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
The Seventh Annual Carroll Round Steering Committee

Sue J. Bai
Stacey L. Droms
LiJia Gong
Brandon A. Feldman
Yasmine N. Fulena
Kory D. Kantenga
Michael K. Kunkel
Sung E. Kim
Fuyang Abby Zhang

Carroll Round Proceedings
Copyright © 2008
The Carroll Round
Georgetown University
Washington, DC

http://carrollround.georgetown.edu

All rights reserved. No part of this publication may be reproduced or transmitted in any form by any means, electronic or mechanical, including photography, recording, or an information storage and retrieval system, without the written permission of the contributing authors. The Carroll Round reserves the copyright to the Carroll Round Proceedings volumes.

Original works included in this volume are printed by consent of the respective authors as established in the “Non-Exclusive Copyright Agreement.” The Carroll Round has received permission to copy, display, and distribute copies of the contributors’ works through the Carroll Round website (carrollround.georgetown.edu) and through the annual proceedings publication. The contributing authors, however, retain the copyright and all other rights including, without limitation, the right to copy, distribute, publish, display, or modify the copyright work, and to transfer, assign, or grant license of any such rights.

Printed in the United States of America.
CONTENTS

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

Carroll Round Proceedings

How do Informal and Formal Restrictions on Women Affect Economic Performance? 3
ADRIENNA HUFFMAN
Washington University

Physical, Informational and Human Distance: A Comparison of the Determinants of Bilateral Portfolio and Foreign Direct Investments Flows 16
MATTHEW PECH
Dartmouth College

Has Foreign Bank Entry Led to Sounder Banks in Transition Economies? 27
HELERI RANDE
New York University

Impact of Innovative Micro-Credit Policy on Rural Credit Markets:
the Case of Thailand’s Million Baht Village Fund 39
YANA MORGULIS
University of Chicago

Fruitless Endeavors: Agricultural Protectionism and North-South Preferential Trade Agreements 51
ALLISON PHILLIPS
Georgetown University

Migration and FDI: Is There a Dynamic Bilateral Complementary Relationship between Foreign Tertiary Student Enrollment and FDI Outflows? 66
JENNIFER NOH
Georgetown University

The Determinants of International Terrorism: Its Perpetrators and their Targets 75
DAVID WOLFF
Dartmouth College
# Table of Contents

The Indirect Impact of Obesity on Economic Growth  
**Tanja Groth**  
*University of St. Andrews*  
83

The Determinants of Latin American Immigrants’ Settlement Locations in the Southwestern United States  
**Ashley Halpin**  
*Dartmouth College*  
89

The Effect of Microfinance Institutions on Income Inequality: Evidence from Ghana  
**R. Priya Mathew**  
*Washington University*  
100

What is the Effect of the Abundance of Natural Resources on Income Distribution?  
**Andrew O’Brien-Penney**  
*Georgetown University*  
111

The Conflicting Effects of Fiscal and Political Decentralization on Corruption: A Cross Country Study  
**Jennifer Xi**  
*Dartmouth College*  
117

Exchange Rate Policy and Economic Growth  
**Yi Sun**  
*Georgetown University*  
126

Joining the European Union: The Real Effect of Foreign Direct Investment  
**Lucia Franzese**  
*Georgetown University*  
137

Determinants of Sovereign Risk Premia in Commodity-Dependent Latin American Countries  
**Stephen Brinkmann**  
*Georgetown University*  
146

Returning Misplaced Mail: Underlying Motives and Influential Factors  
**Abdulla Humaidan**  
*University of Warwick*  
153

The Impact of Stock Market Liquidity on Economic Growth: An Empirical Analysis  
**Angelica da Rocha**  
*University of Warwick*  
167
# Table of Contents

Intrahousehold Bargaining and Divorce Norms in Indonesia  
**Nicholas Hartman**  
*Georgetown University*  

The Effects of the Sarbanes-Oxley Act on U.S. Merger Activity  
**Freddy Tsai**  
*University of British Columbia*  

Oil Price Shocks and Monetary Policy Revisited: An Energy-Augmented New Keynesian DSGE Model for the Analysis of the Role of Systematic Monetary Policy in the Oil Price-Macroeconomy Relationship  
**Sören Radde**  
*University of Bayreuth*  

Fitting Martingale with Horseshoes for a Deterministic Walk towards Chaos: A Model of Heterogeneous Expectations among Traders  
**Erik Eggum**  
*University of Warwick*  

Appendix A: Sixth Annual Carroll Round Presentation Schedule  
Appendix B: Past Speakers  
Appendix C: Former Carroll Round Steering Committees  
Appendix D: Members of the Advisory Panel  
Appendix E: Past Participants  

Note on Paper Submissions
ACKNOWLEDGEMENTS

Over the past six 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 150 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 2006-2007, as well as offered their strategic vision for the evolution of the Conference. We give special thanks to Mr. Mario Espinosa, Mr. Dennis Huggins, Mr. and Mrs. John and Anne Kelly, Mr. and Mrs. Kenneth Kunkel, Mr. Oleg Nodelman, Mr. Scott Pedowitz, and Mr. Geoffrey Yu. 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.

We also had the utmost privilege to share musical moments with The Contractions
(http://contractions.marginalq.com): Dr. Morris Davis (University of Wisconsin-Madison), Dr. Gwen Eudey (University of Pennsylvania), Dr. Peter Norman (University of North Carolina-Chapel Hill), Dr. Martin Schindler (International Monetary Fund), Dr. Andrei Shevchenko (Michigan State University) and Dr. Randall Wright (University of Pennsylvania).

Speakers have shared their wisdom with our conference participants throughout the years. We thank Dr. Grant Aldonas of the Center for Strategic and International Studies, Dr. François Bourguignon of the World Bank, and Dr. Randall Kroszner of the Federal Reserve Board of Governors for joining us for the 2007 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
Dr. Daniel Powers
Dean Robert L. Gallucci, SFS
Ms. Kendra Baity

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

Dr. James Albrecht
Dr. Phil Cross
Dr. Robert Cumby
Dr. Robin King
Dr. Rodney Ludema
Dr. Anders Olofsgard
Dr. Susan Vroman
Dr. Holger Wolf

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 and Mr. Geoffrey Yu.

Lastly, we thank Dean Mitch Kaneda, the Carroll Round Faculty Advisor, without whom this endeavor would not be possible.
Whenever I am asked about the history of the Carroll Round, a story about Oxford and a pub called the Radcliffe Arms pub usually arises. While there is truth in this tale, 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 in 2001.

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 the 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 young 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 made 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 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 universities and the latest research we were conducting at Oxford. As thesis and postgraduate plans matured during these conversations, I appreciated ever more my exposure to alternative experiences and approaches to scholarship. As the year progressed, I worried that these exciting connections would dissolve upon return to the United States.

One evening at the start of my final term in Oxford, I thought about the importance of this dialogue and my 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 concluded 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 a memo that outlined my solution: the Carroll Round. The following paragraph from that proposal captures my motivating thoughts:

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

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 students: Cullen Drescher, Mark Longstreth, Waheed Sheikh, and future Chair Meredith Gilbert to encourage younger students and ensure continuity for the future.

With the unflagging assistance of 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 Vice Chairman 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 at Cafe Milano, and Dr. Edwin M. Truman, former Assistant Secretary of the U.S. Treasury for International Affairs, closed the conference with words of wisdom to students considering careers in academia and policymaking.

A total of 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 Galvez, Avinash Kaza, 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 or surpassed the sophistication of recent graduate-level dissertations. Judging by their comments, the conference brought together some of the best young prospects in economics as they approached the frontiers of research.

I never imagined in March 2001 that the first Carroll Round would attain the heights realized one year later, or for that matter even exist. 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, Stephen Brinkmann, and Yasmine Fulema 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
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
CARROLL ROUND PROCEEDINGS
The Sixth Annual Carroll Round
Undergraduate Economics Conference
HOW DO INFORMAL AND FORMAL RESTRICTIONS ON WOMEN AFFECT ECONOMIC PERFORMANCE?

Adrienna Huffman
Washington University

ABSTRACT

This paper explores how informal and formal restrictions on women affect economic performance. Two indices, created by Jutting and Morrisson (2004), are used to measure the informal and formal restrictions on women in a dataset containing information from 65 developing nations in 2000. The results of the estimated models show that informal restrictions, such as the percentage of women married under the age of 20, and formal restrictions, a property rights structure which assigns property rights to the husband, father, or brother, have a significant and negative affect on the economic performance of the nation, measured by the log of GDP per capita, even when controlling for regional variation. Surprisingly, the model estimates the formal restriction, limiting women’s freedom of movement and dress, to have a significant but positive relationship to the nation’s income. However, when controlling for regional variations, the formal restriction limiting women’s freedom of movement and dress loses significance; no individual region is conclusively found to drive the restrictions affect on GDP per capita for the 65 nations in the dataset.

I. Introduction

This paper explores how informal and formal restrictions on women affect economic performance. In the literature regarding women’s status, economic development was at first considered key to reducing the inequalities between men and women and therefore affording women more economic opportunities. However, other research established that
social institutions were fundamental in constraining women’s status and their participation in the market. More recent research has focused on finding valid indicators to measure gender inequalities as a means to address the origins of such discrimination. These indicators have placed emphasis on access to education, healthcare, and family planning as avenues to increase women’s participation in the formalized labor sector, consequently increasing their economic independence. However, Jutting and Morrisson (2004) found that informal and formal restrictions on women, in a dataset of 65 developing countries, affected their access to resources like education, healthcare, and family planning, and therefore limit their economic opportunities. This paper extends the Jutting and Morrisson (2004) analysis by examining the impact upon GDP per capita of the formal and informal restrictions.

Developmental programs and neoclassical theory suggest that increasing women’s economic opportunities and participation in the market will increase an economy’s productive capabilities and performance. The results of the estimated models show that informal restrictions, such as the percentage of women married under the age of 20, and formal restrictions, a property rights structure which assigns property rights to the husband, father, or brother, have a significant and negative affect on the economic performance of the nation, even when controlling for regional variation. Surprisingly, the model estimates the formal restriction, limiting women’s freedom of movement and dress, to have a significant but positive relationship to the nation’s income. However, when controlling for regional variations, the formal restriction limiting women’s freedom of movement and dress loses significance; no individual region is conclusively found to drive the restrictions affect on GDP per capita for the 65 nations in the dataset.

II. Background

Informal restrictions, like informal institutions, can be defined as norms of behavior, conventions, traditions; these institutions are decentralized and self-enforced by a community (North 1994). Formal restrictions, like formal institutions, however, are rules that humans devise, such as written rules, laws, and constitutions (North 1990; North 1994). O. Williamson (2000) devised four hierarchal-levels of institutions for social analysis. Williamson’s model shows hierarchal, yet interactive, levels of institutions, in which a higher level imposes constraints on the level immediately below it, demonstrated via solid arrows, and reverse arrows connecting lower levels with higher levels signal feedback. Additionally, Williamson includes time horizons for the frequency of change for institutions at each level, where level one institutions, such as customs, traditions, and religion, are very slow to change, from centuries to millennia, level two institutions, like property rights, decades to centuries, and so forth. The time horizons for change designate the location of the institutional level in the hierarchy.

The first and second levels of Williamson’s model formulate the two Jutting and Morrisson (2004) indices measuring informal and formal restrictions on women used in this paper. The first index is called NON-ECO and measures religious and cultural norms, or informal restrictions with regard to women in each of the 65 developing countries.
included in the dataset. The main assumption is that these norms and customs persist for centuries and are slow to change, therefore constituting level one institutions. The second index is called ECO and measures the institutional environment, or the formal restrictions, specifically the property and inheritance rights and freedom of movement and dress, constraining women in each of the 65 developing countries. These constitutional or legal constraints change more quickly than the norms.

As North (1991) describes, and Williamson applies to his model, the informal restrictions that structure political, economic, and social interactions are located, then, within level one institutions, and the formal restrictions are located in level two institutions. The variables comprising the ECO index are the informal restrictions, which as North proposes, structure the formal restrictions of the game, or the variables in the NON-ECO index. This is reflected in the hierarchal design of Williamson’s model: informal restrictions, like religion and cultural norms, shape the design of formal restrictions, like property and inheritance rights, in societies. For example, Bina Agarwal (1994) found that the economic inequalities between men and women in South Asia were intrinsically related to their different rights in land, a formal restriction, and that these differences in land rights were linked to the culture’s gender ideologies, an informal restriction. The correlation matrix (Table 1) containing the variables of the two indices and the GDPPC, show that the correlations between the informal variables and the formal variables are high: the informal restrictions upon women are strongly correlated with the design of the formal restrictions in the developing countries.

III. Jutting and Morrisson Indices

Jutting and Morrisson (2004), using Williamson’s institutional model, calculate two indices measuring women’s discrimination. They use variables quantifying informal and formal restrictions upon women in 65 developing countries. The Non-Economic index, named by Jutting and Morrisson (2004), or NON-ECO, gauges three variables that, in their view, have a non-economic character and have pervaded historically, to determine the informal restrictions with regard to women for each of the 65 nations: female genital mutilation (FGM), marriage before the age of 20 (MARRIAGE), and polygamy (POLYG). All of these variables are constraints or opportunities which have economic consequences. For example, the percentage of women married before the age of 20 is a result of many underlying economic issues. However, the original authors used this nomenclature and for simplicity, this paper will refer to the index as NON-ECO, although this terminology is erroneous.

