negative sign. Though credit markets have been remarkably forgiving of late, a history of default should diminish the rating. These variables alone have been shown to account for over 85% of credit ratings, so the controls are themselves a fairly substantial model.

Regressions and Results

This set of models uses the entire sample of 50 countries, differentiated only by the values on their explanatory variables. After the initial models have been established, the sample is narrowed to try and parse different effects. The initial regression is a near mimic of the reduced-form OLS models in Cantor and Packer (1996) and Alfonso (2002). It is not purely a self-gratulatory indulgence in the stat package’s ability to massage a few vectors into an equation; rather, it is a useful check on the data’s quality. This sample pulls observations from several years and national sources of dubious parole—it is useful to test that the data can replicate prior works. The dependent, AVR, is continuous from 1-16 and determinants include variables which prior research indicates might be good predictors but excludes IPD estimates. The model achieves an R2 of 0.89, which compares well with Cantor and Packer’s 0.92 and Alfonso’s ranging from 0.82 to 0.87.
This model has several interesting facets that foil expectations. The most significant control variable is GDP per capita. Past default was also an important determinant, and most of the other variables are confidently not equal to zero, at least at an 80% level. Though the results are somewhat weak, they are not inordinately worse than prior studies. Standard diagnostics also do not show a specification error. The White Test cannot reject that there is no heteroscedasticity or omitted variable. The n*R-Squared value for this test is 12.6 which is well within any reasonable confidence level on an F-distribution; even at 60% certainty, there is no heteroscedasticity.

Growth and public debt both have the wrong sign and are not very significant. Growth might be explained by its inverse correlation to GDP per capita (-0.48), which is such a strong determinant that it trumps whatever independent influence growth might have. However, in Alfonso and Cantor and Packer the coefficient had the right sign and usually a fair level of significance. Perhaps this is pointing to a flaw in this data: rich, highly rated countries have single-year growth rates from 1996 when global growth was slower than in 2004 when the poorer, lower-rated countries were rated. A better proxy for the trend growth rate of the economy might be an average over a period that spans an entire cycle. Neither prior study used this manipulation, however, and it might be another question entirely as to how much historical context the raters use when making their assessments.

The public debt variable is most interesting, because one must suspect that it displays endogeneity. It is the product of an entirely separate market that sets the quantity of such debt by clearing the cost and demand. The observed effect is that richer countries have a higher demand and lower cost for debt, and thus tend to borrow more, even conditioned as a percent of GDP. This is manifest here as multicollinearity by the correlation of about 0.81
between GDP per capita and public debt. This is only a piece of a circular relationship in which GDP per capita determines much of the credit rating and both rating and income help to determine public debt stock, which confounds debt stock as a regressor for rating. Perhaps the reduced-form function above that uses observed values of public debt does not fully express the structural relationship and the error terms are not entirely stochastic because there is an unreported influence.

The problem might be ameliorated by instrumenting public debt on an exogenous determinant. A potential candidate for this role might be the government bond yield, and this two-staged least squares model follows. To my knowledge, other studies have not attempted this two-staged process, likely because of the difficulty in finding a good manner to demonstrate the full structural relationships. Deferring to the first rule of econometric analysis, that theory ought to drive the conclusions, the two-staged model is more inline with the theory that suggests that public debt does not just appear, but rather is the stock of results from a history of market-clearing transactions. In the market for government debt, higher interest rates move the quantity borrowed backwards along the demand curve. Still, the empirical relationship is not straightforward, for reasons explained earlier. In the first stage, the t-value between public debt and interest rate is only about -1, when a significant coefficient on public debt in the second stage would require a stat closer to 3. As is clear from the second-stage results, the method does little to improve the model’s fit or the significance of the variables. This two-staged least squares should not be seen as a solution to the problem, but rather a first stab at the question.

### Table 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>CURRACTPG</td>
<td>-0.064</td>
<td>0.064</td>
<td>-0.997</td>
</tr>
<tr>
<td>EXDTPX</td>
<td>-0.290</td>
<td>0.211</td>
<td>-1.373</td>
</tr>
<tr>
<td>GDPPCD</td>
<td>0.249</td>
<td>0.053</td>
<td>4.681</td>
</tr>
<tr>
<td>GDPHC</td>
<td>-0.089</td>
<td>0.129</td>
<td>-0.684</td>
</tr>
<tr>
<td>PIC</td>
<td>-0.062</td>
<td>0.040</td>
<td>-1.534</td>
</tr>
<tr>
<td>PUBDTPG</td>
<td>0.012</td>
<td>0.013</td>
<td>0.966</td>
</tr>
<tr>
<td>DEF</td>
<td>-1.761</td>
<td>1.002</td>
<td>-1.757</td>
</tr>
<tr>
<td>DEV</td>
<td>2.229</td>
<td>1.258</td>
<td>1.772</td>
</tr>
<tr>
<td>C</td>
<td>6.138</td>
<td>1.029</td>
<td>5.965</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.895</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Instruments: curractpg, exdtpx, gdppcd, gdphc, pic, pubdtpg, def, dev, inrt

The first incorporation of IPD is a simple OLS model that adds IPD as a percent of GDP and a dummy for a defined contribution scheme. Results are summarized in table #3. They are somewhat disappointing because the IPD variables did not notably improve the fit. Nor are they significant to the extent that one can reject the null hypothesis that their coefficients are equal to zero at any great level of confidence. The estimated coefficient on direct contribution is positive, which suggest some perception bias of direct contribution being more sustainable. This interpretation is strengthened by the near total independence of the two variables. The DC coefficient is not driven by the lower levels of IPD that such a scheme might ultimately encourage. The direct contribution variable is inde-
pendent of IPD at a 90% level of confidence and in this sample it appears even to be associated with more IPD, with a correlation coefficient of -0.23.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>CURRACTPG</td>
<td>-0.060</td>
<td>0.068</td>
<td>-0.881</td>
</tr>
<tr>
<td>EXDTPX</td>
<td>-0.316</td>
<td>0.226</td>
<td>-1.396</td>
</tr>
<tr>
<td>GDPHC</td>
<td>-0.082</td>
<td>0.143</td>
<td>-0.574</td>
</tr>
<tr>
<td>GDPPCD</td>
<td>0.256</td>
<td>0.061</td>
<td>4.197</td>
</tr>
<tr>
<td>PIC</td>
<td>-0.058</td>
<td>0.042</td>
<td>-1.367</td>
</tr>
<tr>
<td>PUBDTPG</td>
<td>0.007</td>
<td>0.014</td>
<td>0.477</td>
</tr>
<tr>
<td>DEF04</td>
<td>-1.377</td>
<td>1.056</td>
<td>-1.305</td>
</tr>
<tr>
<td>DEV</td>
<td>2.731</td>
<td>1.464</td>
<td>1.866</td>
</tr>
<tr>
<td>IPD %GDP</td>
<td>0.149</td>
<td>0.367</td>
<td>0.407</td>
</tr>
<tr>
<td>(IPDPG)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct Contribution</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(DC)</td>
<td>0.587</td>
<td>0.708</td>
<td>0.829</td>
</tr>
<tr>
<td>C</td>
<td>5.353</td>
<td>1.206</td>
<td>4.437</td>
</tr>
<tr>
<td>R-squared = 0.891</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other configurations, using only the externally held government debt in the place of total government debt, excluding the debt stock measure entirely, or various log transformations, do not significantly alter the fit of the data. The R-squared values of all these are similar, while the log transformation drops it considerably.

Despite OLS’ generally good fit, there are several fitted values that are unacceptable in the context of a 16-point discrete scale. As mentioned, Japan’s fitted value is 18.4169, Denmark’s is 17.258 and though none drop below the 0 threshold, the possibility exists. The violation of range constraints coupled with the cumbersome logit of describing a discrete variable as a continuous distribution, leads to the binomial choice models. The dependent, average credit rating, was transformed so that values greater than 9 were 1 and those 9 or below were 0. The model again reverts to the reduced form version, as the logit’s non-linear form precludes the two-staged least squares. The first model, however, keeps the linearity assumption and uses the same regressor variables as in the OLS model that includes IPD. The model now estimates in the language of probability: it predicts the probability of one value over the other. The first is the linear probability model, presented in table 4.
As mentioned, the probability does not have homoscedastic variance, but the estimated $\sigma$ can be calculated as the square root of $[(Y(Y-1)]$, labeled SIGLPM, to perform WLS. However, when the fitted value is less than zero or greater than one, this will entail the square root of a negative number. Such observations are discarded, leaving only 35 remaining in the sample. This key weakness ought not impose a bias on the estimates, as it throws out observations from the high and low end, and not necessarily the most extreme cases either. However, it is a peculiar manner of excluding observations that seems baseless in theory and may distort the estimates. Though the R-squared greater than 95% seems impressive this standard goodness-of-fit measure is not the same as when used in a continuous variable sense. The sum of squared residuals will be smaller, necessarily, because these are fractional values, and squaring the residual does not have the same magnifying impact as in a traditional model’s context. Most of the variables do maintain the correct sign, except for public debt and the direct contribution dummy.

This LPM demonstrates quite clearly why the method has fallen from favor with the proliferation of computer technology that can perform more complicated estimation techniques. These other techniques assume a non-linear functional relationship which asymptotically approaches 0 and 1. The asymptotic CDF avoids the possibility of out of set results, but also imposes a distributional assumption on the data. These assumptions may make both models worthwhile, as there is no a priori reason to believe that the data will take a logistic or linear distribution. The difference between these initial assumptions makes the models basically incomparable, unfortunately, because even a non-nested fit test relies upon the two models having the same basic form.

Additionally, there is no overwhelming feature in the data to dictate a probit or logit model, and both have, in reality, quite similar results. Because these models cannot be compared to their least-squares counterparts I begin from the first steps again. The first two models again exclude IPD variables before completing the expression. Two alternate configurations are tested, one with variables for both total external debt and total public debt and the other has only externally held public debt. The most note-worthy difference

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>CURRACTPG/SIGLPM</td>
<td>-0.020</td>
<td>0.011</td>
<td>-1.849</td>
</tr>
<tr>
<td>EXDTPX/SIGLPM</td>
<td>-0.025</td>
<td>0.043</td>
<td>-0.582</td>
</tr>
<tr>
<td>GDPHC/SIGLPM</td>
<td>0.002</td>
<td>0.023</td>
<td>0.077</td>
</tr>
<tr>
<td>GDPPCD/SIGLPM</td>
<td>0.025</td>
<td>0.011</td>
<td>2.275</td>
</tr>
<tr>
<td>PIC/SIGLPM</td>
<td>-0.010</td>
<td>0.010</td>
<td>-0.946</td>
</tr>
<tr>
<td>PUBDTPG/SIGLPM</td>
<td>0.003</td>
<td>0.003</td>
<td>0.934</td>
</tr>
<tr>
<td>DEF04/SIGLPM</td>
<td>-0.062</td>
<td>0.138</td>
<td>-0.447</td>
</tr>
<tr>
<td>DEV/SIGLPM</td>
<td>0.138</td>
<td>0.289</td>
<td>0.479</td>
</tr>
<tr>
<td>IPDPG/SIGLPM</td>
<td>0.0466</td>
<td>0.045</td>
<td>1.034</td>
</tr>
<tr>
<td>DC/SIGLPM</td>
<td>-0.045</td>
<td>0.108</td>
<td>-0.419</td>
</tr>
<tr>
<td>1/SIGLPM</td>
<td>0.039</td>
<td>0.125</td>
<td>0.312</td>
</tr>
</tbody>
</table>

R-squared: 0.954
is that external debt and public debt both have an unexpectedly positive sign, while external public debt has the expected negative, though none of these estimates can reject a null hypothesis of zero at any significance.

The standard goodness of fit statistics do not apply in the discrete dependent context because the residual will always be less than one, so squaring it will actually have the opposite of the intended effect. Instead, the McFadden R-squared value serves, though it is not strictly analogous to the least-squares R-squared. For this first model the value is 0.836, and the count R-squared, which is simply the ratio of correct fits to total is 0.979. It misspecifies only Chile with a greater than 50% chance of a speculative grade when it is actually observed as investment-grade. The second model has a McFadden R-squared of 0.82 and count R-squared of 0.957. This misses Chile and overestimates Korea. Ultimately, which model is superior seems rather arbitrary—it may only be incidental which are misspecified, especially with only 50 observations. The count R-squared is a rather ‘lumpy’ measure because the difference between one and two makes a considerable difference in the observed statistic but may be the result of inconsequentially small errors in the estimation. While the model with fewer independents might be preferable as a general rule of thumb, theory would suggest that one ought to care both about the total level of public debt and the size of external liabilities that may draw on foreign reserves. The model does not provide any insight into how the ratings agencies actually weight these factors, or if they consider them at all.

IPD variables are added to both configurations and still the goodness-of-fit measures are similar. Again, signs on key coefficients are different: external debt and public debt are unexpectedly positive, but the portion of public debt that is held externally has a negative coefficient.
It is noteworthy that dummies for past default and development have been excluded from the discrete probability models. This is because they are perfectly determinant, which is an interesting result in itself. All countries with a past default are graded as speculative and a country must be deemed developed to qualify as investment grade. When the distinction is drawn between 9 and 10, the current account also has singular covariance. This result may be interpreted as a reverse causation—only countries with strong debt ratings can borrow large amounts and run a large external deficit without the market crimping its borrowing activities by raising rates. This same sort of effect may also explain the positive correlations between debt stock variables and rating. In these variables, they are not determinant because rising debt stocks pose a more creeping problem than the acute deficit of a flow and therefore they elude market notice.

**Tests and Hypotheses**

The most obvious test to undertake is a Wald coefficient constraint test on the OLS equation that includes IPD variables. To see if the two IPD variables have any market impact in the model for credit rating, I simultaneously constrain both of them to zero. The null hypothesis that \( \beta_{\text{IPDPG}} = \beta_{\text{DC}} = 0 \) is set on an F-distribution with 39 degrees of freedom in the numerator (50-11) and 2 in the denominator. This restriction cannot be rejected at the 90% level, indeed there is only a 60% chance that this hypothesis holds. Results are no better if either of the two variables is omitted entirely and the other constrained to zero. With the logit model, the restriction is even less robust, barely significant at a 10% level. The linear probability model yield slightly better results on this test—the hypothesis that \( \beta_{\text{IPDPG}} = 0 \) can be rejected with a 72% level of confidence. Pursuant to Dhaliwal’s study, a further test might check if \( \beta_{\text{IPDPG}} = \beta_{\text{PubDtPG}} \), which would demonstrate whether IPD is considered similarly to other forms of public debt. In both logit and OLS models, the null hypothesis holds with around 75% likelihood. However, interpretation of this is difficult, because the estimates are both positive. So their similarity suggests that they both increase the average rating by the same amount? Really, it may be reflecting that they are both upwardly biased in a similar fashion.

Beyond diagnostic tests, a further step might be to test sample restrictions. The most

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>CURRACTPG</td>
<td>-0.220</td>
<td>0.191</td>
<td>-1.150</td>
<td>-0.297</td>
<td>0.238</td>
<td>-1.251</td>
</tr>
<tr>
<td>EXDTPX</td>
<td></td>
<td></td>
<td></td>
<td>0.287</td>
<td>1.238</td>
<td>0.232</td>
</tr>
<tr>
<td>GDPHC</td>
<td>0.155</td>
<td>0.716</td>
<td>0.217</td>
<td>0.236</td>
<td>0.593</td>
<td>0.398</td>
</tr>
<tr>
<td>PIC</td>
<td>-0.501</td>
<td>0.658</td>
<td>-0.761</td>
<td>-0.487</td>
<td>0.719</td>
<td>-0.677</td>
</tr>
<tr>
<td>EXGOVDTPX</td>
<td>-0.916</td>
<td>3.001</td>
<td>-0.305</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PUBDTPG</td>
<td></td>
<td>0.080</td>
<td>0.389</td>
<td>0.206</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDPPCD</td>
<td>0.555</td>
<td>0.297</td>
<td>1.868</td>
<td>0.591</td>
<td>0.381</td>
<td>1.551</td>
</tr>
<tr>
<td>IPDPG</td>
<td>0.562</td>
<td>1.401</td>
<td>0.402</td>
<td>-0.218</td>
<td>1.447</td>
<td>-0.151</td>
</tr>
<tr>
<td>DC</td>
<td>1.456</td>
<td>2.737</td>
<td>0.532</td>
<td>1.489</td>
<td>2.996</td>
<td>0.497</td>
</tr>
<tr>
<td>C</td>
<td>-6.109</td>
<td>5.996</td>
<td>-1.019</td>
<td>-7.013</td>
<td>5.191</td>
<td>-1.351</td>
</tr>
<tr>
<td>McFadden R2</td>
<td>0.828</td>
<td></td>
<td></td>
<td>0.835</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
obvious of which would be to divide the two observation groups between those whose IPD was estimated by Roseveare in 1996 and those in the Holzmann cohort whose numbers come from 2004. The differences in assumptions may conflate any systematic trends and explain why IPD is so consistently indifferentiable from a random unrelated factor in the combined sample. The differences were discussed in relation to Portugal, but many of the subtle nuances in estimation technique that manifest in that case may also have occurred in many of these other country observations. If IPD in the Roseveare sample were generally higher and this group has, on average, a rating 9 points higher, this would upwardly bias the coefficient on IPD. In reality, the estimates may be incomparable in a unified system.

Though it greatly limits the sample size, the most promising modification seems to be a model restricted to samples in the Holzmann non-OECD sample. The literature suggests that the variables may have a stronger influence because raters seem more sensitive to perceived weaknesses in lower-income/riskier countries. Recall in particular, Alesina et al. found debt stock to be significant for interest rates only in high debt countries, and also the risk ratings studies found that lower income, higher risk countries seemed more closely scrutinized. Even the raters themselves insist that corporate debt must be distinguished along income lines.

The sample limitation also has utility for the regression function because it abrogates some of the difficulty with endogeneity between public debt and credit ratings. Rich countries enjoy good ratings, low interest rates, and easy access to debt, so that they demonstrated a positive correlation between debt and rating. Generally, it is costly for poor countries to run up debt (or has been in the past) and therefore the desired causal relationship whereby debt increases credit risk dominates the correlation. On the margin, higher debt may always diminish the credit rating, but without panel data, this study must look at a single level, which has an unexpected relationship in the total, pooled sample. In the Holzmann group, however, public debt is perceived to increase credit risk and seems to successfully cull out the desired relationship by a manipulation based in theory.

The OLS and logit models are presented in table 7, though LPM had to be omitted because the data set becomes too small. In the process of making the weight factors, observations are shed whose predicted values fall outside of the 0-1 range and when the sample is only 30 observations, such a loss is unacceptable. Within the OLS and logit frameworks of this sample, all of the coefficient signs meet expectations except for GDP growth and external debt though neither is very significant.
The “developed” dummy again must be omitted from the logit model, and the developed dummy is omitted from both as Korea is the only country in the sample with a value of one. Notably, in both of the models, public debt and IPD have negative coefficients. The direct contribution dummy seems to have a strong influence in this group, especially in the OLS model. Though the goodness-of-fit measures are not as strong as the combined sample, they are not inordinately weaker. In fact, the count R-squared of the logit is 0.96, which is similar to the large sample models. Less explanatory power might be expected because this cohort is probably more idiosyncratic, which could confound the normal systematic relationships and make it more difficult to estimate.

Because the logit is non-linear there is no single marginal effect, rather it is different at each level calculated as \( \partial \pi / \partial x \), yielding a marginal response curve. According to this logic, Argentina had a fitted probability of 0.0208 (calculated by \( e^{-3.85} \)), so the marginal influence of an additional 100% of pension debt is to decrease the probability of an investment grade rating by 0.43%. In 2004, its pension debt reduced the probability by ~0.34%. According to the OLS model, pension debt equal to 100% of GDP drops the rating by 0.2 on a 16 point scale, which translates into a rating 0.15 points lower in Argentina’s case. Brazil and Macedonia both lose a by half a point on their rating due to the marginal effect of their exorbitant pension deficits.

Now that the coefficients have the ‘right’ signs, one may make meaningful comparisons. By inspection, it seems that IPD is treated with far less severity than the rest of the public debt stock. This is immediately apparent in the OLS model, where marginal effects are constant and therefore marginal effects are simply the partial slope coefficients. The coefficient is probably going to be smaller on IPD because it is perceived as a more forgiving liability. Still, it is interesting that their magnitudes are so different in both the OLS and logit model configurations. The robustness of this observation can be tested, as with the full sample models, by the Dhaliwal-inspired, constrained F-test that explicit debt is truly weighted significantly heavier than the implicit counterpart. This restriction cannot be rejected and the F-stat is really very low, 0.37 in OLS and 0.49 in the logit. Unfortunately, the standard errors on both of these coefficients are so large that the failure
to reject the null hypothesis is really not very meaningful, the hypothesis that they are both simultaneously equal to zero is even more likely, around 80% in both configurations. A nested-model F-test does not dismiss a null hypothesis that would omit either of these variables outright. Though it is gratifying that their signs are both negative in this estimate, there is little justification to make meaningful comparisons. One would expect that IPD is smaller, but even intuitively, the huge disparity in their coefficients is dubious. Explicit public debt is not eighty-times more influential to a country’s propensity to default on its loans. The most reasonable conclusion might follow that neither variables for debt stock is consistently a hindrance to the sovereign’s rating, but total public debt is more likely to be considered and probably has a stronger impact. Perhaps this confirms the initial suspicion that ratings agencies are not considering implicit liabilities, even in poorer countries where they may care about the overall debt stock.

If we suspend doubts momentarily, one might use trends from the Holzmann sample to make some interested projections about the possible nature of IPD. Assuming that raters were completely ignoring IPD, the coefficient on public debt would be -2.69, and the full results are below.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>CURRACTPG</td>
<td>-0.022</td>
<td>0.091</td>
<td>-0.238</td>
</tr>
<tr>
<td>EXDTPX</td>
<td>0.128</td>
<td>0.580</td>
<td>0.222</td>
</tr>
<tr>
<td>GDPHC</td>
<td>-0.331</td>
<td>0.221</td>
<td>-1.495</td>
</tr>
<tr>
<td>GDPPCD</td>
<td>0.502</td>
<td>0.113</td>
<td>4.461</td>
</tr>
<tr>
<td>PIC</td>
<td>-0.061</td>
<td>0.044</td>
<td>-1.404</td>
</tr>
<tr>
<td>PUBDTPG</td>
<td>-2.690</td>
<td>2.394</td>
<td>-1.124</td>
</tr>
<tr>
<td>DEF04</td>
<td>-1.675</td>
<td>1.111</td>
<td>-1.508</td>
</tr>
<tr>
<td>C</td>
<td>6.056</td>
<td>1.713</td>
<td>3.536</td>
</tr>
<tr>
<td>R-squared =</td>
<td>0.760</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The average fitted rating according to this model is 5.13, but if IPD is added with half the weight of explicit public debt, the sample’s average drops to 3.536 and if they are considered equally detrimental, the average value drops to 1.942. In the OECD sample, the average predicted rating is 14.496, but when IPD is discounted at the rate that public debt is discounted, it drops to 12.671. Because the logit model is non-linear, the marginal effect of IPD must be calculated from a point. Using the mean values from the Holzmann sample, the probability of being investment grade is 15.8%. The probability decreases by 4.25% when the mean IPD value, 116% of GDP, is factored in at the same rate as public debt.