According to Jutting and Morrisson (2004), all variables were selected under the assumption that “these customs constrain women’s freedom to choose the economic activities they wish to pursue” (16). The FGM and MARRIAGE variables are continuous and the POLYG variable is dichotomous (Table 1.1). The coding of POLYG refers to its legality not its frequency because countrywide estimates of how many women live in polygamous households are not available (Jutting and Morrisson, 16). The variables Jutting and Morrisson (2004) use in their Economic index, or ECO, to quantify formal restrictions,
with regard to women are: property (PROP) and inheritance rights (INHERIT), and freedom of movement and dress (FREEDOM).

**IV. Theory**

Informal and formal restrictions on women constrain their access to resources and economic opportunity (Jutting and Morrisson). These restrictions affect how women participate in the formal sector of the economy and may therefore hinder a nation’s economic growth and performance. Williamson’s (2000) model asserts that level one institutions constrain and shape level two institutions, or that informal restrictions shape formal restrictions in societies. Applying this theory to restrictions on women’s status, using the two indices created by Jutting and Morrisson (2004) as measures of informal and formal restrictions upon women in developing countries, this paper formulates a model to examine whether restrictions upon women in developing nations affect the countries’ economic performance, as measured by GDP per capita.

Much of the theory in gender and economic development sought to find valid indicators for gender inequality. The validity of previous indices measuring gender inequalities has been widely debated and questioned (Dijkstra; Bardhan and Klasen; Dijkstra and Hanmer; White). Jutting and Morrisson (2004) developed their indices in response to the criticism raised against the United Nations Development Programme indicators, specifically the Gender-related Development Index and the Gender Empowerment Measure, because of their failure to grasp the institutional frameworks that constrain the economic role of women. The previous indices measure the results of gender discrimination rather than understand the principal causes. Consequently, the Jutting and Morrisson (2004) indices are, at present, more expansive than existing indices in their inclusion of institutional and economic variables which disadvantage women.

**A. Dependent Variables**

The dependent variable is the economic performance of the developing countries, measured by the log of GDP per capita, to compare the relative levels of income for the nations. Although not an exhaustive measure of economic performance, it is the data available for a cross-sectional analysis.

**B. Independent Variables**

The independent variables are the three components of Jutting and Morrisson’s (2004) NON-ECO index, or the informal restrictions on women, and the components of the ECO index, or the formal restrictions on women. The informal restrictions on women variables include: FGM, MARRIAGE, and POLYG. AUTHORITY was eliminated due to its high correlation with POLYG.

The second set of independent variables is the formal restrictions on women: inheritance rights (INHERIT), property rights (PROP), and freedom of movement and dress (FREEDOM).

Additionally, sex ratios will be included as an instrument for the MARRIAGE variable in the instrumental variables regressions and four dummy variables controlling for regional variation will be included in the fixed effects model. The four dummy regional vari-
ables are: Latin America, Asia, the Middle East, and Sub-Saharan Africa.

**C. Model**

Based upon the framework developed above, the following least squares models are estimated:

\[
\begin{align*}
\text{Log}(\text{GDPPC}) &= \beta_0 + \beta_1(\text{FGM}) + \beta_2(\text{MARRIAGE}) + \beta_3(\text{POLYG}) \\
\text{Log}(\text{GDPPC}) &= \beta_0 + \beta_1(\text{FREEDOM}) + \beta_2(\text{PROP}/\text{INHERIT}) \\
\text{Log}(\text{GDPPC}) &= \beta_0 + \beta_1(\text{FGM}) + \beta_2(\text{MARRIAGE}) + \beta_3(\text{POLYG}) + \\
&\quad \beta_4(\text{FREEDOM}) + \beta_5(\text{PROP}/\text{INHERIT})
\end{align*}
\]

Following the least squares estimates, models will be generated for the three equations standardizing the coefficients to examine the relative impact of each independent variable on economic performance. Subsequently, equations (1) and (3) will be estimated with the instrumental variable for MARRIAGE and equation (3) will be estimated including the regional dummy variables in the fixed effects model. Generally, with the variables contained in the Jutting and Morrisson (2004) indices, the closer the variable’s value is to one, the less equality between genders in that institution. When the variables are coded zero, this is interpreted as gender equality with respect to that social institution. This understanding is necessary to predict the results for the estimated model.

**V. Results**

**A. Least Squares and Standardized Estimates**

The least squares results from estimating equation (1), see Table 3.1, show only MARRIAGE as significant at the .01 level. The equation’s adjusted r-squared is 0.45. When the model is estimated standardizing the coefficients, see Table 1, the beta for the MARRIAGE variable is -0.622: an increase in one standard deviation of MARRIAGE results in an average decrease of log(GDPPC) by approximately 63 percent. A high percentage of women married under the age of 20 negatively affects the income of the developing countries contained in the dataset.

The least squares results from estimating equation (2) including PROP, see Table 4.1, show both FREEDOM and PROP significant at the .01 level. The equation’s adjusted r-squared is 0.35. Surprisingly, the coefficient for FREEDOM is positive, unlike the expected negative coefficient for PROP. When standardizing the coefficients, see Table 4.3, the beta for FREEDOM is 0.343 and the beta for PROP is -0.632. Property rights structures which disadvantage women by discouraging them from production and exchange, negatively impact the income of the developing nations. However, constraining women’s freedom of movement and dress is found to positively impact the nation’s relative change in income.

When equation (2) is estimated with least squares including INHERIT, the adjusted r-squared of the equation falls to 0.11, though both FREEDOM and INHERIT are significant at the 0.01 level. Equation (2) including PROP is a better fit for the estimated mod-
els. Yet, when INHERIT is included, the coefficient for FREEDOM remains positive. Unexpectedly, the model estimates imply that constraining women’s freedom of movement and dress positively affects the income of the countries contained in the dataset. This is inconsistent with the original hypothesis.

Because the model including PROP for equation (2) was a better fit than when equation (2) was estimated with INHERIT, equation (3) is estimated only with the inclusion of PROP (Tables 5). The adjusted r-squared is 0.56 and MARRIAGE, PROP, and FREEDOM are all significant at the 0.01 level. Again the coefficient for FREEDOM is positive and the remaining coefficients have the expected negative signs. When the coefficients are standardized, a single standard deviation in PROP affects the economic performance of the country negatively, by approximately 38 percent; for FREEDOM positively by approximately 29 percent; and for MARRIAGE negatively by approximately 19 percent.

The results from the estimates show FREEDOM positively affecting a country’s economic performance. Yet, formal restrictions on women’s movement and dress, in theory, should negatively affect the economic performance because it constrains women’s economic participation and opportunity.

The mean GDPPC for data available in 59 countries is $3237.73 with a standard deviation of $2977.79. The large standard deviation suggests a skewed distribution of GDPPC amongst the included countries. When restricting these 59 countries to nations which have formal restrictions on women’s dress and movement, there remain only 18 countries, mostly Middle Eastern and African nations. The mean GDPPC for the 18 countries is $4021.22 and the standard deviation in income is $3326.60, largely due to four Sub-Saharan African nations with GDP per capitas less than $1000. The average income for the countries with formal restrictions on women’s movement and dress is higher than the average for the dataset including all 59 nations.

Many of the Middle Eastern nations, in the set of 18 which restrict women’s movement and dress, are oil producers, like Saudi Arabia and the United Arab Emirates; this may explain their relatively high GDPs per capita. For example, the United Arab Emirates is one of the 18 countries which restricts women’s movement and freedom of dress and it has the highest GDPPC of all the countries included in the dataset, at $13,857. Consequently, these relatively higher incomes may be skewing the influence of restrictions on women’s movement and dress to positively affect changes in a nation’s GDPPC because the countries are wealthier to begin with: their average GDPPC is higher than the average including the rest of the dataset.

**B. Instrumental Variable and Fixed Effects Estimates**

When estimating equation (1) with the instrumental variable, sex ratios, for MARRIAGE (Tables 6), the coefficient, standard error, and significance at the 0.01 level for the MARRIAGE variable are the same as in the least squares estimates. This is also true when equation (3) is estimated with the instrumental variable for MARRIAGE: the coefficient, standard error and significance of MARRIAGE at the 0.01 level remain the same as in the least squares estimates.

However, when controlling for regional variation with the estimation of equation (3),
not only does the adjusted r-squared increase to 0.59, but the FREEDOM variable has lost significance and PROP has decreased in significance to the 0.05 level. MARRIAGE is still significant at the 0.01 level. Additionally, the dummy regional variable for Asia is significant at the 0.05 level. Thus, when controlling for regional variation, FREEDOM is no longer significant in affecting the countries’ incomes. The countries in the dataset located in Asia are found to be significant in driving the affect restrictions on women have on GDP per capita.

However, when equation (3) is estimated with each of the four regional variables, only the Middle East dummy is significant at the 0.05 level: the dummy for Asia is no longer significant when included alone. When the Middle East regional variable is included, FREEDOM expectedly loses its significance but PROP and MARRIED are still significant. In all four estimations of equation (3), with the inclusion of each regional dummy variable, MARRIAGE and PROP are always significant. However, the affects of regional variation are ambiguous. When all regional dummies are included, Asia is found significant in the estimation of equation (3). But when the regional dummies are included individually in equation (3), only the Middle East dummy is found significant. Therefore, there is not one region driving the affects of formal and informal restrictions on women on GDP per capita for the 65 countries contained in the dataset.

VI. Conclusions

Using two indices created by Jutting and Morrisson (2004) to gauge informal and formal restrictions on women, this paper explores the influence of these restrictions upon economic performance, as measured by the log of GDPPC. The dataset includes 65 developing nations in a cross-sectional analysis for 2000. The results from the estimated models show that high percentages of women married under the age of 20 and a property rights structure which disadvantages women negatively affects the incomes of the developing nations. Surprisingly, formal restrictions upon women’s freedom of movement and dress are found to positively impact a nation’s economic performance. However, when controlling for regional variation, FREEDOM loses its significance in affecting GDP per capita. When sex ratios are instrumented for MARRIAGE, there is no difference in the variable’s coefficient, standard error, or significance from the least squares estimates. The models estimated to control for regional variations found no individual region conclusively driving the restrictions affect on GDP per capita for the 65 nations in the dataset. PROP and MARRIAGE are significant in all models, even when controlling for regional variations.

These results provide a first round estimate of the effect of restrictions on women upon economic performance. They do not, however, conclude that informal and formal restrictions on women cause the economic performance of the included nations. Because cross-sectional analysis is limited in scope, future research in this area should focus on credible instruments to ascertain causality. Part of this could come from historical data. Then a time series model could be estimated to provide a richer perspective of the influence restrictions upon women have on the economic performance of nations over time.
References


RESTRICTIONS ON WOMEN

TABLE 1: CORRELATION MATRIX

<table>
<thead>
<tr>
<th></th>
<th>FGM</th>
<th>FREEDOM</th>
<th>GDPPC</th>
<th>INHERIT</th>
<th>MARRIED</th>
<th>POLYG</th>
<th>PROP</th>
</tr>
</thead>
<tbody>
<tr>
<td>FGM</td>
<td>1.00</td>
<td>0.15</td>
<td>-0.22</td>
<td>0.37</td>
<td>0.18</td>
<td>0.52</td>
<td>0.22</td>
</tr>
<tr>
<td>FREEDOM</td>
<td>0.15</td>
<td>1.00</td>
<td>0.17</td>
<td>0.42</td>
<td>0.04</td>
<td>0.47</td>
<td>0.34</td>
</tr>
<tr>
<td>GDPPC</td>
<td>-0.22</td>
<td>0.17</td>
<td>1.00</td>
<td>-0.20</td>
<td>-0.51</td>
<td>-0.23</td>
<td>-0.39</td>
</tr>
<tr>
<td>INHERIT</td>
<td>0.37</td>
<td>0.42</td>
<td>-0.20</td>
<td>1.00</td>
<td>0.28</td>
<td>0.78</td>
<td>0.66</td>
</tr>
<tr>
<td>MARRIED</td>
<td>0.18</td>
<td>0.04</td>
<td>-0.51</td>
<td>0.28</td>
<td>1.00</td>
<td>0.32</td>
<td>0.44</td>
</tr>
<tr>
<td>POLYG</td>
<td>0.52</td>
<td>0.47</td>
<td>-0.23</td>
<td>0.78</td>
<td>0.32</td>
<td>1.00</td>
<td>0.60</td>
</tr>
<tr>
<td>PROP</td>
<td>0.22</td>
<td>0.34</td>
<td>-0.39</td>
<td>0.66</td>
<td>0.44</td>
<td>0.60</td>
<td>1.00</td>
</tr>
</tbody>
</table>

TABLE 2: MEASUREMENT OF INDEPENDENT VARIABLES

<table>
<thead>
<tr>
<th>NON-ECO</th>
<th>ECO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.) FGM</td>
<td>1.) INHERITANCE</td>
</tr>
<tr>
<td>% of women excision is practiced upon</td>
<td>Unequal inheritance for different sexed children or widow does not inherit after husband’s death = 1, equality in inheritance = 0</td>
</tr>
<tr>
<td>2.) MARRIAGE</td>
<td>2.) PROP</td>
</tr>
<tr>
<td>% of women 15-19 married</td>
<td>Weighted: 0.3 for access to bank loans; 0.3 access to property not including land; 0.4 access to land</td>
</tr>
<tr>
<td>3.) POLYG</td>
<td>Variable =1 when women have no rights</td>
</tr>
<tr>
<td>If legal = 1, if not = 0</td>
<td>3.) FREEDOM</td>
</tr>
<tr>
<td></td>
<td>Weighted 0.5 for prohibitions in clothing and 0.5 for movement</td>
</tr>
</tbody>
</table>
### TABLE 3.1: LEAST SQUARES RESULTS FOR EQUATION (1)

<table>
<thead>
<tr>
<th>LS Estimations</th>
<th>Married</th>
<th>FGM</th>
<th>Polyg</th>
<th>Adj-rsquared</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 n=59</td>
<td>-4.23**</td>
<td>(0.633)</td>
<td></td>
<td>0.43</td>
</tr>
<tr>
<td>2 n=59</td>
<td>-4.01**</td>
<td>-0.624</td>
<td>(0.325)</td>
<td>0.46</td>
</tr>
<tr>
<td>3 n=59</td>
<td>-3.97**</td>
<td>-0.581</td>
<td>-0.055</td>
<td>0.45</td>
</tr>
</tbody>
</table>

*=significant at .05 level  **=significant at .01 level

### TABLE 3.2: STANDARDIZED COEFFICIENT RESULTS FOR EQUATION 1

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>22.34999</td>
<td>3</td>
<td>7.44999</td>
</tr>
<tr>
<td>Residual</td>
<td>24.7748</td>
<td>55</td>
<td>0.45045</td>
</tr>
<tr>
<td>Total</td>
<td>47.1248</td>
<td>58</td>
<td>0.81249</td>
</tr>
</tbody>
</table>

### TABLE 3.3

<table>
<thead>
<tr>
<th>Log(GDPPC)</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t</th>
<th>P&gt;abs(t)</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>POLYG</td>
<td>-0.0548</td>
<td>0.240</td>
<td>-0.23</td>
<td>0.820</td>
<td>-0.0272</td>
</tr>
<tr>
<td>FGM</td>
<td>-0.581</td>
<td>0.377</td>
<td>-1.54</td>
<td>0.129</td>
<td>-0.176</td>
</tr>
<tr>
<td>MARRIED</td>
<td>-3.966</td>
<td>0.659</td>
<td>-6.01</td>
<td>0.000</td>
<td>-0.622</td>
</tr>
<tr>
<td>constant</td>
<td>8.724</td>
<td>0.176</td>
<td>49.46</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 4.1: LEAST SQUARES RESULTS FOR EQUATION (2)

<table>
<thead>
<tr>
<th>LS Estimations</th>
<th>FREEDOM</th>
<th>PROP</th>
<th>INHERITANCE</th>
<th>Adj-rsquared</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 n=59</td>
<td>0.423</td>
<td>(0.441)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 n=59</td>
<td>1.152**</td>
<td>(0.380)</td>
<td>-1.668**</td>
<td>0.35</td>
</tr>
<tr>
<td>3 n=59</td>
<td>0.966*</td>
<td>(0.458)</td>
<td>-0.716**</td>
<td>0.108</td>
</tr>
</tbody>
</table>

*=significant at .05 level  **=significant at .01 level
RESTRICTIONS ON WOMEN

TABLE 4.2: STANDARDIZED COEFFICIENT RESULTS FOR EQUATION 2
SPECIFICALLY FOR ESTIMATION 2 IN ABOVE TABLE INCLUDING PROP

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>17.3244</td>
<td>2</td>
<td>8.6622</td>
</tr>
<tr>
<td>Residual</td>
<td>29.8004</td>
<td>56</td>
<td>0.5322</td>
</tr>
<tr>
<td>Total</td>
<td>47.1248</td>
<td>58</td>
<td>0.8125</td>
</tr>
</tbody>
</table>

Log(GDPPC) | Coefficient | Std. Error | t    | P>abs(t) | Beta  |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FREEDOM</td>
<td>1.152</td>
<td>0.380</td>
<td>3.03</td>
<td>0.004</td>
<td>0.343</td>
</tr>
<tr>
<td>PROP</td>
<td>-1.668</td>
<td>0.299</td>
<td>-5.58</td>
<td>0.000</td>
<td>-0.632</td>
</tr>
<tr>
<td>constant</td>
<td>8.111</td>
<td>0.138</td>
<td>58.76</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 4.4: STANDARDIZED COEFFICIENT ESTIMATE FOR EQUATION (2),
SPECIFICALLY FOR ESTIMATION 2 IN ABOVE TABLE INCLUDING PROP

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>6.5554</td>
<td>2</td>
<td>3.2778</td>
</tr>
<tr>
<td>Residual</td>
<td>40.5693</td>
<td>56</td>
<td>0.7245</td>
</tr>
<tr>
<td>Total</td>
<td>47.1248</td>
<td>58</td>
<td>0.8125</td>
</tr>
</tbody>
</table>

Log(GDPPC) | Coefficient | Std. Error | t    | P>abs(t) | Beta  |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FREEDOM</td>
<td>0.967</td>
<td>0.458</td>
<td>2.11</td>
<td>0.040</td>
<td>0.288</td>
</tr>
<tr>
<td>INHERITANCE</td>
<td>-0.716</td>
<td>0.253</td>
<td>-2.83</td>
<td>0.006</td>
<td>-0.386</td>
</tr>
<tr>
<td>constant</td>
<td>7.939</td>
<td>0.166</td>
<td>47.83</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>
### Table 5.1: Least squares and standardized estimations for Equation (3)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>8.678</td>
<td>0.158</td>
<td>0.000</td>
</tr>
<tr>
<td>POLYG</td>
<td>0.046</td>
<td>0.270</td>
<td>0.866</td>
</tr>
<tr>
<td>FGM</td>
<td>-0.633</td>
<td>0.341</td>
<td>0.069</td>
</tr>
<tr>
<td>MARRIED</td>
<td>-3.074**</td>
<td>0.630</td>
<td>0.000</td>
</tr>
<tr>
<td>PROP</td>
<td>-0.989**</td>
<td>0.308</td>
<td>0.002</td>
</tr>
<tr>
<td>FREEDOM</td>
<td>0.983**</td>
<td>0.338</td>
<td>0.005</td>
</tr>
</tbody>
</table>

### Table 5.2: Standardized coefficient results for equation (3)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>28.2606</td>
<td>5</td>
<td>5.6521</td>
</tr>
<tr>
<td>Residual</td>
<td>18.8642</td>
<td>53</td>
<td>0.3559</td>
</tr>
<tr>
<td>Total</td>
<td>47.1248</td>
<td>58</td>
<td>0.8125</td>
</tr>
</tbody>
</table>

### Table 5.3

| Log(GDPPC) | Coefficient | Std. Error | t     | P>|abs(t)| Beta |
|------------|-------------|------------|-------|-------|------|
| FREEDOM    | 0.983       | 0.338      | 2.91  | 0.005 | 0.293|
| PROP       | -0.989      | 0.308      | -3.21 | 0.002 | -0.375|
| POLYG      | 0.046       | 0.270      | 0.17  | 0.865 | 0.023|
| FGM        | -0.633      | 0.341      | -1.86 | 0.069 | -0.192|
| MARRIED    | -3.074      | 0.630      | -4.88 | 0.000 | -0.482|
| Constant   | 8.678       | 0.158      | 55.04 | 0.000 |      |
Table 6.1: Instrumental variable regression for Equation (1)

<table>
<thead>
<tr>
<th></th>
<th>Coef</th>
<th>Std Error</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARRIED</td>
<td>-3.966**</td>
<td>0.659</td>
<td>0.000</td>
</tr>
<tr>
<td>POLYG</td>
<td>-0.055</td>
<td>0.240</td>
<td>0.820</td>
</tr>
<tr>
<td>FGM</td>
<td>-0.581</td>
<td>0.377</td>
<td>0.129</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>8.724</td>
<td>0.176</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 6.2: Instrumental variable regression for Equation (3)

<table>
<thead>
<tr>
<th></th>
<th>Coef</th>
<th>Std Error</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>FREEDOM</td>
<td>0.983**</td>
<td>0.338</td>
<td>0.005</td>
</tr>
<tr>
<td>PROP</td>
<td>-0.989**</td>
<td>0.308</td>
<td>0.002</td>
</tr>
<tr>
<td>POLYG</td>
<td>0.046</td>
<td>0.270</td>
<td>0.865</td>
</tr>
<tr>
<td>FGM</td>
<td>-0.633</td>
<td>0.341</td>
<td>0.069</td>
</tr>
<tr>
<td>MARRIED</td>
<td>-3.074**</td>
<td>0.630</td>
<td>0.00</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>8.678</td>
<td>0.158</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 6.3: Fixed effects regression controlling for all regions

<table>
<thead>
<tr>
<th></th>
<th>Coef</th>
<th>Std Error</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>FREEDOM</td>
<td>0.628</td>
<td>0.408</td>
<td>0.130</td>
</tr>
<tr>
<td>PROP</td>
<td>-0.837*</td>
<td>0.342</td>
<td>0.018</td>
</tr>
<tr>
<td>POLYG</td>
<td>-0.114</td>
<td>0.313</td>
<td>0.716</td>
</tr>
<tr>
<td>FGM</td>
<td>-0.672</td>
<td>0.340</td>
<td>0.054</td>
</tr>
<tr>
<td>MARRIED</td>
<td>-2.495**</td>
<td>0.712</td>
<td>0.001</td>
</tr>
<tr>
<td>I_REGION ASIA</td>
<td>-0.631*</td>
<td>0.300</td>
<td>0.040</td>
</tr>
<tr>
<td>I_REGION LATIN</td>
<td>-0.388</td>
<td>0.354</td>
<td>0.278</td>
</tr>
<tr>
<td>I_REGION SUB</td>
<td>-0.549</td>
<td>0.297</td>
<td>0.071</td>
</tr>
</tbody>
</table>
PHYSICAL, INFORMATIONAL AND HUMAN DISTANCE:
A COMPARISON OF THE DETERMINANTS OF BILATERAL PORTFOLIO AND FOREIGN DIRECT INVESTMENTS FLOWS

Matthew Pech
Dartmouth College

ABSTRACT

Using data from 14 developed countries, 1989-1996, this paper compares the use of gravity models in explaining the pattern of bilateral portfolio and foreign direct investment (FDI) flows, focusing on the effects of physical and informational distance. Host and source market capitalizations and GDP growth rates, as well as physical and informational distance, can be used to model FDI flows, but the fit is better for portfolio equity flows. This is consistent with the heterogeneous effect physical and informational distance has on the various forms of FDI. While both investment types are decreased by greater physical and informational distance, equity flows are more sensitive to changes in informational distance, whereas FDI flows are more sensitive to changes in physical distance. This is explained in terms of the importance of intermediate good flow, and the value of human interaction in FDI. Finding that informational distance can substitute for physical distance, and that the host country skilled-labor stock strongly attracts FDI (but not portfolio investment), provides some weak evidence for the human interaction hypothesis.

I. Introduction

During the 1990s, the world-wide flow of FDI grew tremendously, and many countries
enacted pro-FDI regulation changes to receive more of it. This speaks to the perceived benefit of being a host country to FDI, such as technological spillovers, and the spread of best practices in business (Razin, 2.) As a result, finding patterns to the global flows of FDI is of great interest.

The ability of gravity models to explain the trade of goods is well known. In Portes & Rey 2005, the authors extended its applicability to bilateral portfolio equity flows. Despite noting that transportation and transaction costs should be relatively small for these “weightless” assets, and despite the potential benefits of portfolio diversification, distance strongly reduced portfolio flows between countries. Its effect decreased, but remained significant, after controlling for a measure of informational capability, bilateral telephone traffic.

In modeling financial asset flows, the authors distinguished portfolio from direct investment by a 20-30% threshold (Portes and Rey, 293). This begs the question—are the results robust not only for higher portfolio investment thresholds, but all categories of foreign direct investment? Do gravity models explain the flow of FDI as well as they have for the goods and portfolio equity trade?

FDI is distinguished from portfolio investments by a number of factors. FDI is a different subcategory of capital flows, requiring a specific relationship between the transacting parties (IMF, 41). Namely, the direct investor must have the objective of establishing a “lasting interest” in the foreign-residing enterprise. The lasting interest should be long-term, and should reflect great influence of the direct investor in the host enterprise; somewhat arbitrarily, the investor must acquire at least 10% of ordinary shares or voting power to establish “lasting interest.” Once the relationship is started, FDI reflects the initial and all subsequent capital transactions between the two entities (OECD; IMF).

One could imagine that the determinants of portfolio equity flows are different from those of FDI flows—investors establishing a lasting interest in a foreign enterprise probably have different motivations, time horizons and expected returns than for other types of foreign investment. For this reason, it seems likely that the effect of (and/or sensitivity to) physical and informational distance will be different for FDI flows than portfolio equity flows.

A priori, it’s unclear what effect physical distance should have on FDI flows, as there are competing factors at play. If most FDI is horizontal, and transportation/time-in-transit costs are high, then horizontal FDI might increase with distance, as multinational firms undertake export-substituting FDI. However, one can imagine that there is a certain inherent complexity to horizontal FDI, making it require face-to-face contact. Such complexities include frequent design changes, implementing improvements in marketing and engineering, and monitoring the affiliate (Redding and Venables, 103). Distance would raise the cost of human interaction, thus decreasing FDI. The picture is as muddled when considering vertical FDI. The greater transportation costs associated with distance could discourage the shipment of the intermediate goods associated with vertical FDI (Lougani et al, 530). But, one could imagine relatively lower-cost inputs that drive vertical FDI to be positively correlated with distance.
The expected effect of informational distance is ambiguous as well. Informational capability should be more important for the longer term, less liquid investment such as FDI. It should make setting up and interacting with foreign subsidiaries easier, encouraging horizontal and vertical FDI. However, the same higher informational capability could make it easier to service the host market via exports, because domestic competitors’ traditional advantage (private knowledge about local markets) could be eroded (Borensztein et al, 133). This effect could dominate at low distances, when transport costs are low.

Given the pattern seen when modeling the flow of goods and portfolio equity, one expects FDI to be no different an asset: greater physical and informational distance should reduce FDI flows. However, with the aforementioned factors acting in opposition in FDI flows, one expects that the magnitude of the effects should be comparatively muted. Additionally, the effects of informational distance should vary with physical distance, in a way that doesn’t occur with portfolio equity.