**Conclusions**

At its most basic, this study was testing the primary economic assumption of rational, informed actors. Granted that there is relevant information that exists are actors systematically incorporating it into their decisions? It ought to have some sort of long-term relationship with risk and returns, so a ‘rational’ actor ought to inquire. In the end, regular
consideration of IPD is limited because the cross-country comparable estimates are really quite difficult and imprecise. If the data exists, it is not widely disseminated, and IPD’s hypothetical influence is not widely understood. It is a useful analogy to compare implicit debt to the better-understood explicit liabilities, but even this position is not a given. Still, those that would argue that IPD may be painlessly defaulted without disrupting the public’s finances are probably stubbornly holding their head in the sand. As the popular presses will attest, the majority of program participants have come to depend upon their benefits to such an extent that default would destroy thousands of lives. Further, a country’s financial credibility is not so segmented that one set of creditors does not worry about the fate of others. IPD is certainly more important than its dubious relationship to credit rating would imply. The mixed result from this model should not suggest that IPD is a not important determinant of sovereign bond debt. Instead, it points to a serious shortcoming in the raters’ effectiveness. There is a need to reform not only the pension systems but also the way they are assessed and considered. The initial projections would indicate that raters have been lax at identifying oncoming threats. Though it is never welcome to be the doomsayer, blithe inaction can, as oft demonstrated, only bring reckoning. As Reisen and Maltzan (1999) have suggested, when raters fail to check overoptimistic investment, bond markets experience undue volatility. Markets have been experiencing exceedingly low yields, the Citigroup World Government Bond Index has returned -7% over the last year, while the EMBI+ is only returning 2.3% better than a US treasury. Perhaps this unusual environment is being driven by an endemic underestimation of the risks associated with implicit liabilities. If history is any guide, this may bring a speedy flight of capital at the first concrete sign of danger, when IPD begins to threaten the global liabilities.

Before jumping to drastic conclusions, there is certainly room for more work in the area. Some of it may be completed in the near term, while others require a larger time horizon and more empirical study. An easy goal for future study is to find better instruments for two-staged least squares that untangle some of the endogeneity that is conflating the relationship of public debt stock and credit rating. Longer-term research might take advantage of a larger set of more comparable IPD estimates. And, perverse as it seems, research would greatly benefit from an empirical case of IPD-driven debt crisis in which explicit liabilities are in direct conflict with implicit ones. Political pressure is everywhere encouraging some degree of IPD reduction, but efforts have not yet had the urgency of a complete inability to pay. Perhaps some of the difficulty with systematically incorporating implicit liabilities into overall credit risk is because the threat is still only hypothetical. When some countries make the desperate calculus between their pensioners’ wellbeing and explicit default, the raters will have a better guide for how to incorporate IPD. Prior work in pension benefits suggest that this point might be nearing quickly, but it seems quite clear that the financial sector’s watchdogs are not adequately prepared.
References:


Ibarra-Salazar, Jorge, Gabriela Garcia-Romo, Lida Sotres-Cervantes. “Determinants of Mexican States


## Data Appendix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current account balance in percent of GDP</td>
<td></td>
</tr>
<tr>
<td>GDP per capita in USD PPP adjusted</td>
<td>IMF World Economic Outlook Database. Data current to September 2005, national accounts vary</td>
</tr>
<tr>
<td>Percent growth in real GDP</td>
<td></td>
</tr>
<tr>
<td>Percent change in GDP price-deflator</td>
<td></td>
</tr>
<tr>
<td>External debt as a percentage of exports in trade-weighted terms</td>
<td>World Bank database of World Development Indicators. Total due to nonresidents payable in foreign currency, goods, or services</td>
</tr>
<tr>
<td>General government external debt as a percentage of exports</td>
<td>World Bank database and national sources for countries: Dominican Republic, Bolivia, Nicaragua, New Zealand, Philippines, Morocco, Iceland, Macedonia, Malta, Moldova, Mauritius, Slovakia, Senegal, Romania</td>
</tr>
<tr>
<td>Public and publicly guaranteed debt as a percentage of GDP</td>
<td>IMF World Economic Outlook, World Bank World Development Indicators, Standard and Poor’s Sovereign Risk Indicators: 2005, and national sources for Germany, Malta and Korea</td>
</tr>
<tr>
<td>Developed country dummy</td>
<td>World Economic Outlook data appendix A</td>
</tr>
<tr>
<td>Past default dummy</td>
<td>Standard and Poor’s courtesy of Frank Packer</td>
</tr>
<tr>
<td>Average government bond yield</td>
<td>Derived from figures in Standard and Poor’s Sovereign Risk Indicators: 2005 and World Bank’s World Development Indicators Database</td>
</tr>
<tr>
<td>Implicit pension debt</td>
<td>Studies by Holzmann et al. (2004) and Roseveare et al. (1996)</td>
</tr>
<tr>
<td>Defined contribution dummy</td>
<td>National Sources and Impavado 2002</td>
</tr>
<tr>
<td>Credit Rating</td>
<td>Unweighted average from Standard and Poor’s, Moody’s and Fitch Ratings</td>
</tr>
</tbody>
</table>

### Numeric Conversion for Ratings

<table>
<thead>
<tr>
<th>AAA</th>
<th>AA+</th>
<th>AA</th>
<th>AA-</th>
<th>A+</th>
<th>A</th>
<th>A-</th>
<th>BBB+</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>15</td>
<td>14</td>
<td>13</td>
<td>12</td>
<td>11</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>BBB</td>
<td>BBB-</td>
<td>BB+</td>
<td>BB-</td>
<td>B+</td>
<td>B</td>
<td>B-</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
Impact of Workers’ Remittances on Developing Countries’ Growth

Salifou Issoufou
University of Wisconsin-Madison

ABSTRACT
Remittances have been subject to a consistent scholarly scrutiny both at the theoretical and empirical level. Theoretically, the emphasis has been on the reasons why migrants remit and the most emphatic reason advanced in the literature is the migrant’s altruism towards his or her family that remains in the country of origin. Empirically, significant research has focused on the economic impacts of remittances, whether positive or negative, on the development of recipient countries.

After suggesting an alternative explanation of what takes place between the migrant and his family regarding the reasons to migrate and those to remit, an empirical model is presented designed to investigate two hypotheses on which this study reposes. The first hypothesis is that remittances positively impact economic growth in developing countries because of their use as private investments in productive ventures. The second hypothesis states that the impact of remittances on growth is through their favorable contribution to the improvement of educational attainment in developing countries.

The estimation results weakly refute the first hypothesis and strongly support the second hypothesis. In fact, an increase in remittances as percentage of GDP by one percentage point is associated with a 0.05 percent economically insignificant decrease in the growth rate of real per capita GDP. When interacted with the measure of educational attainment, the impact of remittances on growth through education is an average of 1.38 percent increase in the growth rate of real per capita GDP associated with every percentage increase in education.

Introduction
Remittances have become increasingly important in recent years both to researchers and to families receiving them. In aggregate terms, remittances mainly go from devel-
oped to developing nations. The amount has increased on a steady trajectory to the extent that it has recently been estimated to be more than the aggregate amount of foreign direct investment received by developing countries (Neyapti, 2004). This increase in the flow of funds has spurred research that analyzes the causes, uses and effects of remittances.

Literature on the causes of remittances is abundant and has given birth to theories explaining why immigrants remit. These theories revolve around three main characteristics: altruism, selfishness, and tempered altruism. Altruism stresses that the migrant remits because he cares about his family left in the native country. The term first appeared in Johnson and Whitelaw (1974). This idea of altruism evolved into theories like those of pure altruism, self-interest, and tempered altruism developed by Lucas and Stark (1985).

Self-interest emphasizes the selfishness of the immigrant remitting only to increase his investment portfolio. The third theory is a combination of the first two in that it pertains to not only the altruism of the migrant but also his desire to invest in his native country. All these theories have, for foundation, Becker’s (1974) theory of social interaction in which family members care about one another to the extent that they find it optimal to maximize a joint utility function.

In addition to the study of the causes of remittances, there has also been much research that looks at the uses of remittances by receiving families. Most, if not all, of these studies are based on surveys of families that receive remittances conducted in a specific country. The results of these studies point to a consumption oriented use of remittances with very few indicating expenditures in schooling and investment in productive ventures as alternatives. Examples of these studies include Stahl and Arnold’s (1986) review of numerous studies of countries in which they concluded that a large portion of remittances goes into expenditures on basic necessities. Another example is the recent study by Yang (2005), using surveys results conducted in the Philippines, which concluded that remittances increase funding for schooling and indigenous entrepreneurship in recipient households.

Although there is an abundance of literature on the causes and uses of remittances, investigations of their effects on developing countries’ economies has been sparse. The main reason for this is the lack of sufficient aggregate data on remittances since most of the financial transactions, between the immigrant and his family, take place through a network of friends, rather than formal institutions like Western Union and banks in order to reduce transaction costs. Added to this is the existence of challenges in the collection of remittance flow in countries receiving the funds. These two factors make the collection of data on remittances flow difficult and incomplete, thus underestimating the real magnitude of the flow. Despite the scarcity of these data, a few researchers have managed to investigate the effects of remittances on economic growth of developing countries.

This scholarship has, however, suffered from a narrow focus. In fact, most of the studies have focused on single-country empirical analyses to measure the impacts on remittances on the country’s economy. These studies use household surveys as datasets and apply the multiplier effect methodology to assess the economic impact. Recently, there has been a drive to look at cross-sections of countries in analyzing the effects of remittances.
An example of this approach is the study by International Monetary Fund economists Ralph Chami, Connell Fullenkamp and Samir Jahjaj. In a recent working paper, the three researchers advanced a theory of causes of remittances based on the altruism of the immigrant and trace its effects through the impact of individual recipient labor effort on economic growth. They found a negative impact of remittances on economic growth using a panel of 113 developing countries (Chami et al, 2003). However, Chami et al.’s research and other studies of this sort leave out two aspects of remittance analysis. First at the theoretical level, the family of the migrant can also be altruistic towards the migrant in addition to the migrant being altruistic towards the family. Second, at the empirical level, the impact of remittances on economic growth can be expressed through the interaction between remittances and educational attainment.

This study addresses these ignored aspects or remittances analysis. It follows the methodological approach developed by Chami et al. (2003) but takes a different direction. First, a new interpretation of the causes, uses, and effects of remittance, rational altruism, is suggested as an alternative to existing views on remittances. Second, an interaction model is used to investigate the relationship between remittances, combined with educational attainment, and growth. In addition, the impact of remittances on growth is investigated by introducing a dummy variable designed to analyze whether larger remittances as share of GDP matter or not. This indicator takes the value of 1 for log of remittances as percentage of GDP greater or equal to 5 percent, and 0 otherwise. It is worth indicating that the choice of remittances as percentage of GDP of 5 percent as point of demarcation is totally arbitrary.

Specifically, the study revolves around two hypotheses. The first hypothesis is that remittances positively impact economic growth in developing countries because of their use as private investments in productive ventures. The second hypothesis states that the impact of remittances on growth is through their favorable contribution to the improvement of educational attainment in developing countries. Both hypotheses are investigated using a generic long-run growth model. This model treats the growth rate of real per capita GDP of 65 developing countries over the period 1970-2000 as a function of economic variables, already established as determinants of growth, plus remittances. What is new in this study is therefore twofold: First, is the suggestion of a new interpretation of the causes and uses of remittances; second, is the introduction of a non linear model, in the form of interaction between education and remittance variables and the introduction of a dummy variable to see whether having larger remittances makes a difference.

The estimation results weakly refute the first hypothesis and strongly support the second hypothesis. In fact, an increase in remittances as percentage of GDP by one percentage point is associated with a 0.05 percent economically insignificant decrease in the growth rate of real per capita GDP. The coefficient on log of remittances, - 0.05, is statistically insignificant. When interacted with the measure of educational attainment, the impact of remittances on growth through education becomes greater and significant at 10 percent level (assuming normality). This impact is an average of 1.38 percent increase in the growth rate of real per capita GDP associated with every percentage increase in edu-
cation when log of remittances as share of GDP takes the value of – 0.37 (which is the average of the natural log of remittances to GDP for the 65 countries sampled in this study). Introducing the dummy variable measuring the scale of remittances returns positive results for countries with remittances to GDP greater or equal to 5 percent and negative results for countries with remittances to GDP less than 5 percent. The coefficient on the dummy variable is statistically insignificant, however.

The remainder of this study has the following structure: chapter one discusses the current status and magnitude of remittances, chapter two reviews the existing literature on remittances, chapter three presents the methodology, chapter four presents and discusses the results, and the last section summarizes the study.

I. Understanding Remittances

A. Conceptual Issues

Reported under the current account in the balance of payment of a given country, remittances are part of financial transactions associated with migration. In fact, remittances are one of three migration related components of the balance of payment statistics. The other two are compensation of employees and migrants’ transfers. The latter are sometimes mistaken for remittances because of the transfer connotation. This denominational issue involving remittances was discussed at the World Bank’s International Technical Meeting on Measuring Remittances. At this meeting, Reinke and Patterson (2005) of the International Monetary Fund specified the different definitions of remittances as well as the other two migration related components of the balance of payment. The definitions are as follows:

Workers’ remittances cover current transfers by migrants who are employed in new economies and considered residents there. A migrant is a person who comes to an economy and stays there, or is expected to stay, for a year or more. Workers’ remittances often involve related persons (Reinke & Patterson, p. 3).

Compensation of employees comprises wages, salaries, and other benefits earned by individuals – in economies other than those in which they are residents – for work performed for and paid for by resident of those economies (Reinke & Patterson, p. 3).

Migrants’ transfers are contra-entries to the flow of goods and changes in financial items that arise from the migration of individuals from one economy to another (Reinke & Patterson, p. 3).

The report specifically stressed the difference between workers’ remittances and migrants’ transfers on the one hand, and the difference between compensation of employees and workers’ remittances on the other. The first difference lies in the fact that workers’ remittances are current transfers involving residents from different countries, whereas migrants’ transfers are part of changes that occur in the capital account because of the change of residence of a given household. The second difference is more obvious in the sense that workers’ remittances and migrants’ transfers are, as their names indicate, transfers, and compensation of employees are just recompenses for work performed.

This study will use workers’ remittances in investigating the impact of financial flows
from migrants on economic growth of destination countries.

**B. Magnitude**

Remittances are important sources of financial in-flow for countries that export migrants. In 2000, the amount of remittances recorded worldwide was $123 billion of which the least developed countries (according to the United Nations’ definition) received $20.2 billion. Whether it is at the worldwide or regional level, the amount of remittances received by countries that export migrant workers has been steadily increasing.

The amount of remittance that goes to labor exporting countries is not evenly distributed, as one might expect. In 2000, India, Mexico, and the Philippines respectively received $12.74 billion, $6.57 billion, and $0.125 billion while Malawi only received 5.3 million of the 123 billion recorded worldwide (all in 2000 $ U.S.). In percentage, the three countries (India, Mexico, and the Philippines) accounted for 16.80 percent of the $123 billion world remittances. The share of workers’ remittances for the three countries was estimated to be one-third of 2004’s aggregate workers’ remittances (Neyapti, 2004).

Aside from these three countries having a large number of their citizens living abroad, another reason for the prevalence of high flow of remittance is the relative ease with which this in-flow of financial transaction is recorded. In fact, it is at least known that remittance going into Mexico benefit from an improved system that allows migrants to remit through formal channels. For example, there is a new method that allows immigrants from Mexico living in the U.S. to make mortgage payments on new houses in Mexico (these schemes are not available to other immigrants living in the U.S.).

Although the figures are seemingly large, the actual amount is underestimated because of the difficulties associated with recording migration related transactions. Indeed, data compilation of workers’ remittances presents tremendous challenges. The compilation is done by central banks of different countries and then transmitted to the International Monetary Fund’s statistics department for publication. Many difficulties emerge with this approach to data gathering of remittances such as how a given country defines his nationals living abroad. This definition varies across countries. In addition, developing countries encounter more problems because of the lack of necessary tools for tracking remittances data. Therefore the numbers should be interpreted with caution.

Given this increasing trend in migrants’ related financial flow, researchers have studied why immigrants remit, analyzed how recipients use the money, and investigated the effects of remittances on the economies of receiving countries. Results from these studies are many as well as various. Therefore, to understand the different views pertaining to the reasons why immigrants remit, how these remittances are used, and how their effects can be traced, an extensive review of literature that has looked at the different angles of the issue is necessary. This review is presented in the following chapter.

**II. Review of Existing Literature**

Since the mid-seventies, workers’ remittances have become an important part of supplying countries’ balance of payments and a hot subject of research for many scholars from a variety of disciplines ranging from sociology to economics. As a result, the abun-
dance of studies that focus on different aspects of remittances is notable. These studies are divided into two major areas: the microeconomic level and the macroeconomic level of understanding the causes, uses, and effects of remittances. In fact, some researchers have put forth theories on the causes of these remittances as well as empirical models, while others have investigated their use, using surveys, and their effects on the overall receiving economy, using empirical analyses. The purpose of this chapter is to review some of the studies that have contributed to a better understanding of remittances and galvanized scholarly efforts for further inquiry. In the first subsection studies on the causes of remittances are reviewed; in the second, an inquiry into the scholarship view on their usage is taking; the last subsection presents the review of studies that covered the overall effect on the receiving economy.

A. Causes of Remittances

1. Theories on the Causes of Remittances

The reasons why many immigrants send part of what they earn, while working in their host countries, back to their countries of origin have been subjects to scholarship scrutiny. This scrutiny has engendered enunciation of theories of the causes of remittances. The most famous of these theories is one that states altruism as the main reason why migrants remit. There are, however, additional theories that were put forth which include pure self-interest and tempered self-interest, among others, as alternative explanations to why migrants remit.

Altruistic motivations as the bases for why migrants remit appeared in studies of remittances of the mid-seventies when it was stipulated that the primary cause of remittances is pure altruism, wherein the migrant cares about his extended family left in her country of origin and therefore remits in order to financially assist them. An example of these studies is that of Johnson and Whitelaw (1974), who mentioned altruism as the driving force behind remittances. Subsequently many studies have taken this direction in theorizing about the causes of remittances. This theory on the causes of remittance is a derivation of Becker's (1974) theory of social interaction in which family members care about one another and are therefore better-off if they all maximize a collective utility function. Using a survey of migration conducted in Botswana, Lucas and Stark (1985) found results not in support of pure altruism. One explanation the two authors gave stems from the complexity of a given family and the possibility that pure altruism alone may not explain the array of causes of remittances. It is in this spirit that alternative theories emerged.

One of these theories pivots around self-interest involving not only the migrant but also her family. This theory of self interest was developed by Lucas and Stark (1985) as an alternative to the pure altruism theory. It is two-dimensional with a tempered side called “the enlightened self-interest” and the extreme part or the “pure self-interest.” Either theories (or sub-theory) entail the migrant and his family but with different level of agreements.

The enlightened self-interest theory is explained as the cause of remittances being governed by contracts between the migrant and his family that remains in the country of origin (Lucas and Stark, 1985). The reasons behind the contracts could be manifold. But one
that stands out is the fact that remittances are a result of risk sharing or investment in access to larger streams of income. With this logic, the family – playing the role of financial intermediary – agrees to invest in the migration process and in return the migrant, once in host country or wherever she is migrating to, remits so as to repay the family (Horowitz, 2002; Gubert, 2002). Because of higher earnings in the place of immigration, the funds remitted are assumed to be larger than what was initially invested, so remittances take the form of redistributive gains. With this scheme, one might wonder why the migrant would abide to the contract once he settles in his place of destination. The answer to the question lies in one of the assumptions of the theory, which states that the migrant returns home after a certain period. For the sake of her dignity when returning home, she would respect the contract. The tempered self-interest theory also suggests that the migrant stops remitting once the contract is broken. Using the same migration survey from Botswana, Lucas and Stark (1985) tested the enlightened self-interest theory and the results corroborated their theory.

When formulating the pure self-interest theory, Lucas and Stark (1985) provided three reasons why a migrant would only remit out of a simple and selfish concern for himself and not his family. The first reason provided is the migrant’s ambitions to inherit from his family. Remittances in this case are geared towards making sure the family keeps the migrant as the favorite of all possible contenders. With the assumption that behavior governs inheritances, the authors concludes that the larger the remittances, the bigger the expected amount to be inherited by the migrant. The second reason given is that the migrant is interested in investing in his country of origin with the guarantee that the investments would be efficiently managed by an intermediary, whom in most cases is a member of the family. The targets of the investments may vary from physical or fixed assets such as land to buying livestock. Lucas and Stark (1985) also indicate the possibility of investment in the education of migrants’ children that stay in the migrant’s place of origin or in the migrant’s own prestige by supporting public works such as building religious structures. The last reason provided for the pure self-interest is the migrant’s intention to return home in which case he would make investments in fixed assets or other types of investments similar to the ones mentioned in the second reason. In all the three reasons why a self-interested migrant only remits for his own sake, the family that stays in the migrant’s place of origin still plays an important role in his decision. Therefore the theme of involvement of family ties in the theories of causes of remittances appears to be pervasive.

In addition to the theoretical analyses of the causes of remittances, studies have dealt with the empirical analysis to determine the forces that drive the motivations to remit. The review of this is presented in the following subsection.

2. **Empirical Analyses of the Causes of Remittances**

Studies that have investigated the determinants of workers’ remittances proceeded using as basis the different theories of the causes of remittances described in the preceding section. Elbadawi and Rocha (1992) summarized the empirical research on the causes of remittances and grouped most of these empirical analyses into two categories: the
The endogenous approach rests on an empirical framework which considers that economic and demographic conditions faced by the migrants and their families affect the decision to remit. The following are the main determinants of migrants’ remittances: composition of the family at home and abroad, the level of disposable income or wage in the host country, the income differentials, and the anticipated length of stay (Elbadawi and Rocha, 1992). Testing the various empirical models that follow this framework, many results were presented. For example, Swamy (1981) found that the ratio of female to total labor population of each nationality has a negative and significant impact on the flow of remittances, while the level of wages and the number of migrants in the host country positively correlated with remittances. Another example is the study by Elbadawi and Rocha (1992) in which, remittances are modeled as a function of certain demographics, income, and some macroeconomic factors. The authors applied the model to six labor exporting countries (Algeria, Morocco, Portugal, Tunisia, Turkey, and the former Yugoslavia) and found that remittances were correlated with income cycle in host countries and to stock of migrant workers.