II. Data, Variables, and Regression

The empirical work was performed on yearly panel data from 1989 to 1996, on bilateral FDI flows between 12 source countries and 14 host countries. The dataset attempts to replicate the country set found in the research of Portes and Rey (2005), but not all of the initial countries report FDI outflows. There were 1344 total potential observations (12x14x8), but due to missing data, n=1025 (or n=970 when only non-negative FDI flow values were considered).

The basic regression is specified as follows:

\[
\log(FDI\_outflow) = \alpha + \beta_1(marketcap_s) + \beta_2(marketcap_h) + \beta_3(gdp\_growth_s) + \beta_4(gdp\_growth_h) + \beta_5(dist) + \beta_6(telephone) + \beta_7(timedummy) + \beta_8(\ldots) + \epsilon
\]

Due to space constraints, detailed information on the data is not shown (information on this and all other omitted information is available upon request to the author), aside from the data sources.

Results

The initial regressions, including just the “mass” and “distance” variables, are given in table 2. The OLS regressions have standard errors clustered by host country to correct for serial correlation.

Column (1) gives the traditional “gravity” modeling of FDI outflows, including the host and source country mass variables (real market capitalizations and GDP growth rates) as well as physical distance. Distance and source country financial size and growth are highly significant and appropriately signed. Host country market cap is less significant, and, surprisingly, host country GDP growth is insignificant. The log-log specification implies that increasing the distance between a country pair 1% decreases FDI outflow by -0.5%.

Column (2) adds the proxy for information capability, bilateral telephone traffic (nor-
The variable has the expected coefficient, and is highly significant. Its addition makes the estimated magnitude of distance fall roughly 20%, while increasing the precision of estimation for all variables. This suggests that there was significant omitted variable bias—the negative correlation between telephone traffic and distance led to the overestimation of distance’s impact.

In column (3), the regression is tested for robustness against a linguistic tie dummy, as this could have been driving the effect of informational capability. This proves not to be the case, as its addition has no effect.

In columns (6) and (7), the regressions are re-estimated using host country fixed-effects. This method, by demeaning each variable, controls for any unobserved time-invariant heterogeneity between host countries. The results are qualitatively unchanged from the previous estimation method. The only difference is a further reduction in the point estimate of the telephone measure.

All in all, the initial gravity-style models follow the patterns seen in equity and good flows, as expected, except for the lack of significance in the host country GDP growth rate. Perhaps this macro-level statistic fails to capture the heterogeneity in growth potential at the firm/industry level. Alternatively, predictability of the growth rate may be more important than its actual magnitude.

The interesting results come from comparisons with the analogous regressions on portfolio equity flows. The basic mass and physical distance specification is in table 2, column (4), with the telephone traffic measure added in column (5). Its point elasticity is 57% higher for the equity flows than FDI flows (.321 vs. .136), while the negative elasticity of distance is 23% larger (-.51 vs. -.39) for equity flows. This qualitative examination is consistent with the hypothesis that FDI is less sensitive to distance measures, due to opposing factors at play. Adding the informational distance variable reduces the negative elasticity of distance 23% for FDI flows, 35% for equity flows, implying that distance may be a better (inverse) proxy for information capability in equity flows. The R² value from the FDI flow regression is lower than for equity flow, and rises much less after the addition of the telephone traffic measure. The inferior fit for FDI outflows is consistent with the hypothesis that the behavior of multinationals and FDI is diverse and complicated, and that “distance” is not uniformly good or bad (Blonigen, 21).

Next, the sensitivities of FDI and portfolio equity flows to physical and informational distance were compared with a statistical approach. This was accomplished by re-running the regressions using log (equity flows) minus log (FDI flows) as the dependent variable. In this specification, when a coefficient is negative and significant, it implies that the variable has a greater effect on FDI outflows than equity flows. The results, measured by OLS with host country clustered standard errors, or by host country fixed-effects, are shown in table 3, columns (1) and (2), respectively.

The regressions show that, like before, equity flows are more sensitive to changes in information capability, as the telephone measure is positive and highly significant, as is the linguistic tie (though this latter result is only significant under fixed-effects estimation). Interestingly, FDI flows are more sensitive to host and source GDP growth rates, whereas
portfolio equity flows are more sensitive to host and source market capitalization. It makes sense that investors take more into consideration their and their host country’s growth prospects when undertaking a long-term, relatively illiquid investment (Razin, 125). Additionally, if market capitalization represents to a certain degree sophistication and liquidity of the financial markets, it is understandable that portfolio flows are more sensitive to these factors.

Most importantly, the regression suggests that FDI flows are more sensitive to changes in distance (with much higher significance in the fixed effects estimation), going against the previous qualitative inspection. This also goes against the original hypothesis.

There are a couple potential explanations for this. One is that the intermediate good flow between these countries predominates, and lower cost inputs are only weakly correlated with distance. Given that the countries in the dataset are highly developed, this view is best supported in a “knowledge-capital model” framework, where corporations parse their knowledge-generating activities from production activities (Carr et al, 694). This hypothesis is also consistent with half of all trade flows being intra-firm (Blonigen, 1).

There is another potential reason for higher FDI flow sensitivity to distance. One can imagine that undertaking and sustaining horizontal FDI involves inherently difficult interactions, which cannot occur using information or communication technology. Such complexities that require human interaction include being able to reliably distinguish good from bad foreign investments; implementing frequent design changes or improvements in marketing and engineering; and monitoring the affiliate to reduce principal-agent problems. (Redding & Venables 2002 p103; Goldberg et al 2005 p915) In this case, distance proxies for greater human interaction cost, and its effect shouldn’t disappear when information capability is considered.

This hypothesis implies that smaller informational distance should weakly substitute for greater physical distance, or, as some have argued, actually complement it (Gaspar and Glaeser 1998, 136). To see how the negative effect of distance changes with higher informational capability, an interaction term between the physical and informational distance variables is added into the FDI outflow regression. The results, estimated by OLS with host-country clustered errors, or by host country fixed-effects, are presented in columns (1) and (2) of table 4. The table also gives the estimated range of the elasticities of distance and telephone traffic.

Upon addition of the interaction term, the telephone traffic elasticity is negative at low distance values, but eventually becomes positive at higher distances. Physical proximity can substitute for low informational capability. Interestingly, at low distances, higher information reduces FDI. Perhaps under proximity, having more informational capability makes it easier to service the market via exports. After a certain threshold distance, the normal relationship between information and FDI is re-established.

The distance elasticity is negative at low telephone traffic values, but less negative at higher telephone traffic values. This suggests that while increased information can substitute for distance, in this set of countries it can never completely make up for distance. This result weakly supports the human interaction cost hypothesis.
The results of introducing the interaction term into the portfolio equity regression are shown in columns (3) and (4) of table 4. The results are qualitatively similar to those in the FDI regression. The main difference is that greater informational capability isn’t a deterrent to equity flows at low distance values. It makes sense that fewer informational constraints are never bad for portfolio equity transactions, as there isn’t another investment type that would be substituted towards at low distances.

III. Robustness tests

As a test of robustness, additional host country characteristic covariates are added to the regression. Time-variant host country heterogeneity could be misattributed to the variables of interest, biasing the estimates. For space reasons, results are not shown. All in all, the distance and informational capability measures are materially unaffected. Only the political corruption index and, surprisingly, the higher education attainment variables are significant (and of the correct sign). The corruption result weakly parallels those seen in Wei (2000), despite all countries in this dataset scoring on the top half of the corruption scale.

The positive elasticity of higher education on FDI flows is, surprisingly, the strongest of all factors for attracting FDI. The result supports the hypothesis the “knowledge-capital model” hypothesis that FDI can vertically fragment to chase cheap skilled-labor inputs. It also supports earlier assertions of the importance of human interaction in the FDI decision—one could imagine that a more educated workforce would require less direct oversight from the source corporation. Consistent with all this, the human capital measure is highly insignificant in portfolio equity flow regressions (results not shown).

IV. Conclusions

This paper compared the role of distance in explaining the size and location of FDI and portfolio equity flows. Gravity models with various measures of distance do a much better job of explaining the variation in portfolio equity flows. This was explained in terms of the heterogeneous effect physical and informational distance can have on the various forms of FDI. The heterogeneity was plainly evident when the two types of distances were interacted, as greater informational capability was found to decrease FDI at low distances.

Portfolio equity flows were found to be more sensitive to changes in informational distance, whereas FDI flows were found to be more sensitive to physical distance. Two explanations were given for the latter result: intermediate good flow predominating between countries, or human interaction being important (and costlier with distance). Finding that informational distance can substitute for physical distance, and that the host skilled labor stock strongly attracts FDI, provides some weak evidence for the human interaction hypothesis.

There are a few caveats to the results. One is that the log specification of the dependent variable drops 120 negative FDI outflow values from the regression, which could
introduce biases. However, various tests show this wasn’t the case. Another potential bias is the nature of measuring foreign direct investment flows. FDI may underestimate the true extent of foreign investment in a subsidiary, as it doesn’t take into account any debt or equity raised in the host market. Systematic differences in this bias could affect the estimates (Borensztein et al, 134). Another issue is the extendibility of the results in this paper to more countries. Such an examination is seriously hampered by a lack of bilateral FDI flow reporting by non-OECD countries.

The results of the paper suggest some avenues for further research. The potential importance of intermediate good flow in explaining the sensitivity of FDI to physical distance could be explored with firm-level data. To further explore the role of human interaction in influencing FDI flows, a measure of airplane person-visits between source and host countries could be added.

References


Yeyati E.L., Stein E., & Daude C. (2002) “Regional Integration and the Location of FDI” Inter-American Development Bank Integration and Regional Programs Department Research Department.
**Table 1: Initial Comparison of FDI and Portfolio Equity Flows:**

<table>
<thead>
<tr>
<th>Estimation Method, and Dependent Variable</th>
<th>OLS† Host country fixed-effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(FDI_outflow$_{ij}$)</td>
<td>log(equity$_{ij}$)</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Log(market cap$_i$)</td>
<td>0.627</td>
</tr>
<tr>
<td>(0.029)***</td>
<td>(0.028)***</td>
</tr>
<tr>
<td>Log(market cap$_j$)</td>
<td>0.405</td>
</tr>
<tr>
<td>(0.209)*</td>
<td>(0.179)*</td>
</tr>
<tr>
<td>GDP growth$_i$</td>
<td>0.08</td>
</tr>
<tr>
<td>(0.022)***</td>
<td>(0.016)***</td>
</tr>
<tr>
<td>GDP growth$_j$</td>
<td>0.039</td>
</tr>
<tr>
<td>(-0.031)</td>
<td>(-0.03)</td>
</tr>
<tr>
<td>Log(distance$_{ij}$)</td>
<td>-0.511</td>
</tr>
<tr>
<td>(0.080)***</td>
<td>(0.076)***</td>
</tr>
<tr>
<td>Log(telephone$_{ij}$)</td>
<td>-</td>
</tr>
<tr>
<td>(0.030)***</td>
<td>(0.031)***</td>
</tr>
<tr>
<td>Linguistic Tie$_{ij}$</td>
<td>-</td>
</tr>
<tr>
<td>-0.228</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>970</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.479</td>
</tr>
</tbody>
</table>

† Standard errors clustered by host country, in parentheses
* significant at 10%, ** significant at 5%, *** significant at 1%

Constant, time dummies not shown.

Data Sources can be found in Table 1. Data is from 12 host countries, 14 source countries, 1989-1996. log(FDI_outflow$_{ij}$) is the log of real $m FDI outflow from the source i country to the host j country in time t. log(equity$_{ij}$) is the log of gross bilateral portfolio equity flows. log(market cap$_i$) and log(market cap$_j$) are the log of host and source country market real $m capitalizations, respectively. GDP growth$_i$ and GDP growth$_j$ are the source and host country GDP growth rates, in percentage terms. Log(distance$_{ij}$) is log of distance between the two countries’ capital cities. Log(telephone$_{ij}$) is the log of bilateral telephone traffic, in millions of call minutes, normalized by real GDP. Linguistic Tie$_{ij}$ equals 1 when host and source country share a common language, 0 otherwise.
### PHYSICAL, INFORMATIONAL AND HUMAN DISTANCES

#### Table 2: Relative Sensitivity of FDI Outflows vs. Portfolio Flows

<table>
<thead>
<tr>
<th></th>
<th>Estimation Method, and Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS† Host country fixed-effects</td>
</tr>
<tr>
<td></td>
<td>log(equity\textsubscript{ij}) – log(FDI\textsubscript{outflow}\textsubscript{ij})</td>
</tr>
<tr>
<td></td>
<td>(1) (2)</td>
</tr>
<tr>
<td>Log(market cap\textsubscript{i})</td>
<td>0.578 (0.069)***</td>
</tr>
<tr>
<td>Log(market cap\textsubscript{j})</td>
<td>0.706 (-0.192)***</td>
</tr>
<tr>
<td>GDP growth\textsubscript{i}</td>
<td>-0.154 (0.034)***</td>
</tr>
<tr>
<td>GDP growth\textsubscript{j}</td>
<td>-0.012 (-0.028)***</td>
</tr>
<tr>
<td>Log(distance\textsubscript{ij})</td>
<td>-0.157 (0.085)*</td>
</tr>
<tr>
<td>Log(telephone\textsubscript{ij})</td>
<td>0.147 (0.045)***</td>
</tr>
<tr>
<td>Linguistic Tie\textsubscript{ij}</td>
<td>0.465 (-0.269)***</td>
</tr>
<tr>
<td>Observations</td>
<td>970</td>
</tr>
<tr>
<td>R²</td>
<td>0.398</td>
</tr>
</tbody>
</table>

† Standard errors clustered by host country, in parentheses
* significant at 10%, ** significant at 5%, *** significant at 1%
Constant, time dummies not shown.