The portfolio approach to empirical analyses of the causes of remittances revolves around the migrant’s decision to either save (or invest) in the host country or to remit to her home country. In this case the relative rate of return, the relative price and uncertainty are said to determine the migrant’s decision to remit. Generally, the models put forth have for dependent variable the flow of remittance and the different independent variables could take the form of incentives that labor exporting countries put in place in order to attract remittances, and the difference between the informal exchange rate and the official exchange rate, among others. Migrant supplying economies in fact give preferential treatment geared at luring in remittance flow. For example, a labor exporting country could set an exchange rate specifically for remittances that is better than the official exchange rate. Examples of empirical analysis within the portfolio framework are many.

Swamy (1981) presented a model of the portfolio approach which has remittances flow as a function of: the interest rates on foreign currencies deposits in the sending country relative to the interest rate on comparable maturity deposit in the receiving countries, rate of real return on real estate in the home country relative to comparable rate of real return on bank deposits in the receiving countries, the difference between the black market exchange rate and the official exchange rate in the home country, and the difference between preferential exchange rate for remittances and the official exchange. He then uses data from Greece, Turkey, and the former Yugoslavia and found that remittances were not affected by the indicated factors.

Katseli and Glytsos (1986) model has remittances as a function of migrant’s wealth, relevant interest rate and exchange rate variability. They tested the model using data from Greek migration and found that per capita remittance is positively related to per capita income and interest rate in the host country, while negatively related to inflation rate in...
home country.

When the reasons why migrants remit are theorized and empirically tested, it is natural to wonder how remittances are used once they arrive in the countries of destination and what are their effects on the economies of receiving countries. This part of the study of remittances is reviewed in the next section.

**B. Uses and Effects of Remittances**

1. **Studies of the Use of Remittances**

Research on remittances’ use posits that the primary role of these financial flows to households with migrants is consumption expenditures, followed by savings and investments. The first use of remittances follows from the view that recipients are in crucial financial needs and therefore direct the funds received to alleviating those needs by buying basic necessities. These necessities cover food and clothing, but some of the consumption items such as jewelry do not fit the “necessity” cataloging. As for savings and investments, they tend to vary with studies and regions. The most common types of investments are in housing and livestock.

There are a multitude of studies that looked at the pattern of remittances use in receiving countries. These studies tend to focus on a specific country or region of a country known for its emigrational importance. Surveys of households with migrants are the basis of most, if not all of the studies.

Stahl and Arnold (1986) presented and extensive review of numerous country studies conducted to determine the use of remittances. One conclusion from the review of studies of six Asian countries (Pakistan, Bangladesh, India, Thailand, Philippines, and Sri Lanka) is the similarity across countries in terms of expenditure pattern. In fact, a big portion of remittances in all six countries goes to expenditure on basic needs, while some of the funds are used in housing (which involves construction, home improvement, and so forth) and debt or mortgage repayment.

Using data from four linked household surveys conducted by the National Office of the Philippines government, Dean Yang (2005) found that remittances received by households increase the schooling of children and boost indigenous entrepreneurship. In a different continent, Hanson and Woodruff (2003) used a sub sample of the 2000 Mexico Census of Population and Housing. This sub sample provides data on the number of households that have migrants residing abroad and remittance received by a given household. The two authors found that a significant part of these remittances go into the education of children in the households, in comparison to households that do not receive remittances. The studies by Yang (2005) and Hanson and Woodruff (2003) bring about another area where remittances funds are channeled. Investment in education has, indeed, not received enough attention from scholars studying the use of workers’ remittances. This is part of the controversy surrounding the study of the effects of remittance on the overall economies of receiving countries, which is reviewed in the next subsection.

2. **Effects of Remittances**

There are two extreme views that dominate the study of the impacts of workers’ remittances on the economies of receiving countries: The negative view, also known as the
“Dutch Disease” or the “Migrant Syndrome,” and the positive one. The negative view is supported by scholars who believe in the detrimental effects remittances have on their recipients as individuals and on the economy as a whole. Individuals receiving remittances are believed to become dependent and reduce their labor effort, while the economy is being negatively affected by this shortage of labor and by the non productivity of investments that remittances go to. On the opposite side of the spectrum, is the positive view which is sponsored and promoted by the New Economics of Labor Migration (NELM). NELM argues that remittances that result from migration create conditions for a positive long-run impact on the development of receiving countries. The reasons given are that remittances alleviate the harsh conditions set by the imperfections of financial markets by providing funds for multiple productive investments that could lead to development or growth in receiving countries.

Support for the negative impact of remittances on long-run growth comes from studies such as that of Kozel and Alderman (1990), in which a labor force participation and labor supply study of Pakistan was performed using 1986 survey data compiled by the Pakistani Institute of Development Economics (PIDE). They found a negative and significant impact of remittances on the labor force participation of males. Using a broader approach, Chami et al. (2003) modeled the growth rate of GDP of receiving countries with workers’ remittances as one of the independent variables. Operating under the assumption that remittances occur under asymmetric information and economic uncertainty, creating the problem of moral hazard, the authors used a panel of large sample of developing countries to test the hypothesis that remittances negatively impact economic growth. Their results corroborate their hypothesis that remittances negatively affect economic growth through reduction of labor effort by the recipients.

Additional arguments in support of the negative view of remittances are based on studies on the use of remittances, which found a large portion going into consumption. This consumption is directed towards non traded goods, and since these goods have a relatively inelastic supply the result is inflationary pressure which is not good for growth (Stahl and Arnold, 1986). Also, scholars in support of the negative view of remittances argue that the effects of remittances, even if positive, do not outweigh the negative effects of loss of income due to a decrease in tax revenues resulting from emigration (Bagwati, 1974). They believe remittances do to not outweigh the negative effects created by the fact that governments are investing in the education of their citizens and are not getting the return on their investment in education when the citizens leave. This has led scholars to argue that the decrease in return on education investment result in the decrease in human capital attainment and therefore a decrease in economic growth. The negative effect could be lessened if migrants return with additional knowledge.

As for the positive view of the impact of remittances on economic development in receiving countries, there are studies that have investigated these effects using the NELM framework. The basic argument in support of this view came with the introduction of the NELM, and is formulated as workers’ remittances being a stimulus to development dynamics by lessening investment and production constraints faced by economies with
imperfect financial markets. The argument continues with the assertion that governments could further these positive effects of remittances by creating favorable environments wherein remittances are used in productive ways. Studies in this domain are multiple and range from ones that use the multiplier effect, which focus on the short run effect of remittances, to those that show the impact of remittance on educational attainment.

In a study of the effects of remittances sent to Mexico, Adelman and Taylor (1990) found that every remittance dollar is associated with an increase, ranging from 2.69 to 3.1, of Mexican GNP. The two authors also found that the effect of remittances is more important in rural areas because rural residents spend their income on locally produced items rather than on imported items. This is not the case with urban areas where goods that are consumed are mainly imported. These findings confirmed the belief that even though remittances funds directly go into consumption, “many of the benefits of remittances accrue to households other than the ones that receive the remittances” (Taylor, 1999 p. 70).

Adams and Page (2003) investigated the impact of remittances on poverty reduction in a cross section of countries. The two authors found a negative and significant effect of remittances on poverty. Specifically, they found that a 10 percent increase in remittance, defined as share of a country’s GDP, leads to a 1.6 percent decline in poverty.

Recently, two studies that investigated the effects of remittances on educational attainment, which in turn impacts economic growth, were conducted. The first study, conducted by Hanson and Woodruff (2003) focused on the relationship between remittances and educational attainment in Mexico using data from the 2000 Mexico Census of Population and Housing. They found that children from households that receive remittances complete more education than those from households that do not receive remittances. Specifically, their OLS estimation reported an extra 0.23 years of schooling associated with a household that has a migrant leaving in the U.S. (that also remits). Instrumental variables estimation produced higher results ranging from 0.73 to 0.89 years. The second study is by Yang (2005) who empirically analyzed the effects of remittances using data from household surveys conducted by the National Statistics Office of the Philippine government. Yang found that remittances have a positive income effects on households. This contributed to more entrepreneurial activities at the local level and to more investment in the education of children. The investment in education resulted in more schooling and less child labor.

These two empirical studies follow the NELM approach to the study of the effects of remittances. However they focus on a specific country and use household surveys as their data source in analyzing the effects of remittances on economic growth. There is a limited number of empirical investigations that take a generalized approach to analyzing the impacts of remittances using a cross section of countries making the repertoire of workers’ remittances studies not well rounded. This study tries to add to this limited repertoire by investigating the impact of remittances on developing countries following a unified approach. It first provides a theoretical framework, which is a synthesis of the existing theories, for understanding the causes and uses of remittances and then specifies an empirical model for tracing the effects of remittances on the economic growth of receiving coun-
tries. This part of the study is presented in the following chapter.

III. Methodology

This chapter first presents a new interpretation of the causes and uses of remittances, and then specifies an empirical model for tracing the effects of workers’ remittances on the economic growth of developing countries before discussing the data to be used in the analysis.

A. New Interpretation of the Causes and Uses of Remittances

The interpretation presented in this section flows from the NELM approach to remittances and development, which stressed the positive influence these funds have on the development of receiving countries. It is also an improved explanation of the “tempered altruism” perspective developed by Lucas and Stark (1985), a theory that has for basis the various contracts, governing the decision to migrate and the one to remit, which the migrant and her family have to respect. The reasons for the improvement lie in the fact that the tempered altruism theory does not preclude the risk for dependency that arises when remittances are received on a constant basis. This dependency is one of the causes that lead to the reduction of the labor force participation by recipients, which negatively affects growth. To alleviate this loophole, I propose rational altruism. In this rational altruism, altruism is bidirectional (or reciprocal), meaning that not only does the migrant care about the family left behind but also the family left in the native country cares about the welfare of the migrant.

First, there are a few assumptions that need to be presented before explaining the basics of the suggested interpretation. The first assumption is related to the family structure and some of the values that the family and the migrant abide to. It is assumed that the family has school aged children living in the household and both the migrant and the remaining family left in the home country value work, education, financial independence and care about one another. The second assumption is the existence of symmetric information between the migrant and his family, key to avoiding the moral hazard issue associated with sending remittances. The third and very important assumption is the existence of imperfect markets for credits limiting the family’s access to borrowing money in addition to the existence of child labor in the country (the market imperfection assumption was first stipulated in the NELM, and the child labor assumption is a new addition). Lastly, I assume a continuous flow of migration from economically disadvantaged countries to those that are more developed. With this in mind, the decision to migrate, the one to remit, and the use of remittances can be explained.

Under harsh economic circumstances, rational altruism leads a family to look for better ways of improving its wellbeing. One way to do this is for a family member to individually, or from a collective family decision, decides to move to a place where he could have a higher return to his labor. Once in the country of immigration, the migrant, knowing that his family back home is living under bad economic conditions and lack of access to credit markets where they can borrow money, decides to regularly remit part of his income because he cares about the welfare of his family. When the remittances are
received, the family experiences relief from the lack of credit markets and directs the money towards various projects. Although there is a strong possibility of expenditure on consumption, this type of expenditure helps the receiving family in relinquishing the pressure of limited access to food and clothing. A big part of the remittances received is invested in entrepreneurial projects and in the schooling of children. Because the family values financial independence, work, and education, and because it cares about the migrant, the money is used for productive investment such as education and project with high return so as to improve the schooling of their children and to become self-sufficient. This availability of funds will allow children to be focus on education instead of work while the adult family members continue working toward a greater improvement of their economic status because they value work. Because of symmetric information and reciprocal altruism, the receiving family will use the funds adequately and the migrant will remit until he knows that the family has improved their living conditions. In the case where the migrant intentions are to return, he may continue remitting in order to create the conditions for a good and economically stable return.

With this investment in the schooling of children and in productive enterprises emerges the improvement of the country’s overall economy. In fact, growth theory tells us that private investment and education attainment are good determinants of economic growth. This allows us to trace the effects of remittances on economic growth using an empirical growth model.

The difference between this interpretation and the existing theories on the causes and effects of remittances is twofold. First rational altruism stresses the bidirectional aspect of altruism. Indeed, previous theories only stress the altruistic nature of the migrant ignoring the family receiving remittances and in this explanation the receiving family is also altruistic towards the migrant. Second, this interpretation assumes the existence of symmetric information between the migrant and the family left in the native country as opposed to the asymmetric information characterizing existing theories.

It is worth indicating that even though this interpretation covers the causes and uses of remittance, the only part that is investigated by this study is their effects on the overall economies of different developing countries. The choice for developing countries is based on the fact that developed countries have strong credit markets and are not characterized by the practice of child labor. The model for estimating the impacts of remittances on growth is presented in the following section.

**B. Empirical Model**

Most studies that looked at the impact of remittances on economic growth used the generic growth model. This model has the growth rate of real GDP (per capita or not) as a function of a series of independent variables known to be determinants of growth including remittances. For example, Chami et. al. (2003) extended the generic growth model to remittances by setting the growth rate of real per capita GDP as a function of the initial level of GDP, the log of investment to GDP ratio, and the log of workers’ remittances to GDP. Specifically, Chami et. al.’s model is of the following form:

$$\Delta Y_i = a + bY_{oi} + c\ln I_i + d\ln WR_i + U_i$$
Where \( Y \) is the log of real GDP per capita, \( Y_0 \) is the initial value of \( Y \), \( I \) is the log of investment to GDP ratio, and \( WR \) is the log of workers remittances to GDP. With this model, the authors estimated the relationship between a change in workers’ remittances and growth using cross-sectional analysis. In addition, they perform a panel estimation including the ratio of net private capital flows to GDP as another explanatory variable. The two authors also used a two-staged least square as a way to deal with endogeneity. Their sample is composed of 113 countries (of which only 50 have no less than 10 years of continuous observation) and the period of analysis is 1970-1998.

This study uses the generic growth model similar to the one in Chami et. al. (2003) for the empirical analysis. However, there are differences in the choice of independent variables, sample size, period, and the estimation techniques. In this study, more explanatory variables are used with a smaller sample size (65 developing countries) and a larger period (1970-2000). It is worth indicating that the sample used in this study has 58 countries with 20 or more observations over the period of 1970-2000.

The impact of workers’ remittances on growth is investigated using the following equations:

\[
\begin{align*}
(2.1) \quad G_i &= \alpha_i + \beta R0_i + \chi Ivi + \delta Pi + \varepsilon Ni + Ui \\
(2.2) \quad G_i &= \alpha_i + \beta R0_i + \chi Ivi + \delta Pi + \varepsilon Ni + \phi LRi + Ui \\
(2.3) \quad G_i &= \alpha_i + \beta R0_i + \chi Ivi + \delta Pi + \varepsilon Ni + \phi LRi + \phi Ni*LRi + Ui \\
(2.4) \quad G_i &= \alpha_i + \beta R0i + \chi Ivi + \delta Pi + \varepsilon Ni + \phi LRi + \phi Si + Ui \\
(2.5) \quad G_i &= \alpha_i + \beta R0i + \chi Ivi + \delta Pi + \varepsilon Ni + \phi LRi + Ui 
\end{align*}
\]

(Note: Definitions of the different variables (dependent and independents) as well as the data sources are presented in Table 1.)

The first equation is the baseline growth model used in many studies of long-run economic growth. Equations (2.2) – (2.5) are extensions of the baseline equation developed to investigate the hypotheses that remittances, as private investments and through education, positively affect growth. In equation (2.2), the impact of remittances on economic growth is estimated using the log of remittances to GDP (LRi) as explanatory variable. The third equation, 2.3, includes the interaction between education variable (Ni) and the log of remittances as share percentage of GDP (LRi). As for equation 2.4, it uses an indicator variable (Si) measuring the scale of remittances to see whether having a larger remittances as percentage of GDP has an impact on economic growth. The last equation, 2.5, is introduced to see whether the results of the regression from the sample in this study match those of Chami et. al. (2003).

The results of the estimation of the baseline equation are expected to confirm the existing growth theory. This means that the coefficient on the initial level of real per capita GDP, R0i, is expected to be negative and so is the coefficient on the average population growth rate. Investment (Ivi) and education (Ni) are keys to economic growth and their coefficients are expected to be positive. As for the equations extending the baseline, the estimation results are expected to support the hypotheses that remittances impact econom-
ic growth in developing countries first as private investments in productive ventures and second as their positive contribution to educational attainment. Consequently, the coefficient on LRi is expected to be positive. In addition, the coefficient for the interaction between remittances and education, and the one for the indicator are expected to support the two hypotheses.

### Table 1. Variables: Definitions and Sources

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definitions and Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gi</td>
<td>Growth rate of real per capita GDP (in 1996 US $), averaged over 1970-2000. Source: Penn World Table</td>
</tr>
<tr>
<td>Ivi</td>
<td>Natural log of Investment as share of GDP (in 1996 US $) averaged over the period 1970-2000. Source: Penn World Table</td>
</tr>
<tr>
<td>R0i</td>
<td>Real per capita GDP (Chain) measured at the beginning of the period, in 1970 (in 1996 US$). Source: Penn World Table</td>
</tr>
<tr>
<td>Ri</td>
<td>Remittances as percentage of GDP, calculated using the following formula: (Remittances/GDP)*100. Remittances and the GDP are both from the World Development Indicators, World Bank. Remittances from the WB were in current dollar and the author converted them to 2000 US constant dollar. GDP was already in constant 2000 US $ from the WB.</td>
</tr>
<tr>
<td>LRi</td>
<td>Natural log of remittances as percentage of GDP (Ri).</td>
</tr>
<tr>
<td>Si</td>
<td>This is a dummy variable referring to the scale of remittances. It takes a value of 1 for Ri of 5.00% and more and a value of 0 for Ri below 5.00%. 5.00% is an arbitrary demarcation.</td>
</tr>
<tr>
<td>Ni</td>
<td>Natural log of Net enrollment rate (%), primary, averaged over 1970-2000. This is the proxy for measuring human capital. Source: The Net enrollment rate is from Education Statistics – EdStats, World Bank and natural log is from the author’s calculations.</td>
</tr>
</tbody>
</table>

### C. Data

Data on remittance are from the World Development Indicators yearly published by the World Bank (WB). Remittances are reported by countries that are member of the International Monetary Fund (IMF) and the WB in current US $ and are first published in the IMF Balance of Payments Statistics Yearbook. Reporting countries have different national accounts base year and certainly use different methods of conversion from national currency into US dollar. This makes data on remittances very shaky.

For the purpose of this study, yearly data on remittances from 1970 to 2000 of 65 countries were gathered and then converted into 2000 US dollar using the US Consumer Price Index (base year: 2000) as deflator. The deflated data were then averaged over the period 1970-2000. In addition, the deflated data on remittances were divided by their respective GDP (in 2000 constant US price) and multiplied by 100 giving the percentage of remittance to GDP, Ri. A natural log of this ratio was taken then averaged over the period 1970-2000; the product of this calculation is represented by LRi. It is from these two variables
that the indicator variable, $S_i$, was created.

Of all the 65 countries used in this study, India has the largest total remittances received over the period 1970-2000. India received $137.11$ billion followed by Egypt with $114.25$ billion, and Portugal with $113.79$ billion. As percentage of GDP, this ranking is different. Indeed, Cape Verde takes the lead in this category with remittances accounting for 21.61 percent of GDP. Cap Verde is followed by Burkina Faso with 13.2 percent and Morocco with approximately 8.9 percent.

Data for education and population growth variables are also from the WDI. The education variable, $N_i$, is the 31-year average of the natural log of the primary net enrollment ratio. The dependent variable $G_i$ and the independent variables $R_0i$ and $I_vi$ are the only variables with data from the Penn World Table.

Basic statistics from the data used in this study are presented in Table 2.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>StDev</th>
<th>Minimum</th>
<th>Median</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_i$</td>
<td>1.59</td>
<td>1.72</td>
<td>-2.43</td>
<td>1.61</td>
<td>6.36</td>
</tr>
<tr>
<td>$R_0i$</td>
<td>2637.00</td>
<td>2002.00</td>
<td>332.00</td>
<td>2018.00</td>
<td>9265.00</td>
</tr>
<tr>
<td>$P_i$</td>
<td>2.16</td>
<td>0.79</td>
<td>-0.09</td>
<td>2.27</td>
<td>3.23</td>
</tr>
<tr>
<td>$I_vi$</td>
<td>2.46</td>
<td>0.55</td>
<td>0.82</td>
<td>2.60</td>
<td>3.43</td>
</tr>
<tr>
<td>$N_i$</td>
<td>4.26</td>
<td>0.35</td>
<td>3.22</td>
<td>4.40</td>
<td>4.59</td>
</tr>
<tr>
<td>$R_i$</td>
<td>2.36</td>
<td>3.70</td>
<td>0.00</td>
<td>0.73</td>
<td>21.61</td>
</tr>
<tr>
<td>$L_R_i$</td>
<td>-0.37</td>
<td>1.95</td>
<td>-4.72</td>
<td>-0.01</td>
<td>3.04</td>
</tr>
</tbody>
</table>

The values of importance here are the mean and median of the growth rate of real per capita GDP ($G_i$), the natural log of net primary enrollment ($N_i$), and the 2 variables representing remittances ($R_i$ and $L_R_i$). The means for these variables are respectively 1.59, 4.26, 2.36, and -0.37.

**IV. Results and Discussion**

A sample of 65 developing countries was used to run an ordinary least square estimation assuming normality. The results from the different regressions are presented in Table 2. The coefficients of the variables already known as determinant of economic growth will first be discussed before assessing the meaning of the coefficients on the remaining equations.
As expected, the coefficient on the start off real per capita GDP (R0i) is negative (-0.0004 for the baseline equation) and statistically significant at a 1 percent level using a two-tailed test in all the equations estimated. This result lines with the previous findings such as those of Barro (1991), which confirm the conditional convergence theory stemming from the neoclassical growth theory addressed by Solow (1956). It means that in the 65 developing countries used in the OLS estimation, the starting level of real per capita GDP is negatively related to the growth rate of real per capita GDP.

The coefficient on investment is positive (approximately 1.5 in the first four equations) and also statistically significant at the 1 percent level using a two-tailed test and assuming normality. This result confirms the theory of endogenous growth that suggests the co-movement of the investment ratio and per capita growth (Barro, 1990). It also stresses the importance of investment in developing countries when it comes to improving their respective economies. In fact, the coefficient indicates that, on average, for every percentage increase in the change of investment as share of GDP, the growth rate of real per capita GDP is expected to increase by 1.5 percent.