Data Sources can be found in Table 1. Data is from 12 host countries, 14 source countries, 1989-1996. log(equity\textsubscript{ij}) – log(FDI\textsubscript{outflow}\textsubscript{ij}) is the log of real $m FDI outflow from the source i country to the host j country in time t, subtracted from the log of gross bilateral portfolio equity flows. log(market cap\textsubscript{i}) and log(market cap\textsubscript{j}) are the log of host and source country market real $m capitalizations, respectively. GDP growth\textsubscript{i} and GDP growth\textsubscript{j} are the source and host country GDP growth rates, in percentage terms. Log(distance\textsubscript{ij}) is log of distance between the two countries’ capital cities. Log(telephone\textsubscript{ij}) is the log of bilateral telephone traffic, in million of call minutes, normalized by real GDP. Linguistic Tie\textsubscript{ij} equals 1 when host and source country share a common language, 0 otherwise.
### Table 3: Testing the Substitutability of Physical and Informational Distance

<table>
<thead>
<tr>
<th>Estimation Method, and Dependent Variable</th>
<th>OLS† Host country fixed-effects</th>
<th>OLS† Host country fixed-effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(FDI_outflow(_{ij}))</td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>(0.031)***</td>
<td>(0.021)***</td>
</tr>
<tr>
<td>Log(market cap(_i))</td>
<td>0.605</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>(0.031)***</td>
<td>(0.021)***</td>
</tr>
<tr>
<td>Log(market cap(_j))</td>
<td>0.394</td>
<td>0.287</td>
</tr>
<tr>
<td></td>
<td>(0.176)**</td>
<td>(0.156)*</td>
</tr>
<tr>
<td>GDP growth(_i)</td>
<td>0.068</td>
<td>0.062</td>
</tr>
<tr>
<td></td>
<td>(0.017)***</td>
<td>(0.017)***</td>
</tr>
<tr>
<td>GDP growth(_j)</td>
<td>0.006</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>(-0.029)</td>
<td>(-0.021)</td>
</tr>
<tr>
<td>Log(distance(_{ij}))</td>
<td>0.826</td>
<td>0.445</td>
</tr>
<tr>
<td></td>
<td>(0.446)*</td>
<td>(0.263)*</td>
</tr>
<tr>
<td>Log(telephone(_{ij}))</td>
<td>-0.302</td>
<td>-0.205</td>
</tr>
<tr>
<td></td>
<td>(0.153)*</td>
<td>(0.098)**</td>
</tr>
<tr>
<td>Log(distance(<em>{ij})) * Log(telephone(</em>{ij}))</td>
<td>0.052</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>(0.018)**</td>
<td>(0.011)***</td>
</tr>
<tr>
<td>Distance elasticity range, low to high informational distance</td>
<td>(-0.22, -0.85,)</td>
<td>(-0.28, -0.72)</td>
</tr>
<tr>
<td>Information elasticity range, low to high physical distance</td>
<td>(-0.02, 0.21)</td>
<td>(-0.009, 0.15)</td>
</tr>
<tr>
<td>Observations</td>
<td>970</td>
<td>970</td>
</tr>
<tr>
<td>R(^2)</td>
<td>0.528</td>
<td>0.507</td>
</tr>
</tbody>
</table>

† Standard errors clustered by host country, in parentheses
* significant at 10%, ** significant at 5%, *** significant at 1%

Constant, time dummies not shown.

Data Sources can be found in Table 1. Data is from 12 host countries, 14 source countries, 1989-1996. log(FDI_outflow\(_{ij}\)) is the log of real $m FDI outflow from the source \(_i\) country to the host \(_j\) country in time \(_t\). log(equity\(_{ij}\)) is the log of gross bilateral portfolio equity flows. log(market cap\(_i\)) and log(market cap\(_j\)) are the log of host and source country market real $m capitalizations, respectively. GDP growth\(_i\) and GDP growth\(_j\) are the source and host country GDP growth rates, in percentage terms. Log(distance\(_{ij}\)) is log of distance between the two countries’ capital cities. Log(telephone\(_{ij}\)) is the log of bilateral telephone traffic, in million of call minutes, normalized by real GDP. Log(distance\(_{ij}\)) * Log(telephone\(_{ij}\)) is the interaction term of the distance and telephone traffic variables.
HAS FOREIGN BANK ENTRY LED TO SOUNDER BANKS IN TRANSITION ECONOMIES? 
THE ESTONIAN EXPERIENCE

Heleri Rande
New York University

ABSTRACT

This paper examines the impact financial sector foreign direct investment (FSFDI) has had on the Estonian economy by comparing the soundness of banks operating in Estonia. Foreign participation in the financial sector increased rapidly during the 1990s in many countries in Central and Eastern Europe (CEE) and foreign banks now account for a major share of total banking assets. It is generally believed that reasonable foreign penetration leads to improved efficiency, greater financial stability and technology transfer. However, the concerns about managerial and operational control over the host country’s economy still remain. This paper uses bank-specific data from the Estonian banking sector from 2000 to 2005 to address this question and prove by comparison that small domestic and foreign-owned banks do not fall behind in financial stability performance of their big market share foreign counterparts, and that extensive foreign ownership is not always a source of threat for the economy.

I. Introduction

Foreign participation in the financial sector increased considerably during the 1990s in many emerging and transition economies of Eastern Europe. These foreign banks now account for more than 50 percent of total banking assets, sometimes even reaching the 80 percent threshold. Financial foreign direct investment (FDI) cross border mergers and acquisitions (M&A) surged from about US$6 billion between 1990-1996, to almost US$50 billion between 1997-2000 (Bank of International Settlements 2004).

Financial Sector FDI (FSFDI) has been primarily fostered by financial liberalization and market-based reforms in many Eastern European countries. After the collapse of the
Soviet Union, the three Baltic countries – Estonia, Latvia and Lithuania – underwent tremendous change and quickly adopted policies that welcomed foreign investment in all sectors of the economy. The banking sector in Estonia experienced the most radical change of ownership – from zero percent in 1990 to 99 percent in 2006 as measured by the percentage of total banking assets. The implications of such pace and depth are important to any small economy and therefore deserve further research.

The empirical literature has been limited on this topic. FSFDI has received far less attention than manufacturing FDI. There have been studies that have focused on the qualitative side and reached policy conclusions. Analysis on the quantitative side, however, has been limited to systematic bank efficiency effects studied by Martinez-Peria and Schukler (1999) and differences in lending patterns studied by Dages, Goldberg and Kinney (2000). A more comprehensive review will follow.

The greater differences in financial stability and risk sharing have been addressed by Crystal, Dages and Goldberg (2001) in their paper Does Foreign Ownership Contribute to Sounder Banks in Emerging Markets? The Latin American Experience. This is the study I will be utilizing to a great extent in my own research. By comparing the financial condition and performance of credit institutions in Estonia, most of them foreign-owned, while using the methodology of Crystal, Dages and Goldberg, I will hope to discover whether foreign banks with greater market share and resources have proven to be financially more sound over the selected five years or whether the differences are not as substantial as one might assume at the beginning.

I decided to focus on Estonia since the country is a perfect case study for market-based reforms that have lead to a considerable growth in FDI. At the end of 2004, the stock of FDI relative to GDP was close to 90 percent (Lane and Milesi-Ferretti). The trends in FDI inflows have changed over the last few years, turning the focus from manufacturing to financial intermediation. The latter now accounts for 47.4 percent of total FDI, while manufacturing only contributes 12.7 percent (Bank of Estonia).

Estonia’s policies have been very liberal in receiving investments from abroad. The Joint-Venture Act of 1987 already set standards to welcoming back Western companies, but the real wave didn’t start until a couple of years later. As of now, the law does not distinguish between citizens, residents and foreigners who want to buy assets in Estonia. Therefore, the relative ease of buying property has led to a substantial foreign ownership of both the manufacturing and financial sector.

The paper will focus on the years 2000-2005 because of the possibility of evaluating all of the banks that have operations in Estonia. Since the earlier years saw tremendous consolidation and bankruptcy activity in the banking sector due to the effects of the Russian Crisis, the analysis wouldn’t be as effective across all banks. All the following banks have majority ownership as of 2005. Hansapank belongs to the Swedish bank Swedbank formerly know as ForeningSparbanken (full ownership in 2005), SEB Eesti Yhispank to Swedish Skandinavska Enskilda Banken (100 percent ownership since end-2000),Sampo Bank to the Finnish Sampo bank (starting date 2000), Credit Bank to the Bank of Moscow (majority stake in 2005) and SBM Bank to the Greek-owned Marfin

The only two domestic credit institutions that have survived are Balti Investeeringute Grupi Pank AS (BIG) that was opened in 1992 and Tallinn Business Bank that started its operations already in 1991. A detailed overview of the banking sector history is available upon request to the author.

Besides the benefits that a foreign-owned entity can bring to the host country, there are also concerns that arise due to a substantial control position. It has been argued that foreign ownership reduces the host country’s control over the management and planning functions, leaving the managerial decision in the hands of foreigners, especially in cases where the majority of the banking assets have become foreign-owned. The reduction of information can affect the host country financial markets of small economies such as Estonia in a very severe way. In addition, the greater exposure to external shocks through the home country’s integration to world markets makes the financial sector of the host country very vulnerable. If all these assumptions were true then the 99 percent foreign control of Estonia’s banking sector should be a potential source of concern. If, however, my findings show that foreign banks have proven to be financially stable and have shown outstanding performance during the years, then such fears should be discarded and support the opposite that substantial foreign ownership of the banking sector is a positive development for small economies such as Estonia.

The main contributions of this paper are as follows. First, it adds to the lacking literature on financial sector foreign direct investment (FSFDI) by comparing small and large foreign-owned and domestic banks according to their financial strength and performance across time. Second, it focuses on Estonia and uses bank-specific data from the Estonian banking system within a new framework. Third, this study might have implications for promoting extreme foreign ownership of domestic sectors as opposed to conventional knowledge that discourages such actions.

II. Literature Review

Literature on the banking side has been less extensive than on the manufacturing side. The major areas of interest have been the expected benefits and drawbacks of foreign entry, the efficiency of banks, the differences in lending patterns, the concentration of home country banks and the determinants of banking FDI.

The main benefits have been examined extensively in the literature by Bonin et al. (1998). Some of these include:
  • The introduction of new banking technology and financial innovation since foreign banks can relatively easily introduce new products and services to the host country market
  • The potential for economies of scale and greater scope
  • Improvement of the conditions for competition
  • Development of financial markets
Improvements of the financial system infrastructure
- The attraction of foreign direct investment by diversifying the basis of capital and lending

The main arguments against these mentioned above have been most thoroughly approached by Buch et al. (2003):
- Fear of foreign control over the allocation of credit leading to substantial economic power over the economy
- Banking is an infant and special industry that needs exclusive protection since the sector has a central role in the economy
- Foreign banks may have different objectives such as taking care of the home country’s interests in the host country
- Regulatory differences. The fact that a weak regulatory control in the host or even home country can lead to an unsound banking sector.

The literature that deals with FSFDI and economic growth focuses on either lending patterns of domestic and foreign-owned banks or credit allocation efficiency. Both are positive factors that contribute to economic growth. Dages, Goldberg and Kinney (2000) in their study on Argentinean and Mexican banks found that although both domestic and foreign-owned banks experienced similar lending patterns in the 1990s, foreign banks showed stronger and less volatile loan growth. As to credit allocation, foreign banks tend to be regulated better and have thus the ability to discipline the host country’s fiscal and monetary policy discrepancies, leading to more efficient credit allocation. See LaPorta, Lopez-de-Silances, Shleifer (2002).

The problem of concentration among the home country banks can provide substantial risks to the host country’s economy. Theory suggests that heavy concentration amongst foreign owners can be potentially hazardous to the host country in case of an external shock. Detragiache and Gupta (2004), using Malaysian banks as an example, discovered that foreign banks with adequate international diversification played a stabilizing role during the Asian crisis, while the foreign banks with operations heavily concentrated in Asia behaved similarly to domestic banks.

The works of Goldberg et al. have focused mainly on foreign banks’ financial strength and performance, also referred to as soundness. Goldberg based her research on Latin American countries, starting with a broader analysis on a country-to-country basis using Moody’s Bank Financial Strength Ratings (BFSR) and then turning her lens to three countries for which she conducted an in-depth analysis of the banking sector. She chose Chilean, Columbian and Argentinean retail banks for the comparison analysis. Her findings were quite interesting. In her analysis she found that foreign-owned banks generally show a slight improvement relative to their domestic counterparts. The more in-depth results do not indicate striking differences in health, however, foreign banks tend to have higher average loan growth, higher average provisioning expense, and greater loss-absorption capacity. Overall Goldberg et al. (2001) concluded that foreign ownership may provide important positive influences on the stability and development of emerging market banking systems (Crystal et al).
Taking the approach and the analysis of the above-mentioned paper by Goldberg et al. (2001) I will attempt to examine the same set of variables with regards to the Estonian banking system.

III. Methodology

In my paper I will borrow the method Goldberg et al. (2001) used to measure financial strength of three Latin American countries in a slightly modified version. The framework is called CAMEL analysis. It consists of individual assessments of five core aspects of a bank’s financial condition and performance: Capital Adequacy, Asset Quality, Management, Earnings and Liquidity. Within the actual framework individual components are evaluated on a rating scale and are then aggregated to arrive at a composite ranking of the institution. Due to the complexities of the real CAMEL I will not adopt the exact approach that Goldberg used, but instead assess the banks in the above mentioned five categories across time utilizing financial statement analysis for credit institutions and Basel

Regulatory literature

Capital adequacy determines how well financial institutions can cope with shocks to their balance sheets. While a general capital over assets ratio is important to start with, it is more important to track capital adequacy ratios (CARs) that take into account the most important financial risks – foreign exchange, credit, and interest rate risk by assigning risk weightings to the institution’s assets. This is also the key parameter for financial managers and supervisory authorities to maintain adequate levels of capitalization. In general it is a measure of the amount of a bank’s capital expressed as a percentage of its risk weighted assets.

Two types of capital are measured in calculating CARs. Core capital (also called basic equity or Tier I) is that which is readily available and can absorb losses without a bank being required to cease trading (e.g. ordinary share capital). Supplementary capital or Tier II capital is that which generally absorbs losses only in the event of a winding-up of a bank, meaning it provides a lower level of protection for depositors and other creditors (e.g. undisclosed reserves or general loan/loss reserves). Tier II capital is also divided into upper and lower tier capital. The upper capital has no fixed maturity, while the lower one has a limited life span, which makes it less effective. The Basel Capital Accord also establishes Tier III capital for short-term subordinated debt, but since there is not requirement to hold capital against market risk this becomes less important (Reserve Bank of New Zealand).

The Basel Committee of Banking Supervision (BCBS) of the Bank of International Settlements (BIS) has a mandatory minimum CAR of 8 percent with Tier 1 capital to total risk weighted credit exposures to be no less than 4 percent. According to the Decree of the Bank of Estonia of 24 April 1997 the required credit institutions’ capital adequacy ratio is 10 percent since 1 October 1997.

Asset quality has a direct impact on the financial performance of a financial entity. The risk management system of an institution plays a significant role in determining the qual-
ity of assets such as loans and investments. While loan assets depend more on the realizable value of collateral, investment assets depend on the market value. The deteriorating value of assets directly pours into other areas (write-offs against capital), which puts the earning capacity of a bank in great danger. In addition, the increasing volume of business activity and the continuing development of technology that renders the environment more complex calls for constant evaluation and revaluation of credit risk management. It is therefore crucial to monitor the exposure to specific risks, trends in nonperforming loans, and the health and profitability of bank borrowers. In quantitative analysis four ratios are used: loan loss provision ratio, portfolio in arrears, loan loss ratio and reserve ratio. In my study I will be using two modified versions of the ratios.

The performance of bank indicators depends heavily on management’s vision, capability and soundness. It goes without saying that a stable management is the key to success, spurring confidence in investors, regulatory institutions and the general public. Although it is primarily a qualitative factor that can be applied to individual banks, there are several indicators such as efficiency measures that serve as indicators for management soundness. Operating Costs over Total Amounts Disbursed (Total Expense) and Total Expenditure over Total Income are the most commonly used numerical measures of management analysis. The first one indicates efficiency in disbursing loans in monetary terms: the lower the cost, the more efficient the bank. Meanwhile, the second one is a more general sign of bank performance affected by management’s decisions.

The role of earnings performance quality and trends can hardly be overemphasized. These measures reflect the institution’s ability to support present and future operations. Furthermore, a bank must earn adequate profit to support asset growth, build up reserves, finance its expansions, absorb losses and enhance shareholders’ value. Good earnings performance, just like sound management, is a positive signal to depositors, investors, creditors and the general public.

Liquidity within the banking sector is undoubtedly an important concept. A bank must be liquid to meet the depositors’ and creditors’ demand to maintain public confidence and prevent bank-runs. In order to obtain an adequate liquidity position, a financial institution must effectively manage its asset and liability systems to minimize the cost of converting its assets or increasing its liabilities when obtaining sufficient funds in case of need. Since liquidity has an inverse relationship with profitability, finding a balance between the two is a challenge for the managers. Even an initially solvent institution may be driven to closure by poor management of short-term liquidity risk.

As of 30 June 2006 there were seven licensed credit institutions operating in Estonia – Hansapank, SEB Eesti Yhispank, Sampo Bank, SBM Bank, Business Bank, Credit Bank and BIG Credit Bank. In addition there were two affiliated branches of foreign credit institutions that have captured a noteworthy share of the Estonian financial market – Nordea Bank from Finland and Parex Bank from Latvia.

Hansapank began its operations in 1992. Over the years it merged with and acquired small local banks in Estonia as well as in Latvia and Lithuania. In 1998 the Swedish-owned ForeningSparbanken acquired 50 percent of the stock, finalizing its 100 percent

Eesti Yhispank was created in 1992. During the next couple of years the bank merged with two big local banks – the North-Estonian Bank and the Tallinn Bank. In 1998 the Swedish banking group Skandinaviska Enskilda Banken (SEB) became the strategic partner of Eesti Yhispank. In 2000 SEB acquired a considerable stake in Eesti Yhispank, leading to a full buyout in 2001. Since April 2005 the name has been changed to SEB Eesti Yhispank.

Sampo Bank, belonging to the biggest Finnish banking group, was introduced to the Estonian market in 2000 when it acquired Optiva Bank from the Bank of Estonia. Nordea Bank opened its first branch office in Estonia under the name of Merita Bank in 1995. In 2002 the bank’s business name became Nordea Bank Finland Plc Estonia Branch, which belongs to the biggest financial group in the Nordic Region.

Credit Bank began its operations in 1992. This small local bank was able to survive the stress years and continue its existence. In 2005, 59.8 percent of its stock was acquired by Latvian Business Bank (LBB), a bank fully-owned by the Bank of Moscow, adding it to the list of foreign-owned banks.

Parex Bank is a subsidiary of the Latvian Parex Banking Group. It opened its representative office in Estonia in 2001 and a full branch in 2004.

Balti Investeeringute Grupi Pank AS (BIG) was founded in 1992 but received its credit institution activity license in 2005. Throughout the years it has expanded its operations both domestically and across borders to the other Baltic countries. It is the largest private credit institution in the Baltics after the universal banks and leasing companies.

Tallinn Business Bank, now part of Tallinn Business Bank Group, was founded in December of 1991 and it is one of the oldest commercial banks operating in Estonia. SBM Bank (Scandinavian-Baltic-Mediterranean Bank) was established in 1999 under the name Preatoni Bank. The current name was adopted in 2004 when Swedish and Greek investors acquired a majority stake in the bank. In 2006 Marfin Financial Group Holdings from Greece acquired a majority stake in the bank (50.12 percent).

For the purposes of this study I will exclude three of the aforementioned credit institutions from my analysis. First, Nordea Bank and Parex Bank will be excluded because they only have affiliated branches operating in Estonia. Second, I will exclude BIG from the analysis since the bank was granted its legal operating license in 2005, making one year of audited data insignificant for this research.

IV. CAMEL Analysis

Capital Adequacy

All of the banks maintained the Basel Capital Accord minimum ratio of 8 percent and the Bank of Estonia minimum requirement of 10 percent throughout the six years. Hansapank with the greatest market share experienced a very healthy growth in its CAR from 2000-2003, rising from 13.5 percent to 20.38 percent. The ratio declined slightly in 2004 to 18.62 percent and then took a huge drop in 2005 reaching 11.14 percent. During
all five years the bank increased its capital considerably through profits retained and profits for the current accounting period, despite the fact that its paid-in share capital and equity, the general banking reserve and other reserves remained the same.

The banks didn’t show a consistency in average CARs for the six years. Four of the banks had an average in the range of 12-17 percent, SBM Bank stood out with an extraordinary 100 percent average and Tallinn Business Bank fell in the middle with a 47 percent average. It is positive that all the banks, with the exception of SBM Bank, experienced very steady movements in the CARs. Stability in one of the most important standard indicators speaks well of the fact that all credit institutions in Estonia had adequate capital to cope with internal and external shocks to their balance sheets.

**Asset Quality**

All of the banks under analysis experienced a healthy growth in total and average loan portfolios, with some minor exceptions. Overall it is self-evident that banks with a greater capital base and a more experienced and knowledgeable parent company can achieve greater credit risk management. On average Hansapank had the lowest allowance to average loan ratio of 0.8 percent with SEB Yhispak following with 1.43 percent. As to the NPL to average loan ratio, Hansapank once again did better with 0.5 percent outperforming SEB Yhispak that had a considerably higher ratio of 2 percent. Sampo Bank had the third lowest ratio in both categories, 2.63 percent in allowance and 2.1 percent in NPL. Credit Bank and Tallinn Business Bank exchanged places after Sampo Bank in both categories. Credit Bank performed better in the NPL ratio with 5.98 percent (Tallinn Business Bank 8.9 percent) and Tallinn Business Bank did slightly better in allowance to average loan ratio with 4.18 percent compared to 4.9 percent. The reason why I excluded SBM Bank from the comparison analysis is simple – on absolute terms the bank would outperform all others with 0.58 percent in allowance ratio and 0.13 percent in NPL ratio. The fact that SBM Bank’s main areas of operations are investment banking and private banking and as a very small niche bank it only focuses on certain customer’s segment and areas of operation, can provide some insights to these staggering numbers.

**Management Quality**

In the expense to total income ratio the clear and stable leader was Hansapank with the lowest average ratio of 57 percent. All the other banks, with the exception of SBM Bank, fall into the 80 percent range. Hansapank was able to maintain its market share in the banking and finance industry due to economies of scale, generation of income, and steady expansion over the years, keeping the cost of unit of money lent at 57 percent on average. Foreign ownership also aided SEB Yhispak, Credit Bank and Sampo Bank, with more of a time lag however, to improve their ratios. Nevertheless, the same can’t be said about SBM Bank. The high ratios as demonstrated above can only lead to the conclusion that management is not undertaking sound practices. Since the acquisition in 2004 expenses grew disproportionately to income, exceeding income by 32 percent in 2005. The only domestic bank under analysis, Tallinn Business Bank, on average did very well in the management quality assessment. It should be noted that as of 2005 the bank’s ratios were down to 49 percent operating to expenditure and 64 percent expenditure to income. With only a
0.4 percent market share (as of June 2006) Tallinn Business Bank was able to remain competitive and produce one of the strongest ratios of all.

**Earnings Performance**

Surprisingly the most stable upward movements in the ROA were those of Sampo Bank and Credit Bank. SEB Yhisperskan would most likely be in this category if it weren’t for the changes in accounting methods that reduced net income. Hansapank experienced a steady move downwards stabilizing in 2004 at 2.44 percent. Both small niche banks, SBM and Tallinn Business Bank, displayed great fluctuations in their ROAs, suggesting very unstable earnings performance over the years. The situation for Tallinn Business Bank improved in 2005 but worsened greatly for SBM Bank. We have yet to discover where they are headed in 2006 as the competition keeps intensifying.

The highest average NIMs were those of SBM Bank that would regularly suggest very good performance. But due to the concerns addressed above this is not a good metric. NIMs for the three biggest banks, Hansapank, SEB Yhisperskan and Sampo Bank declined after 2003 due to competition in the market. Hansapank experienced the most stable decline as expected. There was no consistency in the NIMs of SBM Bank and Tallinn Business Bank suggesting again that jumping income from year to year does not speak positively of stable and healthy earnings performance.

**Liquidity**

On average the three major banks, Hansapank, SEB Yhisperskan and Sampo all experienced the same loan to deposit ratio, relying more on loan than deposit growth. SBM Bank was the more extreme case of 142 percent in the same credit growth category. Credit Bank and Tallinn Business Bank, on the contrary, were funded by deposits.

The liquidity position across banks does not show any consistency. The average ratios vary from 10 percent to 51 percent. All of the banks, excluding Credit Bank, took advantage of the new Bank of Estonia regulation in 2001 that helped to considerably lower their ratios. Although all the banks met the minimum mandatory requirements, the extra liquidity holdings with other credit institutions and cash at hand show very different patterns. One would expect that larger banks would have lower liquidity asset ratios since they always have the ability to borrow from the parent bank. This, however, is not what the numbers are saying. Although Tallinn Business Bank’s results are consistent with this theory (the bank had the highest liquidity ratio of 51 percent on average), Hansapank on the other hand tops the low ratios of SEB Yhisperskan, Sampo Bank and Credit Bank.

**V. Concluding Remarks**

My analysis shows that both small and large foreign-owned and domestic banks do not exhibit considerable differences in stability and performance over the sample period, although some differences do occur.

The best performance on average in all five categories can be attributed to Hansapank, the bank with the largest market share and parent institution to provide support, knowledge and resources. Despite the CAMEL rankings addressed above, following in the footsteps...
of Hansapank was SEB Yhispank that shares many of the same characteristics. The cases for the other banks were more ambiguous, some small banks stood out in management quality category, others in liquidity cushioning. The performance of SBM Bank should be considered an oddball since although characterized as a commercial bank engaged in lending activities, it mainly deals with specific customers in its private banking sector and focuses on certain segments of the business. However, based on the results the theory holds that financial sector foreign direct investment brings greater stability and improved efficiency to the domestic market.

Although I stated earlier that on average the small banks, either domestic or recently acquired, showed strengths in many categories, it is important to note that trends across time did not provide the same results in many cases. Smaller banks experienced greater fluctuations over time and were more prone to suffer from market risk. Big foreign banks were able to recover quickly and keep most ratios within a small range. Nevertheless, the fact that small banks with only a small market share have been able to survive profitably in the highly competitive market is a sign by itself.