---

Table 3. Estimation Results

<table>
<thead>
<tr>
<th>Equations</th>
<th>(2.1)</th>
<th>(2.2)</th>
<th>(2.3)</th>
<th>(2.4)</th>
<th>(2.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>Gi</td>
<td>Gi</td>
<td>Gi</td>
<td>Gi</td>
<td>Gi</td>
</tr>
<tr>
<td>Constant</td>
<td>-3.88 (2.35)</td>
<td>-3.66 (2.40)</td>
<td>-4.32 (2.44)</td>
<td>-4.32 (2.45)</td>
<td>-2.72 (0.83)</td>
</tr>
<tr>
<td>R0*</td>
<td>-0.41 (0.09)</td>
<td>-0.42 (0.10)</td>
<td>-0.43 (0.10)</td>
<td>-0.41 (0.10)</td>
<td>-0.23 (0.10)</td>
</tr>
<tr>
<td>Ivi</td>
<td>1.46 (0.36)</td>
<td>1.44 (0.36)</td>
<td>1.38 (0.36)</td>
<td>1.40 (0.36)</td>
<td>2.00 (0.36)</td>
</tr>
<tr>
<td>Pi</td>
<td>-0.81 (0.24)</td>
<td>-0.83 (0.24)</td>
<td>-0.82 (0.24)</td>
<td>-0.80 (0.24)</td>
<td>--</td>
</tr>
<tr>
<td>Ni</td>
<td>1.10 (0.58)</td>
<td>1.07 (0.59)</td>
<td>1.25 (0.60)</td>
<td>1.18 (0.60)</td>
<td>--</td>
</tr>
<tr>
<td>LRi</td>
<td>--</td>
<td>-0.05 (0.08)</td>
<td>1.44 (1.14)</td>
<td>-0.11 (0.10)</td>
<td>-0.01 (0.09)</td>
</tr>
<tr>
<td>Ni*LRi</td>
<td>--</td>
<td>--</td>
<td>-0.35 (0.26)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Si</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0.60 (0.51)</td>
<td>--</td>
</tr>
<tr>
<td>R–Sq.</td>
<td>48%</td>
<td>48.3%</td>
<td>49.8%</td>
<td>49.4%</td>
<td>34.1%</td>
</tr>
<tr>
<td>R–Sq. (adj)</td>
<td>44.6%</td>
<td>43.9%</td>
<td>43.9%</td>
<td>44.2%</td>
<td>30.9%</td>
</tr>
<tr>
<td>SER</td>
<td>1.28</td>
<td>1.29</td>
<td>1.28</td>
<td>1.28</td>
<td>1.42</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
</tbody>
</table>

Note: The numbers in parentheses represent the standard errors of the coefficients.
* The coefficients and the SE of coefficient for R0i were multiplied by 1,000. Therefore, when interpreting they should be divided by 1000.
The relationship between population growth and economic growth stems from the standard neoclassical growth theory, which asserts that if all countries are in steady state, there should be a one-for-one effect of population growth on the growth rate of GDP and the effect is less than one-for-one if either capital accumulation or labor force growth cannot keep pace with population growth. The effect is negative when growth is measured in per capita terms rather than just the GDP growth rate (Nelson and Singh, 1998). In this study, the growth rate of real GDP is in per capita terms therefore the negative sign of the average annual population growth rate coefficient (-0.81 from the baseline model) supports the argument. This coefficient is also significant at the 1 percent level using a two-tailed test and assuming normality. Since the coefficient is less than one, developing countries’ capital accumulation or labor force is not keeping pace with population growth, at least within the sample used in this study.

The human capital proxy used in this study is the natural log of net enrollment rate at the primary level (Ni). From the baseline equation, the coefficient on this variable is 1.10 and is statistically significant at 5 percent level, using one-tailed test and at 10 percent level using two-tailed test. This implies that on the long run, every percentage increase in the net enrollment rate at the primary level is associated with an average of 1.10 percent increase in the growth rate of real per capita GDP. This results lines with previous findings showing the positive correlation between education and economic growth. Later in this discussion, the effects of the interaction between education and remittances on growth will be discussed.

Equation 2.2 is an extension of the baseline model and is design to assess the effects of changes in remittances as share of GDP on the growth rate of real per capita GDP. Entered as natural log of remittances as percentage of GDP, the variable LRi has a coefficient of -0.05. This coefficient is statistically insignificant. The value of the coefficient implies that when remittances as share of GDP increase by one percent, the growth rate of real per capita GDP is expected to decrease by 0.05 percent. This change does not have a strong economic importance since the average per capita GDP in this sample is 1.59. With this average, a change from 1.59 to 1.54 for every unit increase in remittances as share of GDP becomes barely noticeable.

Even though the coefficient on remittances does not have a strong economic influence, it is worth stressing its sign and what that entails. Recently, Chami et. al. (2003) found a negative relationship between remittances and growth using a sample of 113 countries. The negative coefficient in this study lines with these findings and hence implies a negative impact of remittances on growth. Equation 2.5 was designed to test whether using the same explanatory variables as the ones in Chami et al. would produce a different result. This difference would then signal that the findings of Chami et. al. suffer from omitted variables bias since there are only 3 explanatory variables in their equation (replicated by equation 2.5) as opposed to 5 in equation 2.2 in this study. The coefficient on remittances from equation 2.5 is still negative and statistically insignificant. The only difference is that with more explanatory variables the coefficient on remittances becomes bigger and hence has a stronger effect. Despite the statistical insignificance of the coefficient on log of
remittances as share of GDP, its negative sign supports existing literature’s claims that remittances negatively impact economic growth in recipient countries. The negative sign also refutes the first hypothesis of this study, which asserts that remittances play the role of private investments in productive ventures to positively impact economic growth in recipient countries.

Other results of importance are from equation 2.3, which represents the effect of the interaction between education and remittances on growth in the 65 countries sampled in this study. This model tests the non-linearity of the effect of remittances on per capita real GDP growth rate, which means that these effects are expected to be stronger when interacted with education. Specifically, the second hypothesis of this study is that the effect of education on growth depends on remittances. OLS estimation of this model returned results with the point estimate of the interaction being equal to –0.35. This coefficient is statistically significant at the 10 percent level, using a one-tailed test and assuming normality.

The result for the effect on the growth rate of real per capita GDP of a change in the education variable (Ni), holding the remittance variable (LRi) constant is the following: 1.25 – 0.35LRi. This implies that the impact of one percentage increase in the education variable is a 1.38 percent increase in the growth rate of real per capita GDP when the log of remittances as share of GDP takes the value of –0.37, the sample mean. When compared with the effect of education on growth without interaction, this result becomes stronger because without interaction the effect of a percentage increase in the education variable results in a 1.10 percent increase in the growth rate of real per capita GDP. What this entails is that remittances do have an impact on growth through education in the 65 developing countries sampled here. In other words, education is improved by the flow of remittances. This result supports the second hypothesis stressing the importance of remittances to families who receive the funds. Hence, these remittances constitute a source of funds that are being invested in the education of the recipient families’ children.

The other side of the interaction effect is whether remittances’ impact on economic growth is affected by the education variable. When taking this into account, the result is the following: 1.44 – 0.35Ni. Therefore, a 1 percent increase in remittances as share of GDP will be associated with a 0.05 percent decrease in the growth rate of real per capita GDP when the log of net primary enrollment rate (Ni) takes the value of 4.26, which is the sample mean for Ni. Hence when education is taken into consideration, the impact of remittance in the 65 countries sampled in this study is negative. This points to the following idea: Of the 65 developing countries, those that have a better funded education experience a negative impact of remittances on economic growth.

The last analysis of the impact of remittances undertaken in this study looks at whether the impact of changes in remittances is dependent on their size or scale. To check this, an indicator variable, Si, is used. This variable takes the value of 1 for countries with remittance as share of GDP greater or equal to 5 percent and the value of 0 otherwise. The results presented by equation 2.4 report an effect of changes in remittances as share of GDP of 0.49 percent for countries in the high end as opposed to -0.11 percent for the low
end countries. Therefore, remittances have a positive impact on growth in countries with larger remittances as share of GDP and a negative impact in countries with low remittances as percentage of GDP. These results are, however far from being statistically significant and the difference is very infinitesimal to be accounted for. Therefore the data from this study’s sample do not strongly support the idea that the scale of remittances plays a large role in their impact on economic growth.

Overall, the results stemming from this study weakly refute the first hypothesis and strongly support the second hypothesis. In fact, changes in the growth rate of remittances in the 65 developing countries sampled are negatively related to the growth rate of real per capita GDP. This relationship, however, is statistically insignificant. In addition, the nonlinear relationship between remittances and growth returned results indicating that the effect of education on the growth rate of real per capita GDP, when remittances are taking into account, is larger than without the interaction. When education is considered, remittances positively and significantly affect growth in the 65 countries sampled in this study. Furthermore, developing countries with larger remittances as share of GDP experience a positive impact of remittances as share of GDP on growth, whereas countries at the lower end experience a negative impact. These last results are statistically insignificant, however.

V. Summary and Conclusion

This research’s focus was to test two hypotheses: First, that remittances positively impacts growth because of its role as private investment; second, remittances’ positive contribution to economic growth in developing countries is manifested through its interaction with education.

Before testing the two hypotheses, a look at the recent issues surrounding remittances and their magnitude was presented. Also, an extensive review of literature on the causes, and uses of remittances as well as their effects on the economy of receiving countries preceded the methodology used in this research.

Because the existing literature overlooked some of the theoretical aspects of the causes and uses of remittances, and because of the narrow focus that characterizes the empirical analysis of the effects of remittances on economic growth, a new interpretation of what might be the causes, uses, and effects of remittances was suggested and new empirical specification were presented.

The suggested interpretation, rational altruism, stresses the reciprocity of altruism in the relationship between the migrant and his family left in the native country as opposed to the unilateral altruism which permeates existing theories. The empirical specification was designed to investigate the relationship between changes in remittances as share of GDP and changes in the growth rate of real per capita GDP on the one hand and to address one aspect of remittance analysis left out by the existing literature on the other.

In fact, the relationship between remittances and economic growth is not necessarily linear, which is why an interaction between the remittances and education variables was introduced to test whether the impact of education on growth depends on remittances. In
addition, an indicator variable was introduced as a way of accounting for the scale of remittances and seeing whether larger remittances are of any significance compared with smaller remittances.

The first hypothesis is weakly refuted by the results from the OLS estimation. In fact, every increase in the growth rate of remittances as share of GDP is associated with a 0.05 percent decrease in the growth rate of real per capita GDP. This effect is economically small and statistically insignificant. The second hypothesis is strongly supported by the OLS results. Indeed, the interaction coefficient is statistically significant and implied that when remittances are held constant; education positively affects growth, and this effect is greater than the impact of education on growth without the interaction. Also, the results show that where education better funded, remittances negatively affect growth. The last investigation pertained to whether larger remittances to GDP matter or not. Dummy regression results, although statistically insignificant, reported a positive impact of remittances on growth for countries with high remittances to GDP (greater or equal to 5 percent) and a negative impact for countries with low remittances to GDP.

For further studies of this kind, a few steps could be taken to enhance our understanding of the reasons why migrants remit, estimate the true magnitude remittances, and to analyze their relationship with economic growth. First, one could design a multi-country survey asking immigrants living in countries traditionally known as primary sources of remittances why they remit and about their relationship with their families of origin. The results of such survey could be used to determine the reasons why immigrants send money to their native countries and also assess whether the migrant is aware of how the money is being spent. Second, the World Bank and the International Monetary Fund, in collusion with countries that are members, should strive to create uniformity in the methodology for data collection on remittances. This will make data on remittances more reliable and the research more rounded. Third, different estimation techniques using the most recent data available could be used to analyze the impact of remittances on the countries’ respective economies.
IMPACT OF WORKERS’ REMITTANCES

References


FORECASTING MEXICO'S INFLATION: THE EFFECTS OF AN INFLATION TARGETING REGIME

Juan Carlos Suárez Serrato
Trinity University

I Introduction

This paper studies the effects of the implementation of an inflation targeting (IT) regime on the dynamics of inflation in Mexico. We contextualize the need for a regime change in Mexico and study whether its implementation made a considerable change in the dynamics of inflation. Using an inflation forecasting model we test for a regime change and analyze the variability of the residuals before and after the institution of an IT regime.

Inflation plays a role in the growth of the economy and the development of the quality of life of its inhabitants. High inflation influences the exchange rate by affecting purchasing power, reduces the incentive to saving, increases investor uncertainty, affects the level of unemployment through the Philips curve and, as Martínez Trigueros (1998) describes, increases socioeconomic inequality. Mexico has one of the largest inflation-linked issuances of bonds in the world, which require certainty in the level of prices to attract investors. Although Mexico has enjoyed comfortable levels of inflation in recent years, an accurate forecasting model and understanding of Mexican inflation is desirable for different sectors of the economy. In this paper we analyze some of the causes of inflation, build a tractable model to forecast inflation, and study the effects of the institution an inflation targeting monetary regime.

II. Data Considerations

We use quarterly data encompassing the period 1982Q2—2005Q2. The Indice de Precios al Consumidor (pt) is used as the main price level series for the study. The wage series (wt) emanates from two sources. These are spliced in a manner preserving continuity and the ensuing data are smoothed to adjust for excess seasonal variation. We use the Consumer Price Index (CPI) in the U.S. and multiply it by the exchange rate to generate a version of the import price series (imt). This procedure mirrors Bailliu et al. (2003), who find the inclusion of the exchange rate in the model to be highly desirable. Finally, we use a Hodrick-Prescott filter to create a proxy for the GDP gap (gapt).

Fourth-order Augmented Dickey-Fuller tests for unit root in the price level, wages, import prices and petrol price series reveal the level of integration of each variable. Petrol
prices are integrated of order 1, while the price level, import prices and wages are integrated of order 2. Since the variables are all in log form, inflation is integrated of order 1. “Co-integration analysis helps clarify the long-run relationship between the variables” (Ericsson and de Brouwer (1998)). A Johansen test for co-integration on a vector autoregressive (VAR) model shows at most one co-integrating vector exists. The co-integrated relation is estimated consequently.

III. Long-Run Relationship and Single Equation Dynamic Specification

We estimate the existing long-run co-integrated relation using an economic model based on the markup model. Our analysis is not confined to the markup model as our interest spans the effects an IT regime on the short run dynamics as well as the long run dynamics. We define the error-correction term (ecmt-1) as the residuals of the estimation of this model. Henceforth, the regressions are adjusted for seasonality and periods of crises, economic shocks, and extreme outliers by the inclusion of corresponding dummies in the regression. The single equation dynamic specification we estimate is the following:

$$\Delta p_t = \alpha_1 + \alpha_2 \Delta w_t + \alpha_3 \Delta i_m + \alpha_4 \Delta p_t + \alpha_5 \text{ecm}_{t-1} + \alpha_6 \text{gap}_t + \sum_{i=1}^{3} \beta_i S_i + \sum_{i=1}^{6} \gamma_i D_i + \epsilon_t,$$

where $\epsilon_t$ denotes the error term. Separate joint significant tests confirm the relevance of the seasonal dummy variables and the economic shock-dummy variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta$</td>
<td>0.028</td>
<td>0.004</td>
<td>6.46</td>
<td>0.000</td>
</tr>
<tr>
<td>gapt</td>
<td>0.253</td>
<td>0.102</td>
<td>2.482</td>
<td>0.015</td>
</tr>
<tr>
<td>$\Delta w_t$</td>
<td>0.286</td>
<td>0.101</td>
<td>2.82</td>
<td>0.006</td>
</tr>
<tr>
<td>$\Delta i_m$</td>
<td>0.209</td>
<td>0.042</td>
<td>4.986</td>
<td>0.000</td>
</tr>
<tr>
<td>$\Delta p_t$</td>
<td>-0.035</td>
<td>0.013</td>
<td>-2.729</td>
<td>0.007</td>
</tr>
<tr>
<td>ecmt-1</td>
<td>-0.257</td>
<td>0.029</td>
<td>-8.72</td>
<td>0.000</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.913</td>
<td></td>
<td>F-statistic</td>
<td>59.097</td>
</tr>
<tr>
<td>T (adjusted)</td>
<td>93</td>
<td></td>
<td>Prob(F-statistic)</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Note: Dummy and seasonal variable omitted for brevity

Table 1 presents the results of estimating the ECM model specified in Equation 1. We conduct an omitted variable test on $\Delta p_t$, a dummy for the IT years, and the interaction of this dummy with every other term in the model. The test shows the omitted variables are jointly statistically significant and should be included in the regression. The omitted variable test is equivalent to the Chow Test for a regime change. Table 2 presents the results of this test.
This concludes the first part of our analysis. That is, we have shown that in the years of IT, Mexican inflation follows a statistically different model. The second part of the analysis inquiries further into the properties of the residuals of the regression including the IT variables.

Figure 1 shows the actual-fitted-residual graph of the regression including the IT variables. The distribution of the residuals through time is of concern. The variance of the residuals does not seem constant and several outliers suggest the residuals are not normally distributed. Moreover, the residuals appear to have a significantly lower variance during the IT years. White’s heteroskedasticity test and a Jarque-Bera test for normality reject the null hypotheses of homoskedasticity and normality respectively. Finally, estimating the same model on the sub-sample of the IT years yields normally distributed and homoskedastic residuals. During the IT years, the variance of the residuals of the forecasting model is of a statistically different nature. Thus, we find it necessary to estimate a separate model for this sub-sample.

**IV. Conclusions**

The results presented in this paper show the institution of an IT regime in Mexico had drastic effects on the dynamics of inflation. Namely, a Chow Test for regime change shows the existence of a different model of inflation in Mexico during the IT years. Similarly, we
find the variance of the residuals of the forecasting model is significantly different after
the regime change and appeal for a completely different model for these years. This
process yields a tractable and efficient forecasting model for Mexican inflation.

Results presented here advocate in favor of the notion that the regime change in
Mexico had economy-wide proportions. Most importantly, the Bank of Mexico gained
credibility by successfully curbing inflation. However, as Corbo and Schmidt-Hebbel
(2001) suggest, “ultimately, it has been the combination of progress in achieving fiscal dis-
cipline and the discretion embedded in IT that made possible the success in inflation
reduction.” Similarly, one should take into consideration other factors that have been at
play in these time periods, such as political stability.

References

Bailliu, J., D. Garcs, M. Kruger, and M. Messmacher (2003): “Explaining and Forecasting Inflation in

Bank of Chile 105, Central Bank of Chile.

Papers rdp9510, Reserve Bank of Australia.

investigación no. 9806, Banco de México.

Unpublished working paper, Trinity University, available at http://works.bepress.com/jc_suarez/1/
MEASURING THE WEALTH EFFECT OF HOUSING PRICES:
A REGIONAL APPROACH

John Nesbitt
Georgetown University

ABSTRACT

I empirically determine the relationship between housing prices, other capital gains (namely stock market wealth) and consumption. I use a panel of 360 metropolitan statistic areas from 1975 to 2003 and estimate the marginal propensity to consume out of household wealth as well as other capital gains. I test whether the two are significantly different. I find a 9% marginal propensity to consume out of household wealth. This MPC out of household wealth is larger than that of past literature and significantly higher than the marginal propensity to consume out of other capital gains. I find the marginal propensity to consume out of capital gains and disposable income to be consistent with results of past literature.

1. Introduction

The Permanent Income hypothesis and the ‘Wealth Effect’ state that consumers base consumption decisions on the discounted value of their life-time income. It is clear that consumers increase consumption given a permanent increase in income or a promotion at work; however, it is unclear how consumers respond to short and intermediate term shocks in capital gains. According to the permanent income hypothesis, shocks to capital gains, such as stock market returns or property values, affect current consumption only in their affect on a consumer’s life-time income. Therefore, a one year increase of 10% in the value of a consumer’s home will not substantially affect consumption because it will not substantially affect a household’s lifetime income.

Many analysts would argue that the strength of the US housing market is responsible for strong consumer spending and the consequent sustained growth in the US economy. In the US, home equity accounts for an average of 20% of a household’s gross assets, so any substantial change in the value of a home would dramatically increase the expected life-time wealth of a household. In addition, household’s leverage their housing assets through
their mortgages, thereby exaggerating the effect of home prices on lifetime income. According to the Office of Federal Housing Enterprise Oversight (OFHEO) housing prices have risen by 66% in real terms over the past 10 years. This is approximately a 8% increase in the value of a household’s lifetime consumption. At issue is whether consumers are spending this additional value, what percent of this increase are they consuming, and whether such expenditures wise. I will address the first two of these three issues.

There are several studies which estimate a household’s marginal propensity to consume out of the value of its home to be approximately 0.05 (or, in other words, a household spends 5% of any real increase in the value of its home). These studies either focus on data from the 16 OECD countries or aggregate US data. I propose to use a panel of data of Metropolitan Statistic Areas (MSAs) to explain why consumption has changed over time. I hope to improve on past results primarily by incorporating regional effects.

The use of regional data makes this empirical study unique. Variation of housing prices does not occur across states, it occurs at a more regional level. More specifically, urban areas are driving up the value of property in the US, while rural areas are seeing little and many times negative real growth. Therefore, it makes sense to study variation across counties rather than across states.

The value of the marginal propensity to consume out of household wealth has tremendous significance to the US economy, especially at the present date. If housing prices slow from 66% growth over the past 10 years, consumers may stop increasing their spending at the marginal propensity to consume out of house price wealth. This could shrink consumer spending, which many believe has driven the US economic growth over the past three years.

The following graphs compare consumption and capital gains as well as consumption and housing prices. They motivate any discussion into the marginal propensity to consume out of the value of a home. First, Figure 1 suggests that changes in consumption (retail trade) and capital gains are slightly positively correlated; that is, people spend a small fraction of gains in stock market wealth. Figure 2 suggests that consumption (retail trade) and housing price index (HPI – Housing Price Index) are positively correlated and that the correlation is larger that that of other types of capital gains. The following empirical work will discuss the correlation between consumption and housing prices and compare this to the correlation between consumption and other capital gains.

II. Literature Review

Over the last 40 years there has been a lot of important work to suggest a wealth effect in stock market prices. For example, Starr-McCluer (1998) used Michigan Survey data to understand the impact of stock prices on consumption. Surprisingly, 85% of households said that they did not plan to increase their consumption after recent stock market gains. However, households with holdings of over $250,000 displayed a stronger willingness to spend their capital gains. Starr-McCluer also notes that hesitation to spend capital gains may come from the volatility of stock market returns. Consumers should be less willing to spend capital gains if their portfolios demonstrate greater volatility. Historically the mar
The marginal propensity to consume stock market gains has been estimated to be between 0.03 and 0.07.

However, there is much less work on the effect of housing prices on consumption. Greenspan (2001) estimates that the propensity to consume out of all types of capital gains is between 3 to 5 cents to the dollar. He notes that recently capital gains have mostly appeared in the form of home equity and that theoretically the relation between spending out of wealth in different asset classes should be the same. Since households only began to substantially access their home equity in 1985, there is limited empirical work that measures the marginal propensity to consume out of household wealth.

There has been theoretical debate surrounding whether households consume household wealth. Ludwig and Slock (2002) posit four main ways house values affect consumption. The first is the realized wealth effect whereby homeowners can extract tangible value

\[ \text{FIGURE 1}^6 \]

\[ \text{FIGURE 2}^6 \]
from their home through a second mortgage or selling their homes. The second is the unrealized wealth effect, where a homeowner may consume more but extract no cash from their home because the present value of their lifetime earnings has increased. The third is the budget constraint effect whereby higher property values mean less consumption for renters. The fourth is the substitution effect whereby homeowners who plan to buy another home will consume less because the price of their next home is higher. On aggregate, we would expect the realized and unrealized wealth effects to dominate and consumption to be positively correlated to housing prices. However, the wealth effect and its magnitude are ambiguous.