The economic conditions of Estonia were very favorable for the financial sector throughout the six years. Increasing domestic demand, increasing real incomes and demand for housing boosted the banks’ assets through consumer and mortgage loans. Due to low world-wide interest rates, the competition in the loan market kept intensifying, driving down margins and increasing the range and quality of financial products. The pension acts in 2001 and 2002 also helped to improve savings growth in all of the six commercial banks. When looking at the overall numbers, the growth in loans and deposits across all banks was outstanding. Accompanied by accelerating GDP growth one should ask the question how long can this trend continue in a sustainable manner? The astounding net income numbers in 2005 should perhaps be a source of concern. As the growth in household borrowing exceeds that of their income, a slight downturn in the world markets could have serious repercussions on the Estonian economy as people start defaulting on their loans. This, however, is not the central topic of my paper. Yet, it is a result of the increasing openness of welcoming FSFDI in the earlier years and most definitely a crucial topic for further research.

Taken together, my findings show that overall small domestic and foreign-owned banks do not fall behind in financial stability performance of their big market share and foreign-owned counterparts and that extensive foreign ownership of the financial sector (99 percent) is not a threat for the economy. On the contrary, foreign-owned banks have improved their operations, providing ever more progressive ratios in many categories of the CAMEL framework.
References


IMPACT OF INNOVATIVE MICRO-CREDIT POLICY ON RURAL CREDIT MARKETS: THE CASE OF THAILAND’S MILLION BAHT VILLAGE FUND

Yana Morgulis
University of Chicago

ABSTRACT

The aim of this paper is to understand the impact of micro-credit on aspects of the credit markets in the context of an informationally innovative micro-credit program. In 2001, the Thai government implemented Million Baht Village Fund on a nationwide scale. The fund is the largest government sponsored micro-credit venture yet, and its structure is defined by characteristics prominent in formal and informal credit markets. Similar to the formal credit sector, the fund offers low interest loans. Yet, like the informal sector, its operations are informationally-advantaged through community participation and its credit is collateral-free. As a result of the micro-credit injection, formal credit in default increased significantly, while the size of the informal credit market and informal interest rates exhibit little substantive change. Though, later two measures show evidence of composition effects. This analysis contributes to the evaluation of the program, in particular, and to identification of credit injection tied changes in credit markets, at large.

I. Introduction

Rural credit markets in developing countries are characterized by a puzzling range of interest rates, which can be as low as 6 percent and as high as 80 percent per year. In such markets, lender types are associated with certain interest rates. Government regulated formal credit sources typically provide low-interest loans; these institutions include commercial banks, government organizations, and agricultural cooperatives. Informal sources tend to charge usuriously high-interest rates and consist of friends, relatives, and money lenders. Thus, formal and informal lenders exist side by side, but offer exceedingly unequal terms. Select literature has suggested that credit injections that incorporated inno-
vative methods to affect credit market information asymmetry may diminish market inefficiency (Bell; Besley; Siamwalla et al.). In this paper, I examine the impact of government infused, information-rich credit disbursed by the Thai Million Baht Village Fund (MBVF) on the dynamics of rural credit markets.

This analysis is based on a model proposed by Kaboski and Townsend (2006) who conducted an earlier evaluation of MBVF. I examine a four-year effect of the program and expand upon results of Kaboski and Townsend specifically in rural credit markets. My findings closely follow conclusions of Kaboski and Townsend, showing that program impact on credit markets is sustained over a longer period. I show that the program affected an increase in formal lending beyond the size of its injection, providing no evidence for crowding out or substitution away from other sources. The size of formal credit in default increased substantially, but there was no significant effect on the size of the informal credit market. The injection did not impact informal interest rates significantly, though findings show evidence of a composite effect.

An interview-based case study of MBVF institutions and household credit decisions in two Thai villages informs empirical findings. The study illustrates the effect of participation as an information enriching approach. Findings show that local participation as well as local and informal rules outlining administrative procedures play a large role in household credit decision-making and program outcome. Evidence of active secondary credit markets suggests a potential increase in the use of informal credit and the reason for low incidence of default in the program. The case study depicts local program variability and augments the complexity of econometric results.

The traditional view of rural credit markets claimed that greater availability of formal credit diminished market inefficiency (Siamwalla 1993). The theory asserted that informal lenders functioned as monopolies, suggesting that enhanced competition between lenders would reduce informal interest rates and implying that increased opportunities in the formal credit sector decreased household reliance on the informal credit markets. Credit injection seemed to offer a comprehensive solution for credit market woes, rendering informal lending more competitive, decreasing interest rates, and diminishing rural household dependence on informal credit markets. However, Bell (1993) and Siamwalla et al. (1993) pointed to the failure of credit injection to affect informal credit markets. In fact, both studies demonstrated that government implemented credit policies had negligible impact—informal interest rates did not decrease to match prior expectations, and informal credit markets continued to play a prominent role in rural lending.

Consequently, extensive literature identified the root of credit market inefficiency with information asymmetry (Bell; Siamwalla et al.; Hoff and Stiglitz; Besley). Siamwalla et al. (1993) argued that the informal sector was, in fact, competitive and that high interest rates reflected high information costs facing lenders in rural credit markets. Unlike the informal credit sector which considered information asymmetry, formal lender operations did not address informational problems. Bell (1993) and Siamwalla et al. (1993) showed that a credit injection that disregards the informational base of informal market operations does not decrease credit market inefficiencies. Additionally, Bell (1993), Besley (1994),
and Siamwalla et al. (1993) posited that the incorporation of the informational advantage practiced by the informal sector into formal sector operations could help achieve efficiency in the credit markets. This study investigates the proposition advanced by Bell, Besley, and Siamwalla et al. by measuring the impact of an information-privileged micro-credit program on rural credit markets.

The MBVF was a nation-wide initiative established in Thailand in 2001. Every community or village that organized itself into a committee and fulfilled membership requirements became eligible to form the local MBVF and to receive a government-sponsored grant. The program offered micro-credit at low-interest rates characteristic and matched informational advantage of the informal sector. MBVF involved community residents in its design, leadership, and lending, thus introducing local knowledge into MBVF’s administrative structure. The program represented one of the largest-scale government-sponsored microfinance initiatives of its kind (Townsend 2006). Applied nationwide, MBVF expenditure was equivalent to 1.5 percent of Thailand’s 2001 Gross Domestic Product (GDP).

This study contributes to the evaluation of MBVF as well as to the study of micro-credit in two substantial ways. First, despite the program’s nationwide application and its potential to achieve specific development goals, MBVF was implemented without an evaluatory structure (Chandoevwit 2003). This paper addresses the lack of evaluatory procedures, one of the key weaknesses of MBVF structure, by studying the impact of the program in local financial markets. Second, unlike many studies that analyze the effect of micro-credit initiatives, this study is based on data that exhibit a degree of exogeneity. Each community that applied to participate in the nationwide program received one million baht (US $22,300 in 2001) irrespective of size. Therefore, a resident in a village with a smaller population received a proportionally larger benefit from the program than a resident of a bigger village. Unsurprisingly, the size of household loans was significantly negatively correlated with village size. This study provides a unique opportunity to examine the effect of an innovative approach to micro-credit.

II. Data

This paper utilizes the Townsend Thai Project panel data to evaluate the effects of the program on rural financial markets. The analysis is based predominantly on the household component of the data. The panel dataset spans nine years and includes data for five years (1997-2001) prior to and four years (2002-2005) post MBVF program implementation.

The dataset is a collection of yearly rural surveys conducted in 64 villages across four Thai provinces. In each village, the study selected fifteen households. Each selection below the province level was made at random. Overall, 960 (4 provinces x 4 tambon x 4 villages x 15 households) households originally participated in the survey (Paulson and Townsend). The average household yearly drop out rate was 2.3 percent, such that of the 960 originally surveyed households, 794 participated in the survey for the entire nine year period. The household survey is a stratified, clustered, random sample panel data and is
This paper contributes an in depth analysis of the impact of short term credit on household credit market characteristics. MBVF loans have a government enforced maximum term of 12 months, and according to household surveys, the majority of formal and informal credit exhibit similar term limits. For example, in the 2003 survey year, 1060 of 1237 new household loans had a term length of twelve months or less. The method safeguards against the overwhelming effects of large, long-term loans on the credit market that could drastically diminish, if not altogether remove, the effect of MBVF credit. Ray (1998 p.351) and others (Bell; Siamwalla et al.) use four measures to describe credit markets – formal credit, informal credit, interest rates, and default. I utilize these four measures to study rural credit markets in Thailand. Each variable incorporates aggregate pertinent household data in the designated year. Interest rate is measured as average interest rate weighted by loan size. Formal credit, informal credit, and credit in default are measured in baht.

III. Model

The following model tests the impact of MBVF credit injection on credit market measures (Kaboski and Townsend).

\[ y_{nt} = \sum_{i} \alpha_i X_{i,nt} + \beta_1 \text{MBVF}_{nt} + u_{nt} \]

\( y \) measures a household specific credit market variable, \( n \) denotes the household between 1 and 794, \( t \) indicates the year, between 1997 and 2005 for a total of nine survey periods, and \( i \) refers to each of the control variables. Control variables for each household include number of adult males, number of adult females, number of kids, a dummy for male head of household, age of household head, age of household head squared, years of schooling of household head, gross assets, gross assets squared, income, and inverse of household number in the village. The latter control variable isolates the effect of village size over the full nine years.

Measures of informal credit and credit in default necessitate the use of a lagged MBVF variable. Informal credit coefficient describes ex post household decision to borrow from the informal sector. Siamwalla pointed to a prominence of secondary credit markets, informal sources of credit used to repay other debt obligations (Siamwalla 1993). The coefficient is also the result of a substitution between MBVF and informal credit lending. We utilize lagged MBVF to measure the resulting increase in secondary market presence. We believe that the effect of MBVF on credit in default also becomes visible after one survey cycle. In the regression, we want to isolate the impact of MBVF loans on consequent household decision making to borrow from the informal sector or default in the formal sector by lagging the VFRCR variable.

A two-stage, instrumental variables, fixed-effects estimation approximates the model.
In order to isolate the unadulterated effect of credit injection, we instrument the MBVF variable that measures the size of MBVF loans in household n. \( \theta_t \) and \( \theta_n \) represent time and household-specific fixed effects, respectively.

\[
MBVF_{n,t} = \sum_{i=1} \delta_{i}X_{i,n,t} + \theta_t + \theta_n + \lambda_1 \text{invHH}_{t,n}^*_{t=2002} + \lambda_2 \text{invHH}_{t,n}^*_{t=2004} + \lambda_3 \text{invHH}_{t,n}^*_{t=2005} + \epsilon_{n,t}
\]

Consistency sufficient assumptions are reproduced below.

Error Specification
\[
\epsilon_{n,t} = \varphi_t + \varphi_n + \epsilon_{n,t}
\]

Orthogonality Assumption
\[
\epsilon_{n,t} \perp \text{invHH}_{t,n}^*_{t=2002} \\
\epsilon_{n,t} \perp \text{invHH}_{t,n}^*_{t=2003} \\
\epsilon_{n,t} \perp \text{invHH}_{t,n}^*_{t=2004} \\
\epsilon_{n,t} \perp \text{invHH}_{t,n}^*_{t=2005}
\]

The instrument is an interaction term between the inverse number of households in the village in year t and year dummy variables (invHH\(_{t,n}^*_{t=2002}\)) on the duration of the program between 2002 and 2005. Government grant to each MBVF and, thus, to each village was fixed at one million baht. Therefore, a resident of a larger village receives proportionally less funding than a villager belonging to a smaller village. The instrumental variables model attempts to isolate the effect of the grant on household (n) MBVF loan. Table 1 compares the value of generated instrumental variables with average MBVF loans in each year of the program. Coefficients of instrumental variables are highly significant and describe a majority of instrumented variables as demonstrated with averages in Table 1. The inverse number of households in the village (invHH\(_{t,n}\)) in each year of the panel data set (1997-2005) is included in the control variables of the model in order to control for population-based variation in the villages that is independent of the MBVF effect.

**IV. Results**

Results show the effect of a change in MBVF credit on formal credit, credit in default, informal credit, and interest rate measures. Quantities are coefficients of the MBVF explanatory and instrumented variable. Tables 2, 3, and 4 are based on monetary measurements. The coefficients in these tables explicate the effect in baht of a one baht increase in the household MBVF loan. Table 5 illustrates the impact of a change in a household MBVF loan of 10,000 baht on interest rates in percentage points. Due to sensitivity of regression results to outliers, each regression features three results. The first regression considers the full array of data, the second discards the top and bottom one percent of non-zero dependent variable data, and the third regression removes five percent of non-zero dependent variable data at both extremities.

Credit in default values in Table 3 are measures ex post. The impact on total credit in default attains significance and large scale after removing the first set of outliers. The
result for total credit says that for every baht increase in MBVF loan, a household declares an additional 1.019 to 0.728 baht default. Default in the formal credit sector explains most of the effect on total credit and displays similar empirical patterns. Results for informal and MBVF credit in default are not significant but remain positive. The finding that risk factor increases with increases in credit (Siamwalla 1993) impacted the formal credit sector, in particular.

Table 4 illustrates the effect of MBVF credit on informal credit, also an ex post measure, to be generally positive and insignificant. The data demonstrate that the increase of one baht in MBVF credit, increases informal credit by 0.174 to 0.102 baht in the next period. Informal credit can be further separated by origin with moneylenders or neighbors and relatives. In the case of moneylenders, the effect remains insignificant, but turns negative from 0.108 for the full set of data to -0.015 baht with the removal of outliers. Coefficients for informal credit from relatives and neighbors remains steady in the environs of .07 baht as standard error decreased.