Porterba (2001) stipulates that housing wealth is a less liquid longer term asset and thus the marginal propensity to consume it should be lower than that of stock market gains. On the other hand, Blondel and Girouard (2001) estimate that home equity accounts for on average 20% of a household’s gross assets, so changes in housing wealth have a much more substantial effect on a household’s total wealth than changes in other asset classes. Still, there are many behavioral questions that arise when examining home prices. Do households view home equity as a hedge against catastrophe? Do they view it as transitional? Are household’s even aware of their household wealth given that they do not receive updates on its value?

Case, Shiller, and Quigly (2005) attempted to answer these questions. They completed a recent study in which they use a panel data set of the 50 US states from 1982 to 1999 to estimate the marginal propensity to consume out of household wealth. The authors present 5 models all of which relate consumption and wealth in two forms, stock market and housing market wealth. Their preferred and best fit model is a generalized least squares model which incorporates first order serial correlation. They find the marginal propensity to consume out of household wealth to be between 0.04 and 0.06. The authors present a final regression which introduces a lag on stock market wealth and they find the marginal propensity to consume out of household wealth to be between 0.08 and 0.10. However, this regression does not fit as well as their generalized least squares model which includes serial correlation. In all models, the authors find “at best weak evidence for a stock market wealth effect.”

The results of their OLS model are consistent with expectations and provide a very good fit to the data. However, there are two key issues to be addressed.

First, Case, Shiller and Quigly do not capture regional variation within states. Campbell and Cocco (2005) emphasize the importance of regional data when estimating the effect of house prices on consumption to allow for regional heterogeneity. Many states include quite disparate regions. In New York State, for example, there is the one of the most densely populated cities as well as rural farmland upstate. Clearly, these households have different propensities to consume wealth. Additionally, much of the rise in housing prices has occurred in urban areas. Therefore, the best measure of the wealth effect of housing prices would use metropolitan areas as regions, not states.

Second, the optimal dates from which to measure the effect of housing wealth on consumption are from 1985 to 2005. The elimination of regulation Q and the Tax Reform Act
of 1986 made it significantly easier for home owners to liquidate their housing wealth by expanding banks’ ability to offer consumers home equity lines of credit.

Second, as Greenspan points out, only over the past 10 years have real housing prices increased substantially. From 1975 to 1995 real US housing prices rose by 10%, but from 1995 to 2005 real US housing prices rose by 66% (OFHEO). Therefore, a study should include up to 2005, but also provide a wide range that does not allow the periods 1995 through 2005 to dominate the data set. These two considerations will be explored in testing the wealth effect and measuring the marginal propensity to consume out of the value of a home.

Finally, many have challenged whether it is possible to measure or evaluate the wealth effect empirically. According to Green (2001), three assumptions must hold true to prove a wealth effect in the stock market or in the housing market. First, holdings of an asset must be widespread. Second, changes in the value of an asset must be unanticipated. Third, changes in the prices of that asset must not be a result in a change in the market discount rate. These do hold for stock market wealth, and they seem to hold true for housing wealth as well. Nearly half of Americans own their own home and changes in housing prices cannot be predicted. The discount rate does have a substantial impact on housing prices, but there are certainly other influences such as preferences and county, state and national macroeconomic trends.

The next two sections will empirically determine the marginal propensity to consume out of household wealth. However, it is important to quantify the expected marginal propensity to consume out of household wealth in the context of the permanent income hypothesis. According to the permanent income hypothesis, changes in the value of homes affect consumption in so much as they affect total lifetime incomes. I calculate that the value of a home has a 12% impact on lifetime income. Therefore 12% is used as an initial measure of the marginal propensity to consume out of household wealth and will contextualize empirical results.

However, this number omits several unique characteristics of household wealth that must be considered when interpreting a marginal propensity to consume out of household wealth. First, housing wealth is less liquid than other types of wealth. Consumers do not have the same access to housing wealth that they do to stock market wealth, so the MPC of housing wealth should be smaller. However, as explained later, housing wealth is serially correlated. Therefore, if the value of a house rises (declines) in a given year we can expect the value to rise (decline) in the surrounding years. Therefore a given change in the value of a house for a given year should have a magnified effect. When housing prices rise (fall), people have cause to consume even more (less) than they would for a similar shift of a different, non-serially correlated asset, in a given year. This would cause the MPC of household wealth to be higher than that calculated from the permanent income hypothesis. In the end, there is conflicting reasoning as to whether the MPC of household wealth should be larger or smaller than the permanent income hypothesis predicts. Empirical analysis will help in determining where the actually MPC of household wealth lies, but the 12% predicted by the permanent income hypothesis remains as a general expectation of its
III. Data Sources

Overview

The effects of housing prices, other capital gains and income on consumption are estimated using a panel of 360 Metropolitan Statistical Areas (MSAs) and the years 1975 to 2003. The goal is to accurately measure consumption, housing prices, other capital gains and disposable income over time. House price data is derived from the House Price Index (HPI) published by the Office of Federal Housing Enterprise Oversight (OFHEO). MSA consumption, capital gains and disposable income data are provided by the Bureau of Economic Analysis (BEA), the same group that publishes widely accepted national economic income and product data.

The model will use panel data, where i represents MSA, and t represents year. Regions are defined by the Office of Management and Budget (OMB) by their MSA. These definitions are based on the 2000 census data and are as of December 2005. MSA is the most widely accepted regional division for studying county level data. Not all counties are included, and many of the 360 MSA regions include multiple counties. The use of MSA regions helps to mitigate spacial correlation of errors. County divisions would treat two counties which are highly interdependent as unique observations; MSA definitions tend to group two such counties together.

While regional data is limiting in its availability, it captures regional effects that state and national data do not. There are two major availability issues with regional data. First, I was able to obtain all variables which past literature suggests should be included in a permanent income model; however, I was unable to find many significant variables beyond those. Second, for approximately 10% of counties housing price and other data is unavailable before 1985. Despite data availability issues, using MSAs to regionally divide the data has key advantages to using states. Classifying the entire state of California, which ranges from Orange County to the Mojave Desert, as one economic region would fail to capture dramatic shifts in regional housing prices. While there are limited variables for individual counties, past literature suggests that the relevant exogenous variables are housing prices, capital gains and disposable income. This is likely mitigated by the substantial number of observations. MSA data is able to capture different geographic distributions while not being substantially limited by omitted variables bias.

The panel will range from 1975 to 2003. Although there is less data prior to 1985 and greater limits on the liquidity of home values, the years 1975 to 1985 will be included to diversify the set of market conditions included in the regression. A further discussion of the years included is in the Empirical Results section following Figure 3. Models in this analysis primarily range from 1977 to 2003 due to a lack of data availability from the housing price index in 1975 and 1976.

Consumption is estimated by real retail trade. Most other studies of the permanent income hypothesis have used taxable retail sales to estimate consumption. However, retail sales data is not available for MSAs. Therefore, the aggregate per capita earnings of retail
establishments will be used to estimate consumption for MSAs. This estimate is consistent with past work in that it uses retail transactions as a proxy for consumption. It differs in that it estimates retail transactions from gross earnings rather than sales. A more detailed description of earnings from retail trade is found in Appendix I.

Housing wealth is approximated by the OFHEO’s Housing Price Index (HPI). It uses 1990 as a base year and then creates an index of the changes in housing prices based on transactions reported by Freddie Mac and Fannie Mae. This index has been criticized because it is primarily based on refinancings, not transactions. It also includes only conforming loans, so properties financed by loans that exceed the conforming loan limit, as well as other types of nonconforming loans, are not included in the calculation of HPI. Additionally, there are several omitted HPI data points for certain MSAs prior to 1985.

The HPI is calculated by using a repeat transaction method. It calculates the regional values of homes by measuring how the value of one home has changed from one transaction to another (Case and Shiller method). Therefore, it is able to control for the quality of home. The HPI values for the MSA regions are based on data from administrative records, census data and survey data. Ultimately, the HPI is the most comprehensive data index on housing prices and includes data from other indices.

The HPI will be used to measure the value of housing wealth. It will be deflated by the CPI using 2003 as a base year. It will also be in index units, and its logarithm will be included in the model. Therefore, its coefficient will be interpreted as the percent change in retail trade for a one percent change in housing wealth.

Disposable income is estimated by the per capita net earnings that is received by all persons less contributions for government social insurance. It includes wages and salaries, supplements to wages and salaries (including employer contributions to employee pension) and proprietors’ income.

Disposable income statistics are primarily derived from unemployment insurance records and are published in the Quarterly Census of Employment and Wages by the Bureau of Labor Statistics. However, some industries are not covered by unemployment insurance. These industries primarily include farms, farm labor contractors, private households, and private elementary and secondary schools, and religions organizations. So, in many cases a national estimate of excluded wages is calculated and then proportionally allocated to MSAs.

Disposable income will be deflated by the Consumer Price Index using 2003 as a base year. It will also be in real dollar units, and its logarithm will be included in the model. Therefore, its coefficient will be interpreted as the percent change in retail trade for a one percent change in disposable income.

Dividends, interest and rent is a proxy for capital gains. It includes dividends received by individuals and retirement plans as reported by the individual income tax return form 1040 from the Internal Revenue Service (IRS). Interest includes monetary interest that is reported for income tax purposes; it comprises approximately 10% of personal income. Rental income includes both actual rents and imputed rents. It assumes that in owner occupied homes the owners rent the homes to themselves. It then imputes an annual rental pay-
ment. This data is based on the IRS data for gross rents and royalties and from the Census of Housing.

Dividends, interest and rent will be deflated by the Consumer Price Index using 2003 as a base year. It will also be in real dollar units, and its logarithm will be included in the model. Therefore, its coefficient will be interpreted as the percent change in retail trade for a one percent change in dividends, interest and rent.

Finally, we must consider serial correlation. The HPI is serially correlated and dividends, interest and rent may be to an extent as well. Houses tend to move in multi-year trends and thus this serial correlation must be accounted for in the regression. This may mean that lagged values of housing prices will need to be included to account for multi-year trends in housing prices.

### Table 1: Data & Variables

<table>
<thead>
<tr>
<th>Label</th>
<th>Estimate of</th>
<th>Unit</th>
<th>Value (in real terms)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>RETAILTRADE_{it}</td>
<td>Consumption</td>
<td>Dollars($000)</td>
<td>Earnings on Retail Trade</td>
<td>BEA</td>
</tr>
<tr>
<td>HOUSINGPRICEINDEX_{it}</td>
<td>House Return</td>
<td>Index units</td>
<td>Index which represents changes in prices of homes in an MSA</td>
<td>OFHEO</td>
</tr>
<tr>
<td>DIVINTRENT_{it}</td>
<td>Capital Gains</td>
<td>Dollars($000)</td>
<td>Dividends, Interest and Rental Income</td>
<td>BEA</td>
</tr>
<tr>
<td>DISPOSIBLEINC_{it}</td>
<td>Disposable Income</td>
<td>Dollars($000)</td>
<td>Disposable Income</td>
<td>BEA</td>
</tr>
<tr>
<td>CPI_{it}</td>
<td>Inflation</td>
<td>Index Units</td>
<td>% Change in Consumer Price Index in year t</td>
<td>BLS</td>
</tr>
</tbody>
</table>

All values are real (deflated by the Consumer Price Index with base year 2003), except for the variable consumer price index. Retail Trade, HPI, Dividends, Interest & Rent, and Disposable Income are measured on a per capita basis. The logarithm of all variables, except for consumer price index, is taken so as to interpret coefficients in the following way: a 1% change in the independent variable causes a \( x \) percent change in retail trade.

So, coefficients on Housing Price Index, Dividends, Interest & Rent, and Disposable income can be interpreted as the percent change in retail trade for a given X% change in the independent variable.

### IV. Empirical Results

#### Understanding Relationships between Variables

The tables below illustrate how retail trade, disposable income, housing prices and dividends, interest & rent have moved from 1975 to 2003
Table 2: Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Std. Dev/ Median</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retail Trade ($)</td>
<td>1,868</td>
<td>404</td>
<td>22%</td>
<td>4,452</td>
<td>560</td>
</tr>
<tr>
<td>Housing Price Index (Index units)</td>
<td>129</td>
<td>NA</td>
<td>NA</td>
<td>292</td>
<td>38</td>
</tr>
<tr>
<td>Dividends, Interest &amp; Rent ($)</td>
<td>4,285</td>
<td>1,800</td>
<td>42%</td>
<td>21,976</td>
<td>681</td>
</tr>
<tr>
<td>Disposable Income ($)</td>
<td>16,379</td>
<td>3,759</td>
<td>23%</td>
<td>54,260</td>
<td>7,458</td>
</tr>
</tbody>
</table>

Note: the standard deviation in the housing price index shrinks in 1995 because it uses 1995 as its base year, i.e. where the index value is 100 across all counties.

Summary Statistics
Correlation Table

<table>
<thead>
<tr>
<th></th>
<th>Retail Trade</th>
<th>HPI</th>
<th>Dividends, Interest &amp; Rent</th>
<th>Disposable Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retail Trade</td>
<td>1.00</td>
<td>0.29</td>
<td>0.42</td>
<td>0.60</td>
</tr>
<tr>
<td>HPI</td>
<td>0.29</td>
<td>1.00</td>
<td>0.08</td>
<td>0.30</td>
</tr>
<tr>
<td>Dividends, Interest &amp; Rent</td>
<td>0.42</td>
<td>0.08</td>
<td>1.00</td>
<td>0.35</td>
</tr>
<tr>
<td>Disposable Income</td>
<td>0.60</td>
<td>0.30</td>
<td>0.35</td>
<td>1.00</td>
</tr>
</tbody>
</table>

From Figure 3, retail trade, housing prices and disposable income move similarly over time. Based on the assumption that retail trade is endogenous, these graphs motivate the theory that changes in housing prices and disposable income have a positive effect on changes in retail trade. Disposable income has the strongest correlation to retail trade, while HPI still has a noticeable correlation. The changes in dividends, interest & rent are noticeably different from those of retail trade over time. The effects of changes in these three variables on retail trade will be analyzed empirically. Past literature has used a model where consumption is a function of capital gains, housing wealth and disposable income, and this paper will do the same.

The illustrations of the standard deviation of each variable in Figure 3 motivate the use of MSA data to measure the wealth effect. The standard deviations are quite large relative to medians indicating that there is substantial variation within counties. In Figure 3, it is clear that this variation is constant and substantial over time. MSAs consistently display wide ranges in their consumption, return on capital gains, and disposable income. Therefore, we will account for this variation using regional data.

Finally, there are three issues to note regarding the period 1975 to 1985. First, even though variables are deflated by the CPI, the substantial inflation during this period may have caused the real values of each variable to fluctuate due to large changes in the CPI. Second, there is less housing data available in this range. Housing data is available for almost every county post 1985, but it begins for many between 1975 and 1985. Additionally, as stated in the literature review section, prior to 1985, there was less liquidity in the value of a home. Due to limits on borrowing, homeowners had greater difficulty spending out of their household wealth. However, I will generally use models which run from 1975 to 2003. This is because the range of 1995 to 2003 exhibited historically excessive real returns in the HPI (total of 66% nationwide) and may tend to dominate a regression from 1985 to 2003. Therefore, the period of 1975 to 1985 will diversify the market conditions contained within the model of consumption.

Model:

\[
\log(\text{RETAILTRADE}_{it}) = C_0 + \beta_1 \log(\text{HOUSINGPRICEINDEX}_{it}) + \beta_2 \log(\text{DIVINTRENT}_{it}) + \beta_3 \log(\text{DISPOSABLEINCOME}_{it}) + \beta_4 \text{CPI}_{it} + u_{it}\]

MSA*: i=1,...,359
Year*: t=1977,...,2003

*May be limited in some models due to data availability.

This model is intended to examine the impact of movements in disposable income,
housing prices and other capital gains on retail trade. Therefore, the variables are measured as logarithms. CPI is not a function of a logarithm because it is a year over year percent change. As a result, coefficients on HPI, dividends, interest & rent and disposable income are interpreted in the following way: an \( X\% \) increase in independent variable \( Z \) leads to a \( \beta \times X\% \) percent increase in retail trade.

*Hypothesis:*

\[
\begin{align*}
\text{Hypothesis 1} & : & H_1 : \beta_1 < 0 \\
\text{Hypothesis 2} & : & H_2 : \beta_1 \leq \beta_2 \\
H_a : \beta_1 & \neq 0 \\
H_a : \beta_1 & > \beta_2
\end{align*}
\]

I will test two hypotheses. First, I will test whether people consume out of household wealth. Second, I will test whether people consume more out of household wealth than they do out of stock market wealth.

**Table 3**

**Method: Panel Least Squares**

*MSA fixed effects and second degree serially correlated errors*

*All variables are deflated by the CPI and are per capita logarithms*

*Dependent variable: Retail Trade*

*(T-statistics are in parenthesis below coefficients)*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing Price Index</td>
<td>0.101 (12.07)</td>
<td>0.065 (7.59)</td>
<td>0.114 (10.87)</td>
</tr>
<tr>
<td>Dividends, Interest &amp; Rent</td>
<td>0.033 (4.32)</td>
<td>0.043 (5.72)</td>
<td>0.039 (4.95)</td>
</tr>
<tr>
<td>Disposable Income</td>
<td>0.556 (38.77)</td>
<td>0.561 (38.69)</td>
<td>0.440 (27.34)</td>
</tr>
<tr>
<td>CPI</td>
<td>-0.324 (-11.60)</td>
<td>-0.452 (-13.77)</td>
<td>-0.301 (-5.19)</td>
</tr>
<tr>
<td>Lag* Years</td>
<td>No Lag</td>
<td>1 year</td>
<td>No Lag</td>
</tr>
<tr>
<td>Time Trend</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.975146</td>
<td>0.975093</td>
<td>0.976810</td>
</tr>
<tr>
<td>Observations</td>
<td>6,639</td>
<td>6,282</td>
<td>5,955</td>
</tr>
<tr>
<td>Durbin-Watson stat</td>
<td>2.00</td>
<td>2.02</td>
<td>2.05</td>
</tr>
</tbody>
</table>

For the set of panel data, Table 3 employs fixed effects by MSA. The HPI is serially correlated and in order to capture this serial correlation (obtain a Durbin-Watson statistic near 2.0) second degree serially correlated errors are employed. Tables 4 and 5 below will also include fixed effects by MSA and second degree serially correlated errors.

Table 3 includes two differences among regressions. Models 3.A and 3.C contain no lag, meaning that all variables are for the same year \( t \). Literature is split as to whether to lag variables which fall into the category of capital gains when measuring their effects on consumption. As for the time interval \( i \), all models except model 3.C contain data from 1977 to 2003. Models 3.A, 3.B, and 3.C include a time trend which captures the concept...
that these variables tend to rise over time.

Model 3.C illustrates the larger affect of changes in housing prices on consumption from 1985 to 2003. As previously stated, I will reference models which include the years 1975-1985 in order to account for more variable market conditions. I include a time trend as it is significant in all models. Its coefficient is approximately 0.002 for each regression. One might expect real consumption to rise substantially over time; however, the coefficient on the time trend variable as well as inspection from Figure 3 leads to a small coefficient. In fact, consumption increases on average .2% year controlling for other effects.

Past work has found that households consume approximately 40% to 60% of changes in disposable income, 3% to 5% of changes in capital gains, and 3% to 10% of changes in housing prices. My results are very much in line with these estimates. Table 3 finds that, regardless of whether we lag HPI, people are consuming approximately 56% of changes in disposable income and 5.5% of changes in dividends, interest & rent. However, the consumption out of household wealth depends on whether we lag the HPI. If we do not lag the HPI, people consume approximately 10% of household wealth and 6.5% if we include a lag. The R-squared value is high and the F-statistic and t-Statistics are very high. In fact, these results for model C in Table 1 are very similar to results done on state panel data.

I will expand on Table 4 to include lagged and non lagged regressors together to reach a model with the greatest scope in explanatory power.

Table 4 incorporates lagged and unlagged capital gains into the model for consumption. It is the most explanatory model and yields similar results as would be interpreted from Table 1. Unlagged housing price index is larger than lagged housing price index, and the coefficient on dividends, interest & rent is similar regardless of a lag.

Ultimately, the results in regression 4.A will be accepted. It is the most explanatory regression and includes a time trend as well as the widest range of time (1977-2003). The hypothesis that the coefficient on HPI is zero is rejected, and the hypothesis that the coefficient on the HPI is less than or equal to the coefficient on dividends, interest & rent is rejected (T-Stat of -4.19). Therefore, we accept the alternative hypotheses, concluding that people consume positively out of changes in household wealth and that the MPC out of household wealth is greater than the MPC out of other capital gains.

These results fit with literature. The permanent income hypothesis roughly predicts a 12% MPC out of household wealth. I find an MPC out of household wealth of 9%. As theory suggests, liquidity and elasticity issues should push the MPC of household wealth below what is predicted by the permanent income hypothesis. We do not expect households to have complete access to household wealth, and we do not expect them to be perfectly sensitive to changes in the value of their home when making consumption decisions.

Finally, there is the issue of endogeneity. One way to evaluate the merits of lagging housing prices is as a control for endogeneity. However, when examining endogeneity, the question is whether housing prices affect consumption or whether consumption affects housing prices. Therefore, including a lag is not a convincing method to control for endogeneity. Ultimately, we would like a way to simply measure the affect of housing prices on consumption while controlling for the affect of consumption on housing prices. A the-
oretically more precise alternative to accounting for endogeneity is the two-stage least squares model which uses the HPI lagged one year as an instrumental variable.

Table 4
METHOD: PANEL LEAST SQUARES
MSA fixed effects and second degree serially correlated errors
ALL VARIABLES ARE DEFATED BY THE CPI AND ARE PER CAPITA LOGARITHMS; LAGS ARE 1 YEAR
DEPENDENT VARIABLE: RETAIL TRADE
(T-STATISTICS ARE IN PARENTHESIS BELOW COEFFICIENTS)

<table>
<thead>
<tr>
<th></th>
<th>2.C</th>
<th>2.B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing Price Index</td>
<td>0.090</td>
<td>0.099</td>
</tr>
<tr>
<td></td>
<td>(10.22)</td>
<td>(9.19)</td>
</tr>
<tr>
<td>Housing Price Index Lag</td>
<td>0.058</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>(6.84)</td>
<td>(5.39)</td>
</tr>
<tr>
<td>Dividends, Interest &amp; Rent</td>
<td>0.42</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>(5.34)</td>
<td>(6.02)</td>
</tr>
<tr>
<td>Dividends, Interest &amp; Rent Lag</td>
<td>0.051</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>(6.71)</td>
<td>(6.99)</td>
</tr>
<tr>
<td>Disposable Income</td>
<td>0.523</td>
<td>-0.451</td>
</tr>
<tr>
<td></td>
<td>(35.01)</td>
<td>(-7.35)</td>
</tr>
<tr>
<td>CPI</td>
<td>-0.416</td>
<td>-0.451</td>
</tr>
<tr>
<td></td>
<td>(-12.67)</td>
<td>(-7.35)</td>
</tr>
<tr>
<td>Time Trend</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.977043</td>
<td>0.977146</td>
</tr>
<tr>
<td>Observations</td>
<td>5,744</td>
<td>6,282</td>
</tr>
<tr>
<td>Durbin-Watson stat</td>
<td>2.01</td>
<td>2.06</td>
</tr>
<tr>
<td>Test H2 (T-stat) no lag</td>
<td>-4.19</td>
<td>-3.72</td>
</tr>
</tbody>
</table>

Finally, there is the issue of endogeneity. One way to evaluate the merits of lagging housing prices is as a control for endogeneity. However, when examining endogeneity, the question is whether housing prices affect consumption or whether consumption affects housing prices. Therefore, including a lag is not a convincing method to control for endogeneity. Ultimately, we would like a way to simply measure the affect of housing prices on consumption while controlling for the affect of consumption on housing prices. A theoretically more precise alternative to accounting for endogeneity is the two-stage least squares model which uses the HPI lagged one year as an instrumental variable.

Table 3 accounts for endogeneity by using the lagged HPI as an instrumental variable. The results are non-conforming. The coefficient on disposable income and dividends, interest & rent agree with the results of Tables 3 & 4 and past literature; however, the coefficient on HPI is substantially larger than that of Tables 3 & 4 and past literature. This is surprising; accounting for the possible positive effect of retail trade on housing prices using instrumental variables causes a substantial rise in the effect of housing prices on consumption. It would seem that a decline or even slight rise in the marginal propensity to consume out of household wealth would result from a two-stage least squares model.
TABLE 5
METHOD: PANEL TWO-STAGE LEAST SQUARES
MSA FIXED EFFECTS AND SECOND DEGREE SERIALLY CORRELATED ERRORS
ALL VARIABLES ARE DEFLATED BY THE CPI AND ARE PER CAPITA LOGARITHMS
INSTRUMENTS: C ; HOUSING PRICE INDEX (LAGGED 1 YEAR) ; DIVIDENDS, INTEREST & RENT ; CPI
DEPENDENT VARIABLE: RETAIL TRADE
(T-STATISTICS ARE IN PARENTHESIS BELOW COEFFICIENTS)

<table>
<thead>
<tr>
<th></th>
<th>5.A</th>
<th>5.B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing Price Index</td>
<td>0.258 (12.17)</td>
<td>0.277 (13.30)</td>
</tr>
<tr>
<td>Dividends, Interest &amp; Rent</td>
<td>0.053 (6.55)</td>
<td>0.046 (5.74)</td>
</tr>
<tr>
<td>Disposable Income</td>
<td>0.455 (26.13)</td>
<td>0.375 (20.86)</td>
</tr>
<tr>
<td>CPI</td>
<td>-0.277 (-9.04)</td>
<td>-0.156 (-2.54)</td>
</tr>
<tr>
<td>Time Trend</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.973486</td>
<td>0.975777</td>
</tr>
</tbody>
</table>

The rise in the MPC out of household wealth when using a lag of HPI to control for endogeneity may be due to the serial correlation of consumption. If both consumption and housing prices are serially correlated we cannot say that the lag of HPI has no effect on consumption; the lag of HPI affects the lag of consumption which then effects current consumption. Therefore, the difficulty lies in finding an instrumental variable that affects current HPI but not current consumption. To date, I have not run across an empirical analysis that finds this instrumental variable.

V. Conclusion

I have examined the effect of housing prices and other capital gains on consumption using the most comprehensive US data set. While it may suffer from greater measurement error, it compensates in its ability to account for regional effects.

The results vary based on econometric interpretation; however, they are generally in line with past results, especially in the cases of the affects of capital gains and disposable income on consumption. Regardless of method, I find that households’ marginal propensity to consume out of changes in disposable income is between 50% and 60%, well within the 40% to 60% range that past literature dictates. I find that households’ marginal propensity to consume out of capital gains to be between 4% and 6%, well covering the range of past literature.

The results for the marginal propensity to consume out of household wealth are larger than that of past literature but not substantially larger. It ranges from 6%, if we lag HPI to correct for endogeneity, to 10%, with no lag. Model 1 in Table 4, which I choose as the strongest model, reports a 9% marginal propensity to consume out of household wealth with no lag. As Case, Quigley and Shiller (2001) find, I find that people consume significantly more out of household wealth than has been previously estimated (9% vs. 5%,
respectively).

The permanent income hypothesis roughly predicts the MPC out of household wealth to be 12%; I empirically find the MPC out of household wealth to be 9%. There are liquidity and volatility issues surrounding this discrepancy. Home values are not liquid nor are consumers always aware of their homes’ values. Since house prices are serially correlated, we expect to see a higher MPC out of household wealth. In the end liquidity issues do substantially reduce the MPC out of household wealth, but the MPC out of household wealth is only 25% below what we might expect (9% out of 12%). Perhaps consumers do spend more out of their homes because they realize that their homes’ prices are serially correlated over years.

What does this mean on a macroeconomic level? Consumer spending has driven the recovery in the US economy since 2001. Therefore, it is important to understand what drives consumption outside of disposable income, namely capital gains and housing wealth. Substantial work has been doing on capital gains’ affect on consumption because segments of households have been receiving substantial capital gains for a long period of time. However, mean US housing prices have only recently started to substantially rise over the past 10 years, and many people are predicting that the growth is not sustainable. If this is the case, strong consumer spending driven by housing wealth could substantially decline.
References


THE WINTER FUEL ALLOWANCE AND PENSIONERS: WARMING THE HOME OR COLD COMFORT?

Pratik Dattani
University of Warwick

ABSTRACT
Using the British Household Panel Survey and the Expenditure and Food Survey, this paper examines the effect of the Winter Fuel Allowance and methods of payment for fuel on fuel expenditure. Prospect theory and mental accounting are applied to the field of public economics as has not been the case before. Evidence shows £6.20 of the Allowance is worth as much as £1 of any other income in influencing fuel expenditure decisions. This is evidence against the concept of having a Winter Fuel Allowance; as well as against mental accounting theory that says labelling a payment in a certain way should stimulate recipients to integrate that money into the mental account recommended by that labelling. However, it is argued this effect can only be robustly tested if the Allowance was greater than fuel expenditure. Additionally, preferences for methods of payment vary for different income groups. Those with higher incomes prefer to maximise hedonic efficiency, while poorer or more loss averse households underconsume fuel by choosing to maximise decision efficiency.

The Winter Fuel Allowance (WFA) is a government transfer paid to people aged 60 and over. It is paid every November in addition to other pension transfers. It was introduced to help tackle perceived pensioner underconsumption of fuel, characterised by the high incidence of excess winter fuel deaths amongst the elderly in Britain compared to the rest of Europe. Since the money can be distributed more easily in other, more costefficient ways, there cannot be an economic justification for such a transfer under the normal Hicksian demand framework. However, the government must see transfer t of WFA as having a greater impact on fuel expenditure than, say, adding t onto the BSP. Labelling a payment thus must influence consumption decisions in a way not captured by conventional economics. Why else would the government adopt a measure that incurs it at least five
percent in extra costs to implement, even though it is essentially a free cash payment that can be spent on anything? That it is free cash makes this analysis distinct from that on food or education vouchers widely considered successful in the USA. The paper also analyses the way households pay for their fuel, and tests mental accounting theory on what determines choices whether to pre- or post-pay. The conclusion is that the free cash element of it, alongside its payment in November or December, prompt pensioners to use the money for uses other than fuel. Additionally, distinct types of consumers prefer the different payments options available.

Most pensioners receive their payments automatically if they received a payment last year and their circumstances have not changed. Although its introduction in 1997 saw some recipients receive only £20, since then it has evolved into a major piece of government expenditure, forecast to reach £1.9 billion in 2003-04. Currently, the WFA benefits around two million households.

**Table 1: Winter Fuel Allowance: An Introduction**

<table>
<thead>
<tr>
<th>Year</th>
<th>Number in receipt (thousands individuals)</th>
<th>Government expenditure (£ millions)</th>
<th>Rules for payment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997-98</td>
<td>9,759</td>
<td>191</td>
<td>£20/household if they receive Income Support; £50 otherwise</td>
</tr>
<tr>
<td>1998-99</td>
<td>10,191</td>
<td>199</td>
<td>£20/household if they receive Income Support; £50 otherwise</td>
</tr>
<tr>
<td>1999-2000</td>
<td>10,225</td>
<td>774</td>
<td>£100/household</td>
</tr>
<tr>
<td>2000-01</td>
<td>11,123</td>
<td>1,749</td>
<td>£150/household</td>
</tr>
<tr>
<td>2001-02</td>
<td>11,202</td>
<td>1,682</td>
<td>£200/household</td>
</tr>
<tr>
<td>2002-03</td>
<td>11,322</td>
<td>1,710</td>
<td>£200/household</td>
</tr>
<tr>
<td>2003-04</td>
<td>na</td>
<td>1,900*</td>
<td>£200/household; £100 more if person is over 80 years of age</td>
</tr>
<tr>
<td>2004-05</td>
<td>na</td>
<td>1,900*</td>
<td>£200/household; £100 more if person is over 80 years of age</td>
</tr>
</tbody>
</table>

*predicted

Because it is universal, non take-up rates of the WFA are lower than with means-tested benefits, where pensioners are likely to feel stigmatised by receipt of these state “hand-outs”, driving non take-up rates higher. Several medical articles advocate its age-contingent nature, adding that it should be aimed at especially women and people living alone.

There has been significant praise for the scheme from interest groups such as Help the Aged and it is popular in the pensioner community. It is not, however, to be confused with Cold Weather Payments, which is a flat-rate payment of £8.50/ week made to vulnerable groups for every week of continuously cold weather. Fuel poverty, defined as a household spending more than ten percent of their income on fuel, has been decreasing – in seven years from 1994-95, this proportion for the poorest 30 percent of households has fallen from 9.4 to six percent, although in the general population this has also fallen from 4.3 to
2.6 percent. In the same period, real gas prices have fallen by 19 percent and electricity prices by 26 percent. Since then, as these prices have increased 11 and seven percent respectively, the number of vulnerable fuel poor has gone up by 400,000, although this has been offset by increases in income contributing to a decrease in fuel poor of 300,000.5 Higher energy prices could force another 250,000 into fuel poverty during 2006, especially as fuel firms have announced severe price increases recently.

Much of the previous reduction in fuel poor households is “thought to be due to changes in incomes and energy prices.” The cost of tackling fuel poverty could be as high as £3.2bn for 1.7m households (indicative maximum total cost from DTI (2003) plus their projected “further 20 percent overheads”). The introduction of the WFA was, in conjunction with a raft of measures such as free TV licenses and greater help towards home insulation for the poorest pensioners, intended to raise the living standards of pensioners. This is often most acute for older pensioners. Wilkinson et al. show the probability of winter death amongst those aged 85-89 is three times as for those aged 75-79. Interestingly, they also show moderate consumption of alcohol cuts this probability substantially.

The normal Hicksian demand function says demand depends on price and income, and is frame invariant. This paper shall relax this invariance assumption, by suggesting labelling a payment $t$ as ‘Winter Fuel Allowance’ changes behaviour to stimulate greater expenditure on fuel than would otherwise be the case. We use Kahneman and Tversky’s prospect theory and Richard Thaler’s mental accounting ideas in an area (public economics) largely ignored by behavioural economists, in Sections A. Section B considers a parallel explanation for the underconsumption of fuel suggested by Prelec and Loewenstein (1998) involving the trade-off between pleasure of consumption and pain of payment. Section C presents regressions looking at the actual WFA data, and we finish with suggestions for extensions and our conclusions in D.

Section A: Theory: Mental Accounting

The impact of a £200 benefit on fuel expenditure is minimal under normal Hicksian conditions, where it would be interpreted solely as an age-contingent lump sum. Munro (2005) finds an income elasticity of fuel expenditure $S$ of only 0.28 using Expenditure and Food Survey (EFS) data for 2001-02 and Blundell et al.’s (1998) estimates imply a similar level. Therefore, for all agents concerned, adding $(1.1)t$ to the BSP would lead to greater expenditure on fuel than having a WFA of level $t$, which incurs implementation costs of 0.1$t$. Politicians’ utility is a function of their citizens’ utility. By introducing the latter rather than the former, utility for the government (and, by implication, for pensioners) must be greater through WFA than through an increase in the BSP. As Munro explains, there must be some bounded rationality on the part of both the players.

The explanation analysed here is given by prospect theory, which shows how individuals evaluate gains and losses. Different consumption decisions are first ordered heuristically. Evaluations about losses and gains are then developed starting from a reference point (an outcome considered neutral/normal, typically the status quo). The utility function passing through this point is s-shaped with concave positive and negative regions with dif-
ferent slopes, which indicate losses produce greater disutility than gains produce utility (loss aversion). Therefore frame invariance no longer applies.

We look at net utility for an individual, starting from:

\[
\sum_{i=1}^{N} v(C_i) = v \left( \sum_{i=1}^{N} C_i \right)
\]

where \( v(C_i) = \text{utility from consumption } C \text{ of good } i \)

Equation under Hicksian framework (1)

But now, having seen gains plateau quickly, while losses don’t, the implication is:

\[
\sum_{i=1}^{N} v(C_i) > v \left( \sum_{i=1}^{N} C_i \right)
\]

Equation under prospect theory framework (2)

This shows that for consumption or gains in income, individuals would be better off by segregating their gains rather than integrating. Conversely for losses, they should integrate rather than segregate. Thaler (1985) shows four simple rules:

1. \( v(x+y) < v(x) + v(y); \ x, y \text{ gains} \)
2. \( v(x+y) > v(x) + v(y); \ x, y \text{ losses} \)
3. \( v(x+y) < v(x) + v(y); \ x \text{ large gain, } y \text{ small loss (mixed gain)} \)
4. \( v(x+y) < v(x) + v(y); \ x \text{ small gain, } y \text{ large loss} \)

\[ v(x) + v(y) = \text{utility from two goods consumer sees as separate transactions} \]
\[ v(x+y) = \text{utility from two goods consumer sees as single, combined transaction} \]

He says “if a gain can be broken down into smaller and smaller gains, utility can be still higher,” because each small payment is transferred into mental accounts earmarked for certain activities i.e. a person subjectively frames a transaction in their mind, which then determines the utility they receive or expect. This suggests a pensioner income made up of a number of smaller transfer payments (BSP, WFA, Pension Credit, free TV licenses, free bus passes) gives greater utility than the same income made up of just one large transfer payment (BSP). The key is the lack of fungibility. By allocating different monies into different accounts, no one pound is (necessarily) the same as any other, which is why the government could label the Winter Fuel Allowance thus and expect that money to be spent disproportionately on this rather than, say, alcohol. If the money is not used for winter fuel, it is more likely to be used for something similar (rival account). This would then mean the mental account for fuel is integrated (Case 1) to offset loss aversion. Net benefit from fuel consumption is made up of the utility of consumption \( v(C) \), government transfer \( t \) (WFA) and the disutility of expenditure \( E(C) \). Cases and Diagrams 1 and 2 show the cases where the WFA is larger than fuel expenditure (Rule 3) and vice versa (Rule 4) respectively.
Case 1: \( v(C) + E(t-C) \) if mental account integrated
Case 2: \( v(C) - E(C) + T(t) \) if mental account segregated

Remember under a Hicksian framework, we would always segregate. Figure 2 shows two cases: where loss aversion is particularly low (nearing Hicksian conditions) and where it is especially high. We can see segregation is clearly preferred in the former case, whereas in the latter, the individual’s pain of paying for fuel is so great, the gain can do little to attenuate it. If we assume this pain is inversely related to income, we can surmise pensioners, who tend to be relatively poorer, are more likely to follow (2*). Additionally, people aged 60 and over spend around £535 on average on fuel every year. Say half of this is in winter. Then on average, \( E(C) > t \). In both cases, we would not expect WFA monies to be integrated into a fuel account, enabling us to reject the null hypothesis of mental accounting versus the alternative of Hicksian demand properties. Thus there is no economic rationale to having a separate government transfer entitled WFA. On the other hand, if \( t \) was sufficiently large, recipients would integrate and \( S \) would then be higher than 0.28.

![Figure 1: Mental account for fuel is integrated (t is large relative to c)](image)

But some budgets are set too low to help deal with “self-help problems.” Consumers want to consume a small amount of a good, but self-control problems prevent them from not consuming far more (especially for luxury goods). Heath and Soll (1996) agree, saying mental accounts cause people to underconsume. Existing mental accounting theory suggests pensioner fuel underconsumption (characterised by excess winter deaths) is because pensioners recognise high fuel costs, and hence label it mentally as a luxury. Thaler (1985) suggests that “the gift of a small portion” would solve this problem by
removing the guilt of consumption. Clearly, this argument lacks satisfying theoretical or empirical impetus in our case in public economics.

Munro (2005) is the only paper that specifically deals with the WFA, although Walker et al.’s (2004) look at Child Benefit (CB) surprisingly finds the marginal propensity to spend CB on children is zero while the marginal propensity to spend it on alcohol may be as high as 0.6, although not because parents don’t care for their children. Their interpretation is that parents care so much for their children, they insulate them from all risk – parents spend enough on their children before they receive the benefit. When the parents receive it, they decide to replace the expenditure they had refrained from before. Both estimate Engel curves analysing how the share of income spent on fuel and children changes due to the WFA and CB respectively. Engel curves show demand for one good as a function of income, with all prices held constant.

In this way, Munro’s paper is a value-added extension. He discusses the Prelec and Loewenstein (1998) theory without delving into empirics, although he does find significant evidence in favour of mental accounting. For a household of income £3000, the effect of the Allowance is to consume approximately £17.30 more fuel a year, or around 8.65 percent of the current £200 payment. Increasing the income to £8000, this effect falls to £11.60, but is still above the 90 pence increase that would result from an ordinary increase of £200 in the BSP.
Whether the actual positive effect of labelling is this high may be down to misspecification. The use of his dummy “for Winter Fuel payment ... [of] £100 or more” is, upon first sight, a good specification, but ignores the impact of the level of WFA, particularly relative to income. Also, consider that only those 60 or older from 1999 received a sufficiently high level of WFA to be included in the test group. The control group should only be those not receiving the Allowance, however because he ignores those who received £20 or £50 in preceding years, the control group includes many who do. The interpretation of the variable is therefore ambiguous.

He uses the council tax band to act as a proxy for house size in a period when house prices increased by 92 percent unequally across the country. He does not control for each region, instead only for Scotland, even though data shows there were two English regions that had an excess winter death count due to the cold higher than that of Scotland. A temperature variable would have been useful here – clearly heating consumption is greater in colder temperatures. Healy’s (2003) study into factors explaining this phenomenon across Europe found the most important determinant to be variations in mean winter temperature. Seasonal variation in mortality was highest amongst (generally warmer) southern European countries; but elsewhere, Northern Ireland and Britain were highest. This could be due to a lack of protection indoors from the cold, as well as income poverty. The first factor makes intuitive sense for a country like Portugal, where house-building best complements its warm weather. Wilkinson et al. (2004) find “the risk of winter death seems to be widely distributed in elderly people,” or that income isn’t an important determinant of excess fuel deaths, but that women and those with respiratory diseases are more likely to die.

Accurate specification for an Engel curve is important too, either to accurately model consumer responses to policy reforms; our work on the changes in relative demand for fuel; or in estimating the impact of demographic change and equivalence scales. Munro uses a squared income specification for the Engel curve, which fits the data well for British Household Panel Survey (BHPS) households. Blundell et al.’s (1998) Family Expenditure Survey (FES) semiparametric estimation of consumer demand however cannot conclusively say whether a linear or quadratic functional form would be more suitable for fuel expenditure. They truncate their data very parsimoniously, cutting merely the top and bottom 2.5 percent of incomes, causing their simple relationship to break down for rich people.

Section B: Theory: Payment Methods

By using transaction-specific accounts, consumers first debit the account by the amount paid for the item, and later credit the account by the benefit they get from using the item. This accounting relation is used by Prelec and Loewenstein to understand how recipients pay for expenditures. When an individual makes a purchase, he experiences an immediate pain of paying, which can undermine pleasure derived from consumption. They propose a theory of prospective accounting which argues if a product has already been paid for, then individuals do not have thoughts of payment when consuming and can therefore
enjoy the product more: "the pain associated with payments prior to consumption is buffered by the thoughts of the benefits that the payments will finance. Another important concept is coupling, which refers to the degree to which consumptions calls to mind thoughts of payment and vice versa.”

A WFA recipient would have the decision criteria of whether to consume an extra unit of fuel as:

$$\delta \sum_{b=1}^{B} U_b - \lambda \delta \sum_{c=1}^{B} p_c > 0$$

where $U_b =$ utility from the future benefits at time $b$ of consuming a unit of fuel now (e.g. being warm now and not getting ill tomorrow)

$p_c =$ the disutility from all future costs of payment for a unit of fuel now

$\delta =$ discount rate by which future benefits and costs are discounted

$\lambda =$ Lagrange multiplier; marginal utility of money when consumer optimises

Prelec and Loewenstein distinguish between different types of (dis)utility, in the form of the ‘consumption experience’ and ‘payment experience.’ To avoid their complex algebraic route, we simplify here. (3) expands to:

$$\delta \sum_{b=1}^{B} (U_b + U_c) - \lambda \delta \sum_{c=1}^{B} (p_b + p_c) > 0$$

$U_b =$ utility drawn from the consumption e.g. the house getting warmer.

$U_c =$ utility drawn from making the payment e.g. relief the future debt has been paid.

$p_b =$ cost of sacrificing future consumption; opportunity cost.

$p_c =$ the pain of payment, of actually handing the hard-earned money over.

As should be obvious, these terms are endogenous – for example there is a trade-off between consumption now and in the future:

$$p_b = \alpha \left( \frac{U_b}{\sum_{t>b} U_t} \right)^{r} \sum_{t=b}^{T} p_t$$

$$U_c = \beta \left( \frac{p_c}{\sum_{t>b} p_t} \right)^{r} \sum_{t=b}^{T} U_t$$

where $\alpha =$ degree to which payments temper the pleasure of consumption ("tempering")

$\beta =$ the degree to which consumption buffers the pain of payment ("buffering")

The above relations show Prelec and Loewenstein’s imputation of these non-linear linkages between the four terms in (4). For example, “pb is now the imputed cost of consumption at time $b$ and is the sum of payments still due at $b$, over the aggregated over $p_t$, prorated over future consumptions still remaining, the summation of $U_t$ over $t>b$ to $T$ adjusted for the degree of coupling, $\alpha$.” Someone with a low price elasticity of demand for fuel would have a low $\alpha$, while individuals who are employed and therefore stay out of the house most of the day, would not derive much pleasure out of keeping the heating on dur-
ing the day and would have a low $\beta$. Substituting these into (4), we get a prediction of how consumption and payment are actually experienced by pensioners:

$$
\delta \sum_{b=1}^{2} U_b + \beta \left( \frac{p_t}{\sum_{r=b}^{t} p_t} \right) \sum_{r=b}^{t} U_t - \lambda \delta \sum_{c=1}^{2} \alpha \left( \frac{U_b}{\sum_{r=b}^{t} U_t} \right) \sum_{r=b}^{t} p_t + p_c
$$

When deciding upon how to pay for fuel bills, the consumer would try to maximise the above term. Hedonic efficiency (consumers want to enjoy the good rather than have nagging thoughts of payment diminish enjoyment) and decision efficiency (knowing what the price of each unit is, so that informed decisions over consumption can be made) are key to maximising utility. Needing to know the costs, yet not wanting to means there is a trade-off between both. Therefore a one-off payment, in November when temperatures start getting colder, could help decouple the consumption-payment relationship and improve hedonic efficiency, thereby increasing fuel expenditure.

<table>
<thead>
<tr>
<th>Method of Payment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prepayment</td>
<td>Slot meter</td>
</tr>
<tr>
<td></td>
<td>Similar payment to mobile phone. Top-ups purchased ex-ante and inserted into meter. Strong coupling.</td>
</tr>
<tr>
<td>Gas/Electricity card</td>
<td>Similar to slot meter</td>
</tr>
<tr>
<td>Budgeting scheme</td>
<td>Fixed monthly direct debit is paid of pre-agreed amount, with any under-/overpayment resolved only at the end of the year. Strong decoupling.</td>
</tr>
<tr>
<td>Postpayment</td>
<td>Account</td>
</tr>
<tr>
<td></td>
<td>Quarterly bills. Some degree of coupling.</td>
</tr>
<tr>
<td>DSS pays whole bill</td>
<td>For poorest households. Strong decoupling.</td>
</tr>
</tbody>
</table>

Consider the payment options for fuel bills in Table 2: pre-payment and post-payment. Post-payment promotes decision efficiency because the consumer consciously thinks about his future bill (pc), and can therefore better control “self-help problems.” Prelec and Loewenstein say pre-payment encourages hedonic efficiency because thoughts of payment at the time of consumption are neglected. Consider an avid internet surfer paying fixed fees for internet access: $U_c$ is high because he can surf to his heart’s content, safe in the knowledge he won’t have to worry about the ticking meter. Cheema and Soman (2006) show controlling spending also depends on how “malleable” different accounts are, which follows on from our earlier fungibility discussion. Say an individual is, at the end of a month of heavy partying, agonising about whether to go for an expensive dinner with friends. His ‘entertainment’ account is in significant deficit. By simply redefining the dinner as part of his ‘food’ budget, not ‘entertainment’, he may justify the spending more easily to himself. Cheema and Soman say creating clear definitions of their spending categories is essential in decision efficiency.
We show how consumers choose their method of payment (from (4) and (5)):

Present Value (Prepayment) = \[ \delta \sum_{b=1}^{B} (U_b + 0) - \lambda \delta \sum_{c=1}^{B} (\alpha p_b + p_c) \]

Present Value (Postpayment) = \[ \delta \sum_{b=1}^{B} (U_b + \beta U_c) - \lambda \delta \sum_{c=1}^{B} (0 + p_c) \]

With prepayment, there is no utility derived from paying in advance (\(U_c = 0\)) but the payment would reduce the pleasure from consumption initially. With postpayment, the pain of payment is buffered by consumption because the foregone interest due to early payment is saved (\(p_b = 0\)). Instead it preys on debt-averse individuals’ minds throughout consumption that the payment is still to come. Prepayment will be preferred if:

\[ -\lambda \delta \sum_{c=1}^{B} (\alpha p_b) > \delta \sum_{b=1}^{B} (\beta U_c) \]

\[ \Rightarrow \delta \beta \sum_{b=1}^{B} (U_c) + \lambda \delta \sum_{c=1}^{B} (p_b) > 0 \]

We can now treat this as the condition for whether coupling is preferred, rather than if prepayment is preferred. This is because we have departed sufficiently from Prelec and Loewenstein’s original theory, where prepayment strictly maximised hedonic, and post-payment decision, efficiency. Prelec and Loewenstein hinted at the trade-off shown in Table 2 whereby there are prepayment methods that maximise hedonic (budgeting) as the theory predicts, but also prepayment options which maintain a strong coupling (slot meters, cards). Through (10), we can explicitly see poor households look to achieve the optimal trade-off between hedonic and decision efficiency, whereas rich households emphasise maximising the former.

This is likely if:

1. \(\lambda\) is high i.e. the individual is poor because he cares a lot for each additional pound – remember our discussion in Figure 2.
2. \(\alpha\) is high i.e. coupling is high and individual is poor. Tempering is greater when you are poor because it is harder to part with the money. Alternatively, “tightwads are less likely to decouple payments and consumptions.”
3. \(\beta\) is high i.e. consumption is viewed as “utilitarian”/ “relatively virtuous.”
4. \(\delta\) is high i.e. individual is patient. This is unlikely – a 60 year-old would typically be less worried about saving for the future than a 40 year-old. When considering paying for a product, consumers are debt-averse so they would prefer to repay bills as quickly as possible. But, as Prelec and Loewenstein show, they display a hyperbolic discounting function, which creates an incentive to constantly postpone painful payments, such as
clearing bills. In this case, they would prefer postpayment.

Section C: The Evidence

We look at both BHPS (1997-2003) and EFS (2001-2003) data to analyse mental accounting empirically. The EFS has the advantage of recording the method of payment for fuel, but is a cross-section, whereas the BHPS is more robust because it is a panel. Using similar controls, we compute the ‘WFA effect’ across both, and additionally the effect of payment methods with the EFS, but since one uses random effects models and the other OLS, some results do vary. As Munro illustrates, take-up rates of WFA are very high and it is reasonable to assume a 100 percent take-up in the following analysis, since neither survey asks explicitly asks upon the receipt of the Allowance.

Figure 2: Yearly expenditure on fuel as a proportion of income (1997-2003)

After trimming households with reasonable assumptions about especially income and expenditure for most of the following regressions, we have an unbalanced panel of 22,107 observations and 3,778 households, of which we have 1,604 pensioner households for the BHPS, while for the EFS, we have 20,047 and 6,579 respectively. For both datasets, we exclude households whose rent includes heating and lighting bills and those reporting fuel expenditure over 30 percent of income. In Figure 2, it is unclear whether a \([\log(\text{income})]^2\) term for the EFS as well as the BHPS is important, but there does seems to be a weak non-linear relationship. Blundell et al. find a similar relation similar. Both graphs suggest possible heteroscedasticity issues, which we control for using heteroscedasticity-consistent standard errors.

Figure 3 confirms our predictions at the end of Section B. A budgeting scheme is most popular amongst wealthier households, because it reduces coupling and the benefit of consumption is less tempered by the pain of payment. Therefore \(\lambda\) and \(\alpha\) are low and inequality (10) is less likely to be fulfilled. Even though prepayment meters charge a higher tariff than for post-payment alternatives, poorer households prefer this method. Paying by DD (account or budgeting) is usually ten percent cheaper than paying by cash/cheque and more so compared to prepayment. In a Scottish Executive survey, 22 of a sample of 40
households who had switched their method of payment for electricity said it was to increase convenience and ease of payment, implying they were looking to reduce the pain of paying $p_e$.

Figure 4 shows that the traditionally cold areas of the country, such as Scotland spent more on fuel than the rest of the country. The poorest households spend the least, while the temperature is higher in congested areas such as London. Using trend average monthly days of frost instead of temperature also provides similar results. There is a correlation between fuel expenditure and temperature of -0.54 and of 0.24 between expenditure and frost.

Section C1: The Evidence: PHPS

We test for the ‘WFA effect’ in two ways: looking at factors affecting the amount of fuel expenditure, and Engel curves with income shares of fuel. (11) uses Munro’s Winter dummy as discussed above. (12) more robustly allows these coefficients to explicitly differ and derives from a Piglog or Almost Ideal Demand System (AIDS) specification like Blundell et al. and forms the basis of proceeding analysis, where the $\beta_{WFA}$ determines the
Figure 4: Average trend temperature 1971-2004 and yearly fuel expenditure by region

The effect of the WFA. Notice the coefficients on the components of Income are 1 and $\beta_{\text{WFA}}$ respectively. If money was fungible and we were to reject the null of mental accounting, both coefficients would be equal. Piglog neatly circumvents the aggregation problem suggested in Figure 2a of translating five societal preference relations into one macro Engel curve by allowing consistent non-linear aggregation through its assumptions of additivity, homotheticity and symmetry. Equation C goes one step further, forming a quadratic AIDS model.

$$\log(\text{FuelExpenditure}) = \beta_0 + \beta_1 \log(\text{Income}) + \beta_2 \log(\text{Income})^2 + \beta_0 \text{Winter}$$

$$w = -\beta_1 \log(\text{Income}) + \text{other}$$
$$\Rightarrow -\beta_1 \log(\text{Income}) = \beta_1 \log(y + \beta_{\text{WFA}} \text{WFA})$$
$$\Rightarrow -\beta_1 \log \left( \frac{y + \beta_{\text{WFA}} \text{WFA}}{\text{Income}} \right)$$
$$\Rightarrow -\beta_1 \log \text{Income} + \beta_1 \log \left( \frac{y + \beta_{\text{WFA}} \text{WFA}}{\text{Income}} \right)$$
$$\Rightarrow -\beta_1 \log \text{Income} + \beta_1 \log \left( \frac{y + \beta_{\text{WFA}} \text{WFA}}{\text{Income}} + \frac{\beta_{\text{WFA}} \text{WFA}}{\text{Income}} \right) \approx -\beta_1 \log \text{Income} + \beta_1 \log \left( 1 + \frac{\beta_{\text{WFA}} \text{WFA}}{\text{Income}} \right)$$

$$[\log(1 + x) \approx 1 + x]$$
$$\Rightarrow -\beta_1 \log \text{Income} + \beta_{\text{WFA}} \beta_1 \left( \frac{\text{WFA}}{\text{Income}} \right) = -\beta_1 \log \text{Income} + \beta_2 \left( \frac{\text{WFA}}{\text{Income}} \right)$$

where $w = \text{Fuel expenditure}/\text{Total income}$
$\text{WFA} = \text{Winter Fuel Allowance amount}/\text{Total income}$
$y = 1 - \text{WFA}$
TABLE 3: REGRESSIONS SHOWING EFFECT OF WINTER FUEL ALLOWANCE ON FUEL EXPENDITURE

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.  Std. Err  t-stat</td>
<td>Coef.  Std. Err  t-stat</td>
<td>Coef.  Std. Err  t-stat</td>
</tr>
<tr>
<td>log(Income)</td>
<td>0.2864  0.0705  4.06</td>
<td>-3.3031  0.0966  -34.2</td>
<td>-36.4881  1.4395  -25.4</td>
</tr>
<tr>
<td>[log(Income)]^2</td>
<td>-0.0125  0.0036  -3.48</td>
<td>1.71106  0.0741  23.1</td>
<td></td>
</tr>
<tr>
<td>WFA/Income</td>
<td>0.5326  0.0676  7.88</td>
<td>0.1290  0.0664  1.94</td>
<td></td>
</tr>
<tr>
<td>Winter</td>
<td>2.4907  1.1163  2.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winter * [log(Income)]</td>
<td>-0.4853  0.2241  -2.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winter * [log(Income)]^2</td>
<td>0.0236  0.0112  2.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>3.8276  0.2508  10.91</td>
<td>32.3108  0.9646  33.5</td>
<td>193.339  7.0307  27.5</td>
</tr>
<tr>
<td>Hausmann test (κ^2)</td>
<td>148.5</td>
<td>110.55</td>
<td>99.13</td>
</tr>
<tr>
<td>No of observations</td>
<td>22160</td>
<td>4726</td>
<td>4726</td>
</tr>
<tr>
<td>Overall R-squared</td>
<td>0.3284</td>
<td>0.4407</td>
<td>0.4886</td>
</tr>
</tbody>
</table>

*Multiplied by 100.
Income over £50,000 and under £2000, fuel expenditure under £5,000, rooms over ten and households with more than nine people have been excluded.

Income and expenditure are both deflated by national RPI. The WFA is left in nominal terms since the payment is so small, deflating would not affect our results. Appendix 2 gives all variable definitions. S in Equation A is similar to Munro’s estimate, suggesting fuel expenditure is price inelastic. From B, we see our results follow Engel’s Law, that the proportion of total expenditure devoted to a good declines as income rises, as w falls by 3.3 percent when income goes up by one percent. Furthermore, we can calculate WFA as 0.53/3.3=0.16. This is highly significant and is evidence against mental accounting. The interpretation is that £1 of other income (y) is worth the same as 16p of WFA to an individual receiving the Allowance. In other words, raising the βWFA by £6.20 is the same as raising the BSP by £1 in terms of impact upon fuel expenditure, notwithstanding the ten percent running costs of the Allowance we assumed at the start. This suggests individuals do not actually respond to labelling by integrating those monies into the fuel account. Instead, they spend the majority of it elsewhere and segregation is distinctly preferred. To confirm this:
Equation C shows similar results, but because of the log(income)^2 term, the determining $\beta_{WFA}$ is nontrivial. Both results are left in for comparison: B offers an easier interpretation for the relevant coefficient; C is the specification recommended by Figure 2 but offers few further insights. Equation A conversely suggests a positive effect of the Allowance for the lowest income decile that rapidly disappears – a household on income £3000 would expect fuel expenditure to increase by 4.46 percent if the extra money was the WFA, 2.14 percent otherwise. But for income £8000, this drops to 2.45 and 2.42 percent instead, far less than Munro predicts.

We know generally $t>c$ i.e. Case 2 applies. Removing the restrictions placed above WFA is still below one, increasing to 0.43. Testing Equation B just for those households where $t>c$ (where it would be more likely to accept the null of mental accounting), $\beta_2$ is 1.41 and $\beta_{WFA}$ 0.4, implying mental accounting does not hold in either case. However, we cannot conclusively say Case 1 is refuted because these results should be treated with caution – placing such a restriction removes half of the sample in B, while these are likely to be smaller households with employed house members, who are home less and therefore require heating less. The other coefficients stay broadly similar, although the temperature variable is significant and more pronounced, indicating 1°C above trend decreases fuel expenditure by an unlikely 11 percent.

Also significant are number of rooms and house members. Both are quadratic, with each additional room/member adding less to the fuel bill than the last. Only having a detached house adds 6-8 percent to fuel bills as a proportion of income compared to a terraced in our specification, thus leaving neither semi-detached or flats as significant effects. Types of fuel not as prevalent around the country such as oil and solid fuel increase bills substantially, as does region. Most region variables are not significant in determining share of income spent on fuel, although living in Scotland adds around 6.5 percent relative to London, while in A, regions matter greatly with Scottish residence adding 20 percent to bills. This makes sense: people spend more in absolute terms across different regions, but mean incomes also vary, meaning the w variable is not as affected by region. Wilkinson et al. too show living in West Scotland increases chances of winter death by 14 percent. The BHPS (and EFS) regions are recoded to reflect the regions in this paper, to simplify comparisons between both works. Living in Scotland adds nearly 20 percent to bills in Equation A, while conversely in Wales, the second greatest amount is spent on fuel, and yet their incidence of excess winter deaths is the second lowest.

As Munro, we also include dummies for the month interviews are conducted. 86 percent of interviewees are asked during September and October and both exhibit a considerably lower fuel expenditure reported than other months, reflecting they are myopic and
remember recent transactions better than older ones. The same phenomenon applies to income, so deflating by income in B and C makes the effect disappear.

The temperature variable is defined as average temperature minus trend winter temperature over 30 years and is derived from Met Office data from 35 weather stations. We find 1°C of unseasonably high temperature reduces \( w \) by around one percent. The real effect may be greater; instead the possible computational errors mentioned above may have introduced a high standard error and downward bias in the value of the coefficient. It performs better though than \( \min(0, \text{average}-\text{trend}) \) or \( \max(0, \text{average days of daily frost}-\text{trend}) \).

The ‘accommodation problems’ variables for leaky roof, damp walls, rot all have powerful implications for policymaking, even though the \( w \) specification withdraws explanatory power from them. They are jointly highly significant only in A, which is likely since the problems are indicative of general state of housing and income too. From a policy perspective, the data shows these problems add around eight percent to bills. Clearly then, tackling these problems in the home would save a lot in fuel. The average person suffering from all these problems spends £736 on fuel a year (0.6 percent of BHPS dataset), compared to £626 for others. Each would save £59 a year if these problems were tackled. Extrapolating for an estimated 15 million households in Britain this would be a saving of £5.3 million a year. Table 4 provides further details, but since health is self-reported and qualitative, we cannot read too much into its insignificant in our regressions. Schemes that help fuel poor households with loft insulation etc. are more effective in tackling fuel underconsumption than free WFA cash because a) their effects are permanent; and b) they do not suffer from Cheema and Soman’s malleability problem. The advantage of food/education vouchers in the USA is they force the recipient to integrate with the relevant account and consumption increases to an efficient level.

**Table 4: Accommodation problems and their effects on fuel expenditure and health**

<table>
<thead>
<tr>
<th>Accomodation problem</th>
<th>Yearly fuel expenditure (£)</th>
<th>Percentage of people with ‘good health’</th>
<th>Percentage of people with ‘poor health’</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condensation</td>
<td>601.02</td>
<td>74.95</td>
<td>13.33</td>
<td>2790</td>
</tr>
<tr>
<td>Leaky roof</td>
<td>722.95</td>
<td>78.08</td>
<td>11.15</td>
<td>780</td>
</tr>
<tr>
<td>Damp walls, floors etc.</td>
<td>653.11</td>
<td>75.50</td>
<td>11.72</td>
<td>1612</td>
</tr>
<tr>
<td>Rot in windows, doors</td>
<td>646.30</td>
<td>75.56</td>
<td>12.59</td>
<td>1747</td>
</tr>
<tr>
<td>Average across population</td>
<td>625.82</td>
<td>78.44</td>
<td>10.34</td>
<td>22482</td>
</tr>
</tbody>
</table>

**Section C2: The Evidence: EFS**

In formulating a similar regression with the EFS, we also add variables for ‘type of payment’ to test the prepayment theory. Notice F is quadratic in income, as in Figure 2, without taking away explanatory power from the log(income) term as is the case with Blundell et al.’s FES work. This may be due their parsimonious trimming, whereas here we consider only those individuals with \( w < 0.3 \).
Table 5: Regressions showing effect of different payment types

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Equation D</th>
<th>Equation E</th>
<th>Equation F</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(Income)</td>
<td>0.0457</td>
<td>0.0084</td>
<td>5.42</td>
</tr>
<tr>
<td>log(Income)-squared</td>
<td>0.3054</td>
<td>0.0118</td>
<td>25.94</td>
</tr>
<tr>
<td>WFA/Income</td>
<td>1.3185</td>
<td>0.0329</td>
<td>40.02</td>
</tr>
<tr>
<td>Slot</td>
<td>-0.0559</td>
<td>0.0246</td>
<td>-2.08</td>
</tr>
<tr>
<td>Budget</td>
<td>-0.0471</td>
<td>0.0294</td>
<td>-7.57</td>
</tr>
<tr>
<td>Card</td>
<td>-0.2225</td>
<td>0.0294</td>
<td>-7.57</td>
</tr>
<tr>
<td>Gas-Slot</td>
<td>0.1120</td>
<td>0.0636</td>
<td>1.76</td>
</tr>
<tr>
<td>Gas-Budget</td>
<td>0.0865</td>
<td>0.0121</td>
<td>7.12</td>
</tr>
<tr>
<td>Gas-Card</td>
<td>0.1005</td>
<td>0.0190</td>
<td>5.29</td>
</tr>
<tr>
<td>Electricity- Slot</td>
<td>-0.0549</td>
<td>0.0430</td>
<td>-1.28</td>
</tr>
<tr>
<td>Electricity - Budget</td>
<td>0.0478</td>
<td>0.0112</td>
<td>4.25</td>
</tr>
<tr>
<td>Electricity - Card</td>
<td>-0.0056</td>
<td>0.0152</td>
<td>-0.37</td>
</tr>
</tbody>
</table>

| Number of observations | 18914 | 18914 | 18914 |
| Number of variables | 38 | 35 | 36 |
| F-Test | 160.11 | 307.46 | 328.24 |
| Adjusted R-squared | 0.2422 | 0.3619 | 0.3838 |

Control for payment methods is Payment by Account for all equations.
All standard errors are Huber/White/sandwich estimators of variance to attempt to counter heteroscedasticity.
Sources: Expenditure and Food Survey (2001-2003)

The results should show decision efficiency is promoted by having slot meters or cards, thereby reducing expenditure on fuel. However, we see gas expenditure is 11 percent higher for slot users compared to account holders, while for electricity users this effect is less pronounced (however both are insignificant). If true, this could be because the greater per-unit expense for this method outweighs the gain in decision efficiency. Remember in (10) we ignored price differences, assuming they were constant. E and F offer results more consistent with theory. The most pronounced effect is of gas cards – 22 percent less is spent as a proportion of income. On the other hand, budget customers, buoyed by cheaper bills, spend between 5-9 percent more on bills. This is evidence that α, measuring coupling is low while λ is high i.e. the budget customers prefer to maximize hedonic efficiency because they are richer and therefore the price pressures are not as acute, something Figure 3 confirms. The former result is likely to exacerbate over time as gas prices rise at a record pace. Because of these cyclical effects we can still assert Prelec
and Loewenstein’s theory holds in practice.

We also consider WFA/ Income again. The results approach Munro’s results, with $\beta_{EWFA} = 1.87$ providing evidence for mental accounting. The reason may be that in the panel, I controlled for individual heterogeneity, whereas for a cross-sectional study this is not possible. Following expenditure behaviours over time is important here since the WFA has only recently become important government expenditure.

The coefficients on the rooms and type of house variables corroborate with the BHPS results, although living in a flat seems to lower fuel bills by a larger 15 percent. This effect disappears when determining w because flat inhabitants are generally poorer households that would also have a correspondingly lower income. An additional adult in the house adds more to heating bills than an additional child. The employment variables are interesting: if there is someone employed in the house, they would spend more time out of the house and would therefore require less heating; however if the head of household is unemployed, the income effect reduces expenditure by five percent, while expenditure as a proportion of income is cut by even more. Myopic interviewees interviewed in the first two quarters report a fuel expenditure of 10-12 percent more than if they are asked in the fourth quarter, which is what we expected earlier, while the regional dummies are highly significant, as with Equation A.

Section D: Conclusions and Extensions

This paper finds strong evidence for Prelec and Loewenstein’s theory on preferences for prepayment. Further research could attempt to quantify the (dis-)utilities shown there, as well as incorporate the price effect that distorted more precise conclusions from our analysis. This factor will be significant in the near future as energy prices increase. The FES 2000 and the EFS 2004 ask WFA-specific questions which cast doubt upon our assumption of 100 percent take-up. Due to the absence of this information in our data, we have not considered the implication of backdated WFA payments (these may mean $t>c$ in more cases), which appear to be slow to materialise. Better computation of the temperature and health variables would be desirable. A comparison with similar government transfers in a mental accounting framework would make the results more robust e.g. cash help towards childcare provision, or German Kindergeld. A major research could be to separate the different Engel curves in Figure 2a, determining which types of households correspond to which demand preferences. This could potentially be invaluable in determining which sections of society transfers such as the WFA should be aimed at the most.

We have gained a preliminary insight into public economics mental accounting. The WFA does not seem to have an effect on fuel expenditure. The timing of the payment could have affected our result – if recipients view it as a Christmas windfall, they could put a disproportionate amount into their ‘entertainment’ or ‘presents’ accounts instead of ‘fuel’. This is again a topic for further research. As this study and Walker et al.’s have shown, free cash, however labelled, does not produce the desired results. Currently, the WFA is viewed as free cash, fit to spend freely. Integration is only preferred (Case 1) if the WFA gain is larger than the fuel expenditure. Instead expenditure is more than three times the
Allowance level, so this paper cannot give a clear-cut rejection of the null of mental accounting, simply saying the conditions are not right to be able to test the magnitude. Testing this places too severe restrictions on the data.

If the government does not raise the WFA level substantially (as would seem) to counter the energy price increases and this issue, it would be far more cost-effective to provide pensioners the payment through a simple increase in the BSP. If the goal is still to tackle fuel underconsumption, there ought to be a restriction placed on the WFA, whereby it cannot be used for anything other than fuel. This ‘voucher’ could then be distributed just in time for winter. Alternatively, this money could be used to target only fuel poor, vulnerable households to eradicate problems such as leaky roofs and damp. This one-off cost would cost the same as one year’s expenditure on the Allowance, and would target the neediest far more effectively (note: this is merely a very high-level analysis). Overall, tackling fuel underconsumption and addressing Britain’s bad record in excess winter deaths is admirable, but the devil lies in the detail. Of course, it may be that the Allowance’s introduction had nothing to do this noble aim and was just to win marginal pensioner votes.

References


BBC News Online, “Brown moves to cool house prices” (17 March 2004).


Heath, Chip and Jack B. Soll, “Mental Budgeting and Consumer decisions,” Journal of Consumer Research


NEAa, Winter Fuel Payments, Policy position papers (March 2004).

NEAb, Prepayment Meters, Policy position papers (March 2004).

Pensions Service, Online Q&A: ‘Winter Fuel and Age-related payments.’

Pollak, Robert A. and Terence J. Wales, “Demand System Specification and Estimation,”


uSwitch, ‘Energy bills to soar this Winter,’ Press Release, 29th January 2006


DOES MONEY BUY HAPPINESS?

Ruth Coffman
Georgetown University

Introduction
This paper explores what determines variation in levels of reported happiness, or “subjective well-being,” across nations. Literature on the subject diverges around two disciplines. Sociology has attempted to correlate happiness with various cultural or psychological traits (such as “collectivism,” “Protestantism,” or “self-esteem”). Economists have investigated more quantitative determinates such as income or income inequality. The problem with the former is that they are open to accusations of cultural bias. Economists, for their part, stand accused of ignoring something that seems intuitively obvious: answers to the question “how happy are you, over all?” will be linguistically and culturally skewed.

This paper rests firmly within the bounds of economic theory but recognizes that “something else” influences subjective well-being. We can control for this “something else” without erroneously defining it or worrying too much about its specifics. Controlling for culture, this analysis more accurately investigates the implications of economic theory for happiness.

Model
This paper investigates the utility of income using happiness data, a basic relationship which many scholars have explored. Here, however, we assume that cultural variables might be relevant and control for them. The model is designed to investigate two relationships:

1) the relevance of economic predictors versus cultural predictors
2) the relationship between income and happiness

The paper shows there is a false distinction between cultural and economic variables. Simply put, there is not an either/or relationship but rather a both/and. The paper proposes that studies have misreported the relationship between income and happiness by not including cultural data. Scholars have suggested the existence of a “happiness threshold,” a level of income (usually GDP per capita) after which increases in income do not affect happiness levels. The paper attempts to find an “income threshold,” that is independent of cultural effects on happiness.
Variables

The model incorporates the following variables:

1) GDP per capita, at purchasing power parity, is used as a proxy for income. World Bank estimates are used.

2) Gini indices are used to proxy for income inequality, with high values corresponding to more inequality. A variable for inequality was included to avoid under or over estimating income’s relationship to happiness. (For example, if two countries with the same income have different levels of inequality, it is quite possible that their happiness levels will correlate to how many people in the country actually have that amount of income). The analysis used data from the United Nations’ inequality database.

3) Countries are put into cultural groups (which are “cultural dummies” in the regression) formulated from political scientist Samuel Huntington’s The Clash of Civilizations.

Why Samuel Huntington?

As aforementioned, papers investigating cultural influences on happiness have often defined a cultural characteristic, given countries “scores” on this characteristic, and then correlated that characteristic with happiness. This methodology adds more subjectivity to the data and risks constructing an index that closely correlates with a specific variable in expected ways – introducing an experiment-produced endogeneity. Huntington’s analysis has several advantages: it is rather simplistic and not produced by sociological methods; instead, Huntington uses history to group countries. The groups have the advantage of being designed independently for a completely different purpose, and there is no “score.” Countries are grouped rather intuitively.

Results

A panel least squares model is used. The dependent variable is the percentage of people in the country reporting “very happy” as their state of being to a world values survey. The dependent variable is also modified to produce a logit model inside the least squares analysis. The two hypotheses were:

1) If cultural variables are included in a regression with happiness as a dependent variable, then more of the data will be explained. Cultural variables will have statistically and economically significant effects on happiness data.

2) If we control for culture, then income will still have a positive and economically significant effect on happiness. This effect will be “lower” than the literature suggested.

Both hypotheses are correct. Cultural variables have increased the R2 of the data, and, in the logit model, all cultural dummies except Japanese and Slavic Orthodox are significant at the 5% level. As for the second hypothesis, income is positively correlated with happiness, although the relative magnitude of the coefficient is small. (As a side note, income inequality is negatively related to happiness, but with diminishing marginal returns).

The results suggest that income and income inequality have different relationships to happiness depending on i)the level of income already present; ii)the level of inequality already present; and iii)the cultural group to which a country belongs. These relationships are explored more fully in the long version of this paper. The graph of the logit results
indicates that the income threshold, for most cultural groups, is at $8,000 GDP per capita (at purchasing power parity), a significantly lower threshold than previous literature suggests.

**Conclusion**

Only so much money buys happiness, and at different cultural exchange rates. Future economic investigations into cross-national happiness should include cultural variables, which have significant explanatory power. In addition, after a threshold, variations in income no longer explain variations in happiness. More investigation is needed into the other correlates of happiness, keeping in mind that cultural variation will also have an impact. Finally, my model is a humble attempt to make economic theory more robust to criticisms concerning its disregard of people and cultures. This paper has shown that these factors can be incorporated, and will enhance and improve our models.

**Table 1: Very Happy on Income, Income Inequality**

<table>
<thead>
<tr>
<th>Panel Ordinary Least Squares</th>
<th>I % very happy</th>
<th>% very happy logit</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent Variable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>35.820</td>
<td>-1.137</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.55)</td>
<td>(-0.96)</td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>0.001356</td>
<td>0.000883</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.60)</td>
<td>(5.98)</td>
<td></td>
</tr>
<tr>
<td>Gini Index</td>
<td>-2.6284</td>
<td>-0.1460</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.93)</td>
<td>(-2.31)</td>
<td></td>
</tr>
<tr>
<td>Gini Index Squared</td>
<td>0.0437</td>
<td>0.0023</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.46)</td>
<td>(3.20)</td>
<td></td>
</tr>
<tr>
<td>Confucian</td>
<td>15.8255</td>
<td>1.4271</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.07)</td>
<td>(2.86)</td>
<td></td>
</tr>
<tr>
<td>Hindu</td>
<td>31.3912</td>
<td>2.4371</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(10.38)</td>
<td>(10.24)</td>
<td></td>
</tr>
<tr>
<td>Islamic</td>
<td>40.9725</td>
<td>2.6667</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(14.37)</td>
<td>(11.69)</td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latin American</td>
<td></td>
<td>0.8882</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.89)</td>
<td></td>
</tr>
<tr>
<td>Slavic Orthodox</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Western</td>
<td></td>
<td>0.3546</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.28)</td>
<td></td>
</tr>
<tr>
<td>R^2 (adjusted)</td>
<td>0.62</td>
<td>0.71</td>
<td></td>
</tr>
</tbody>
</table>

*Base Group is Slavic Orthodox and Western. All variables are significant at or above the 5% level.
APPENDIX A: Fifth Annual Carroll Round Presentation Schedule

Session 1A Chair: Phil Cross (Assistant Professor, Georgetown University)

Xiaoti Zhang (University of Warwick)
*A Natural Experiment on Reference Dependent Model - A Test of Alternative Theories: Traditional Expected Utility Theory vs. New Prospect Theory*

Emy Reimao (Georgetown University)
*Money Stamped for Development: A Study on the Elasticity of Consumption with Respect to International Remittances*

Sarah Carroll (Stanford University)
*Local Divorce Norms, Intrahousehold Bargaining, Poverty and Development in Rural Ethiopia*

Session 1B Chair: Anders Olofsgard (Assistant Professor, Georgetown University)

Vikram Shankar (Georgetown University)
*Bankruptcy and Industry Performance: Is Government Protection Hurting the Airline Industry?*

David Wiczer (Carleton College)
*The Impact of Implicit Pension Debt on Sovereign Credit Ratings*

Evgeniya Petrova (Dartmouth College)
*EU Integration and Stock Market Performance in Accession Countries*

Session 1C Chair: Raj Desai (Assistant Professor, Georgetown University)

Svetoslav Roussanov (Columbia University)
*The Currency Board in Bulgaria: A Promoter of Economic Growth Rather than an Obstacle to It*

Stella Klemperer (Brown University)
*Dollarization in Ecuador: Analysis of Motives and Outcomes*

Jennifer Dawson (Illinois Wesleyan University)
*The Effect of Oil Prices on Exchange Rates: A Case Study of the Dominican Republic*

Session 2A Chair: Garance Genicot (Assistant Professor, Georgetown University)

Oyebanke Oyeyinka (Carleton College)
*Can Development Cause AIDS?*

Shanthi Manian (Georgetown University)
*The Determinants of Participation and Success in the Olympic Games: A Cross Country Analysis*

Corinne Low (Duke University)
*The Senegalese Experience: Rethinking Fertility Theory for Highly Religious Societies*

Session 2B Chair: Douglas Brown (Associate Professor, Georgetown University)

John Nesbitt (Georgetown University)
*Measuring the Wealth Effect of Housing Prices: A Regional Approach*

Dubravka Colic (Wellesley University)
*Minimum Wages in Eastern and Central Europe: Effects on Employment and Inequality*

Pratik Dattani (University of Warwick)
*The Winter Fuel Allowance and Pensioners: Warming the Home or Still Out in the Cold?*
APPENDICES

**Session 2C** Chair: Arik Levinson (Associate Professor, Georgetown University)

Geoffrey Yu (Carleton College)
*The Impact of Government Intervention on Foreign Exchange Market Efficiency: Japan 1995-2005*

Juan Carlos Suarez (Trinity University)
*Forecasting Mexico's Inflation: the Effects of an Inflation Targeting Regime*

Daniel Kurland (Dartmouth College)
*The Determinants of Cross Border Equity Flows: Gravity, Friction, Momentum*

**Session 3A** Chair: Robin King (Assistant Professor, Georgetown University)

Sherri Haas (Illinois Wesleyan University)
*Economic Development and the Gender Wage Gap*

Jen Hardy (Georgetown University)
*Multilateral vs. Bilateral Aid: Differing Effects on Growth?*

Lauren Iacocca (University of California - Los Angeles)
*Exaggerated Currency Union Benefits: A Comparison of International Trade Increases in Cases of Currency and Trade Union Formation*

**Session 3B** Chair: Adhip Chaudhuri (Assistant Professor, Georgetown University)

Michael Monteleone (University of Chicago)
*Knocking from Within: Contemporary Social Unrest and its Consequences for a Stable China*

Natasha Nguyen (University of California - Berkeley)
*Regionalism and Racism as a Barrier to Economic Development*

Ruth Coffman (Georgetown University)
*Does Money Buy Happiness?*

**Session 3C** Chair: Jonathan Heathcote (Assistant Professor, Georgetown University)

Héber Delgado-Medrano (Georgetown University)
*Economic Determinants of Dropout Rates in Mexico*

Salifou Issoufou (University of Wisconsin - Madison)
*Impact of Worker's Remittances on Developing Countries' Growth*

Austin Vedder (Dartmouth College)
*The Effect of Immigration on Growth in Texas Counties*
APPENDIX B: PAST SPEAKERS

First Annual Carroll Round (April 5-7, 2002)
Roger W. Ferguson, Federal Reserve Board of Governors
Donald L. Kohn, Federal Reserve Board of Governors
Lawrence B. Lindsey, Assistant to the President and National Economic Council
Edwin M. Truman, Institute for International Economics
John Williamson, Institute for International Economics

Second Annual Carroll Round (April 11-13, 2003)
R. Glenn Hubbard, Council of Economic Advisers and Columbia University
Donald L. Kohn, Federal Reserve Board of Governors
John Williamson, Institute for International Economics

Third Annual Carroll Round (April 15-18, 2004)
Donald L. Kohn, Federal Reserve Board of Governors
John F. Nash, Jr., Princeton University (1994 Nobel Laureate)
Peter R. Orszag, The Brookings Institute

Fourth Annual Carroll Round (April 22-24, 2005)
Ben S. Bernanke, Federal Reserve Board of Governors
William Easterly, New York University
Maurice Obstfeld, University of California at Berkeley
Edwin M. Truman, Institute for International Economics

Fifth Annual Carroll Round (April 28-30, 2006)
Kemal Dervis, United Nations Development Programme
Thomas C. Schelling, University of Maryland (2005 Nobel Laureate)
APPENDIX C: Former Carroll Round Steering Committees

First Annual Carroll Round
(April 5-7, 2002)
Christopher L. Griffin, chair (SFS ’02)
William B. Brady (SFS ’02)
Cullen A. Drescher (COL ’04)
Meredith L. Gilbert (COL ’04)
Joshua M. Harris (SFS ’02)
Andrew T. Hayashi (SFS ’02)
Mark R. Longstreth (SFS ’04)
Kathryn E. Magee (SFS ’02)
Ryan F. Michaels (SFS ’02)
J. Brendan Mullen (SFS ’02)
Scott E. Pedowitz (SFS ’02)
Waheed A. Sheikh (SFS ’04)

Second Annual Carroll Round
(April 11-13, 2003)
Seth M. Kundrot, chair (SFS ’03)
Nada M. Abdelnour (SFS ’03)
Maria M. Arhancet (SFS ’04)
Victoria E. Bembenista (SFS ’03)
Michael J. Callen (SFS ’05)
Eric M. Fischer (SFS ’03)
Daphnney Francois (SFS/GRD ’04)
Meredith L. Gilbert (COL ’04)
Jeffrey M. Harris (COL ’03)
Robert S. Katz (COL ’04)
Marina Lafferriere (SFS ’06)
Lu Shi (SFS ’03)
Stacey H. Tsai (SFS ’03)
Robert T. Wrobel (SFS ’03)
Erica C. Yu (COL ’05)

Third Annual Carroll Round
(April 15-18, 2004)
Meredith L. Gilbert, chair (COL ’04)
Héber M. Delgado-Medrano (SFS ’06)

Ryan V. Fraser (SFS ’04)
Tetyana V. Gaponenko (SFS ’07)
Yunjung Cindy Jin (SFS ’05)
Sarah H. Knupp (SFS ’04)
Robert S. Katz (COL ’04)
Marina Lafferriere (SFS ’06)
Alia F. Malik (SFS ’04)
Susan M. Work (SFS ’04)
Beatka J. Zakrzewski (SFS ’04)

Fourth Annual Carroll Round
(April 22-24, 2005)
Erica C. Yu, chair (COL ’05)
Jasmina Beganovic (SFS ’05)
Lucia Franzese (SFS ’07)
Dennis L. Huggins (SFS ’05)
Yunjung Cindy Jin (SFS ’05)
Jonathan W. Kirschner (SFS ’05)
Susan C. Kleiman (SFS ’05)
Yousif H. Mohammed (SFS ’06)
Amy M. Osekowsky (SFS ’07)
Daniel P. Schier (SFS ’05)

Fifth Annual Carroll Round
(April 27-30, 2006)
Marina Lafferriere, chair (SFS ’06)
Irmak Bademli (SFS ’06)
Stephen Brinkmann (SFS ’07)
Heber Delgado (SFS ’06)
Lucia Franzese (SFS ’07)
Yasmine Fulena (SFS ’08)
Jen Hardy (SFS ’06)
Michael Kunkel (SFS ’08)
Yousif Mohammed (SFS ’06)
Emy Reimao (SFS ’06)
Tamar Tashjian (SFS ’06)
APPENDIX D: Members of the Advisory Panel

Meredith L. Gilbert, Yale University
Christopher L. Griffin, Andrew W. Mellon Foundation
Andrew T. Hayashi, University of California at Berkeley
Mitch Kaneda, Georgetown University
Robert S. Katz, World Resources Institute
J. Brendan Mullen, The Advisory Board
Scott E. Pedowitz, University College London
Erica C. Yu, University College London
APPENDICES

APPENDIX E: Past Participants

First Annual Carroll Round (April 5-7, 2002)

Azhar Adbul-Quader Columbia University
Santosh Anagol Stanford University
William Brady Georgetown University
Daniel Braun Oberlin College
Jacqueline Bueso University of Pennsylvania
Karla Campbell University of Virginia
Benn Eifert Stanford University
Courtney Fretz University of Pennsylvania
Carlos Galvez Stanford University
Aniruddha Gopalakrishnan Duke University
Christopher Griffin Georgetown University
Casey Hanson Lehigh University
Joshua Harris Georgetown University
Andrew Hayashi Georgetown University
Marco Hernandez Massachusetts Institute of Technology
Katia Hristova Illinois-Wesleyan University
Maria 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
APPENDIX E: *Past Participants*

Second Annual Carroll Round (April 11-13, 2003)

Nada Abdelnour
Amanda Barnett
Andrea Bell
Patrick Byrne
David Chao
Sylvia Ciesluk
Adam Doverspike
Benn Eifert
Adam Engberg
Alexandra Fiorillo
Eric Fischer
Zlata Hajro
Samina Jain
Avinash Kaza
Eric Kim
Seth Kundrot
Lada Kyi
Lee Lockwood
Sunil Mulani
Holly Presley
Duncan Roberts
Lu Shi
Shanaz Taber
Jiang Wei

Georgetown University
Emory University
Wellesley College
University of Colorado
Cornell University
Lehigh University
Georgetown University
Stanford University
Georgetown University
Connecticut College
Georgetown University
Wellesley College
Georgetown University
Stanford University
George Washington University
Georgetown University
Rice University
Northwestern University
New York University
Vanderbilt University
University of California at Berkeley
Georgetown University
Barnard College
University of Michigan
## APPENDIX E: Past Participants

### Third Annual Carroll Round (April 15-18, 2004)

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jeffrey Arnold</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Julia Berazneva</td>
<td>Mt. Holyoke College</td>
</tr>
<tr>
<td>Mehmet Cangul</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Richard Carew</td>
<td>University of Virginia</td>
</tr>
<tr>
<td>Ashley Coleman</td>
<td>Vanderbilt University</td>
</tr>
<tr>
<td>Dilyana Dimova</td>
<td>Stanford University</td>
</tr>
<tr>
<td>Fernando Galeana</td>
<td>Stanford University</td>
</tr>
<tr>
<td>M. Blair Garvey</td>
<td>Emory University</td>
</tr>
<tr>
<td>Meredith Gilbert</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Adam Greeney</td>
<td>Oberlin College</td>
</tr>
<tr>
<td>Asim Gunduz</td>
<td>University of Virginia</td>
</tr>
<tr>
<td>Marc Hafstead</td>
<td>Northwestern University</td>
</tr>
<tr>
<td>Andrew Hayashi</td>
<td>University of California at Berkeley</td>
</tr>
<tr>
<td>Katherine Howitt</td>
<td>McGill University</td>
</tr>
<tr>
<td>Sohini Kar</td>
<td>Columbia University</td>
</tr>
<tr>
<td>Josh Lewis</td>
<td>McGill University</td>
</tr>
<tr>
<td>Satish Lohani</td>
<td>Illinois-Wesleyan University</td>
</tr>
<tr>
<td>Alexis Manning</td>
<td>Illinois-Wesleyan University</td>
</tr>
<tr>
<td>Sara Menker</td>
<td>Mt. Holyoke College</td>
</tr>
<tr>
<td>Elizabeth Mielke</td>
<td>Vanderbilt University</td>
</tr>
<tr>
<td>Stratos Pahis</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Alicja Pluta</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Adam Raymakers</td>
<td>Dalhousie University</td>
</tr>
<tr>
<td>Caroline Schmutte</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Matt Sekerke</td>
<td>Johns Hopkins University</td>
</tr>
<tr>
<td>John Soleanivic</td>
<td>Columbia University</td>
</tr>
<tr>
<td>Kai Szakmary</td>
<td>Columbia University</td>
</tr>
<tr>
<td>Brandon Wall</td>
<td>Yale University</td>
</tr>
<tr>
<td>Kenneth Ward</td>
<td>University of Chicago</td>
</tr>
<tr>
<td>Susan Work</td>
<td>Georgetown University</td>
</tr>
</tbody>
</table>
APPENDIX E: Past Participants

Fourth Annual Carroll Round (April 22-24, 2005)

Lidia Barabash
Jasmina Beganovic
Xun Bian
Michael Fuchtgott
Michael Gechter
Kevin B. Goldstein
Michael Haase
Dennis Huggins
Michael Insel
Jonathan Kirschner
Shiying Lee
James Liao
Brian Lichter
Wee Lee Loh
Alice Luo
Katharine Mullock
Jose Mustre del Rio
Leah Nelson
Ee Cheng Ong
Matthew Phan
Nina Rendelstein
David Rogier
Ana Maria Romero
Nathan Saperia
Bogdan Tereshchenko
Olga Timoshenko
Tom Vogl
Kenneth Ward
Jonathan Wolfson
Suzanne Zurkiya

Dartmouth College
Georgetown University
Illinois-Wesleyan University
Columbia University
Pomona College
Dartmouth College
University of Copenhagen
Georgetown University
Claremont McKenna College
Georgetown University
Duke University
Dartmouth College
Washington University
Cornell University
Duke University
University of Western Ontario
Ohio State University
Georgetown University
Wellesley College
Columbia University
Washington University
Washington University
Illinois-Wesleyan University
Dartmouth College
Georgetown University
University of Western Ontario
Princeton University
University of Chicago
Washington University
Emory University
APPENDIX E: Past Participants

Fifth Annual Carroll Round (April 28-30, 2006)

Sarah Carroll             Stanford University
Ruth Coffman             Georgetown University
Dubravka Colic             Wellesley University
Pratik Dattani          University of Warwick
Jennifer Dawson            Illinois Wesleyan University
Héber 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        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
**Note on Paper Submissions**

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