Lastly, Table 5 demonstrates the effect on interest rates as a result of a 10,000 baht increase of an MBVF loan. The first regression says that for every 10,000 baht increase in MBVF loans, aggregate interest rate increases by 2 percentage points. However, the top one percent of interest rate data (interest rates that exceed 40 percent annually) drives the result as the effect decreases and becomes insignificant in subsequent regressions. Interest rate data can be divided into formal and informal categories. The impact of MBVF credit on formal interest rates is small, decreasing further and changing signs as outliers are removed. Interest rates from the informal sector remain insignificant, but large and positive throughout the analysis. As outliers are removed, the effect shifts from 4.15 to 7.36 percentage points representing a substantial shift with average informal interest rate at 16.4 percent. A further breakdown of informal interest rates between moneylenders and relatives or neighbors shows evidence for composite effects. Both effects are large and opposite in sign, and, in the case of interest rates from relatives or neighbors, even significant for the full data regression. An additional 10,000 baht of each of the two variables exhibits a large effect, yet due to opposite signs, the sheer size becomes obscured as categories are merged.

As a result of MBVF credit injection, total credit increased significantly, predominantly due to increases in formal credit. Similarly, the substantial increase in ex post credit in default can be attributed to the formal sector. Ex post effect on informal credit was not significant, but generally positive suggesting that an increase in MBVF borrowing affected a smaller positive change in the informal credit market. Impact on interest rates when distributed between formal and informal sectors evidenced opposing effects. Interest rates on loans from moneylenders decreased as contrasted with loans from neighbors or relatives, which increased as a result of a positive change in MBVF borrowing.

The impact on credit market measures illustrates that not only do increases in credit and default follow from credit injections, but each effect as demonstrated with informal credit and interest rates may be the result of composite effect (Bell 1993). Unsurprisingly, credit influx also increased the ex post default as noted by Ray (p.365 1998).
Siamwalla (1993). The fact that most of the default increase was in the formal credit market may be the result of information privileging and consequent enforcement advantage in the informal and MBVF credit sectors.

V. Case Study

To illuminate factors that shaped empirical findings, I conducted a case study that consisted of two Thai villages. This study examined the place of MBVF’s informationally-rich participatory approach in program impact on the credit markets. In fact, local participation was a prerequisite of program institutionalization. National policy required that community members organize a program committee and membership base in order to receive the MBVF government transfer. Thereafter, the committee was responsible for creating program policy based on national guidelines and enforcing rules in the lending process. Local participation introduced local knowledge into the policy-making and lending processes of the program.

The study is based in Thailand’s Chachoengsao Province. This province is one of the four provinces sampled in the Townsend Data Project from which data for the empirical study is derived. The province’s location in the Central region of Thailand allows it to benefit economically from a spatial proximity to the capital of Thailand. Two villages used in the study were located close to each other, nevertheless, MBVFs in the villages differed substantially in their policies. Significant differences in program structure suggested the presence of more prominent policy variability between more distant MBVFs and consequent adjustment to local preferences. This observation followed with the Thai government’s limited policy guidance that encouraged policy adjustment to local conditions (“National Procedure for Village and Urban Food Committee: Formation and Administration of Village and Urban Fund”).

The policies adopted by local institutions may have induced increased volume in the informal credit sector. Both institutions lent for the maximum, nationally dictated loan term of 12 months and required lump sum repayment. Such focus may have contributed to greater interaction between MBVF borrowers and the informal credit sector, which borrowers accessed to aid repayment. Leadership of both institutions encouraged use of loans for entrepreneurship, yet, despite, local information on use of borrowed funds, they did not enforce their preference. This loose policy led to a less financially productive personal investment of MBVF funds, highlighting another factor that strengthened formal and informal intermediation with regard to loan repayment. Locally placed pressure to repay and prominence of the secondary credit market, a credit sector used as means to repay, in this case, MBVF loans, removed default as a tenable category in the context of MBVF. Each interviewee noted a dearth of default cases. Numerous aspects of the program may have strengthened the ties between MBVF and informal credit markets.

On the other hand, local participation in administrating the program created a knowledge base which may have increased competition in the informal credit market. Informational basis for MBVF lending paralleled lending methods in the informal credit
sector. As a result, better loan conditions offered by MBVF allowed for competition for borrowers between the two credit markets. As a result, not only could borrowers use MBVF credit to substitute away from informal credit markets, but MBVF markets selected credit worthy borrowers, raising the average informal market borrower’s risk rating. The study demonstrates that the program resulted in a wide ranging set of, at times, opposing effects, especially in its borrowers’ interactions with the informal credit sector.

This study brings to the forefront a multiplicity of effects in the rural credit markets predominantly arising from local policy-making and informational advantage. Here, I have begun to demonstrate that various aspects of program design, specifically, structure of lending cycle, focus of lending, accumulation of local information, and definition of default that affect program impact in intricate and complex ways. Details of program design offer a more nuanced picture of the program and provide insight into empirically deduced results.

VI. Conclusion

This empirical study points to a complex financial market response that resulted from an injection of locally-administrated formal credit. MBVF credit injection triggered a change in rationed formal credit markets and increased formal lending in rural environments. The expansion of credit was accompanied by an increase in default ex post, predominantly in the formal credit sector. Ex post impact on informal credit was weakly positive, while contemporaneous impact on interest rates was largely insignificant. The effect on interest rates exhibited dual characteristics: on the one hand, impact on interest rates derived from credit from moneylender was negative, and, on the other hand, impact on interest rates on credit from relatives or neighbors was positive. The study demonstrated a range of significant effects and negating composite influences.

Although MBVF’s design addressed Siamwalla’s (1993) proposition for an innovative method to address informational asymmetry in credit markets, empirical findings on informal credit markets agreed with the results of earlier micro-credit ventures. Informal interest rates and informal credit markets did not significantly decrease – inefficiencies in aggregate credit market measures remain prominent. The case study nuances these conclusions by revealing contrasting effects influencing subcategories of credit market measures.


### Table 1: Instrumental Variables Coefficients

<table>
<thead>
<tr>
<th>Year</th>
<th>Instrumental Variable</th>
<th>Average MBVF Loan</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>7,677</td>
<td>9,635</td>
</tr>
<tr>
<td>2003</td>
<td>5,945</td>
<td>9,245</td>
</tr>
<tr>
<td>2004</td>
<td>8,934</td>
<td>10,009</td>
</tr>
<tr>
<td>2005</td>
<td>6,347</td>
<td>10,198</td>
</tr>
</tbody>
</table>

### Table 2: Total and Formal Credit

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>VFCR Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td><strong>Total Loans</strong></td>
<td>1.931***</td>
</tr>
<tr>
<td></td>
<td>[0.199]</td>
</tr>
<tr>
<td><strong>Total Loans (no MBVF)</strong></td>
<td>0.931***</td>
</tr>
<tr>
<td></td>
<td>[0.199]</td>
</tr>
<tr>
<td><strong>Formal Loans</strong></td>
<td>0.575***</td>
</tr>
<tr>
<td></td>
<td>[0.141]</td>
</tr>
<tr>
<td><strong>Formal Loans from Agricultural Bank</strong></td>
<td>0.227***</td>
</tr>
<tr>
<td></td>
<td>[0.052]</td>
</tr>
<tr>
<td><strong>Formal Loans from BAAC</strong></td>
<td>0.221*</td>
</tr>
<tr>
<td></td>
<td>[0.126]</td>
</tr>
<tr>
<td><strong>Formal Loans from PCG</strong></td>
<td>0.034***</td>
</tr>
<tr>
<td></td>
<td>[0.013]</td>
</tr>
<tr>
<td><strong>Formal Loans from Commercial Bank</strong></td>
<td>0.099***</td>
</tr>
<tr>
<td></td>
<td>[0.028]</td>
</tr>
<tr>
<td><strong>Formal Loans from Rice Bank</strong></td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>[0.003]</td>
</tr>
</tbody>
</table>

(1): All Data
(2): 1% of extreme non-zero dependent variable observations excluded
(3): 5% of extreme non-zero dependent variable observations excluded
* significant at 10%; ** significant at 5%, *** significant at 1%
### Table 3: Credit in Default

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>VFCR Coefficient (Lagged)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Total Credit in Default</td>
<td>0.384</td>
<td>1.019***</td>
<td>0.728***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.406]</td>
<td>[0.286]</td>
<td>[0.192]</td>
<td></td>
</tr>
<tr>
<td>Formal Credit in Default</td>
<td>0.367</td>
<td>0.936***</td>
<td>0.692***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.305]</td>
<td>[0.169]</td>
<td>[0.128]</td>
<td></td>
</tr>
<tr>
<td>Informal Credit in Default</td>
<td>0.232</td>
<td>0.108</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.225]</td>
<td>[0.210]</td>
<td>[0.017]</td>
<td></td>
</tr>
<tr>
<td>MBVF Credit in Default</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.006]</td>
<td>[0.006]</td>
<td>[0.006]</td>
<td></td>
</tr>
</tbody>
</table>

### Table 4: Informal Credit

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>VFCR Coefficient (Lagged)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Informal Loans</td>
<td>0.174</td>
<td>0.043</td>
<td>0.102</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.207]</td>
<td>[0.156]</td>
<td>[0.084]</td>
<td></td>
</tr>
<tr>
<td>Informal Loans from Moneylenders</td>
<td>0.108</td>
<td>0.122</td>
<td>-0.015</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.184]</td>
<td>[0.162]</td>
<td>[0.084]</td>
<td></td>
</tr>
<tr>
<td>Informal Loans from Relatives and Neighbors</td>
<td>0.066</td>
<td>0.079</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.091]</td>
<td>[0.066]</td>
<td>[0.045]</td>
<td></td>
</tr>
</tbody>
</table>
### Table 5: Interest Rates

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>VFCR Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>0.023***</td>
</tr>
<tr>
<td></td>
<td>[0.00834]</td>
</tr>
<tr>
<td>Formal Interest Rate</td>
<td>0.00604</td>
</tr>
<tr>
<td></td>
<td>[0.00496]</td>
</tr>
<tr>
<td>Informal Interest Rate</td>
<td>0.0415</td>
</tr>
<tr>
<td></td>
<td>[0.0481]</td>
</tr>
<tr>
<td>Informal Interest Rate from Moneylenders</td>
<td>-0.103</td>
</tr>
<tr>
<td></td>
<td>[0.0812]</td>
</tr>
<tr>
<td>Informal Interest Rate from Relatives and Neighbors</td>
<td>0.120*</td>
</tr>
<tr>
<td></td>
<td>[0.0689]</td>
</tr>
</tbody>
</table>

Note: *** indicates statistical significance at the 1% level.
I. Introduction

Many argue that the divisions between the global “north” and “south” can be lessened through more liberalized trade, particularly in agricultural products, in which most developing nations concentrate production. However, agricultural tariffs were one of the prime sticking points in the recent “death” of the World Trade Organization (WTO) Doha trade round, and it does not seem that many developed countries are poised to make drastic changes in their agricultural protection regimes. While multilateral efforts have stalled, there has been a simultaneous influx of preferential trading agreements (PTAs), creating a “new world trade order” of hub and spoke agreements, often linking large developed economies with developing nations. Since the 1970’s, the European Union (hereafter referred to as the EC, as it is referred to by WTO documents) has signed over a dozen preferential trade agreements with developing nations, especially in North Africa and the Middle East.

But to call these agreements “free trade agreements” would be a misnomer since several agricultural tariffs, such as those embodied by the EC’s embattled Common Agricultural Policy (CAP), are still included in the agreement. Using the data from a particular EC North – South PTA, the PTA signed between Morocco and the EC in 2000, we have constructed a model based on that of Messerlin (2001) and Hufbauer and Elliot (1994) to estimate the net welfare effects of the EC’s agricultural protectionism on a specific Moroccan export, olive oil, both on EC consumers and Moroccan producers. Using this model, we replicated two different scenarios – shifting from MFN tariffs to the PTA, and then from the PTA to a free trade scenario between the EC and Morocco, concurrently performing a welfare analysis in both the domestic (EC) and import markets for Moroccan olive oil. While there were gains from instituting the PTA, much larger gains for both producers and consumers were obtained from completely free trade in olive oil.
Despite the rather small size of the gains, trade creation was found to occur in both shifts, as the welfare gains exceeded the loss of tariff revenue.

The null hypothesis is that the presence of the agricultural tariffs has no effect on either producer or consumer surplus, with the alternative being that there is an effect due to the presence of tariffs. Through this research model, we hope to expose the contradictory nature of these PTAs – not to completely devalue their good faith efforts, but to illustrate their shortcomings, hopefully suggesting paths towards future liberalization and greater welfare gains.

Following Peter Messerlin’s model of calculating the cost of protectionism in various sectors of the EC economy, my analysis will take his work an additional step forward by applying it to a specific EC PTA. We calculate the welfare effects of at first instituting the EC-M, and then shifting to a free trade scenario. The model applies a partial equilibrium model used by Hufbauer and Elliot in their study of US protectionism (Hufbauer and Elliot 1994). Concerning data, we use time series (1979 – 2005 for the domestic market analysis, 1962 – 2005 for the import market analysis) trade volume and unit price data from the U.S. Department of Agriculture (USDA) (for the domestic market) and UN Comtrade Database (for the import market), and EU tariff lines from the International Trade Administration. We then estimate the supply and demand curves in a partial equilibrium model under three price scenarios (Before EC-M, Under EC-M, and Free Trade), estimating the elasticity of domestic supply, import demand, and domestic demand. In addition, we construct estimates of the welfare triangles, as in Messerlin’s analysis, to illustrate the overall costs of protection (Messerlin 2001). The welfare analysis found that while a shift to the EC-M agreement resulted in a consumer surplus gain in both markets of 8,149.191, shifting to completely free trade produced a consumer surplus gain of 469,976.227. Changes in producer surplus also follow a similar pattern: instituting the EC-M results in an increase of .000319, but shifting to completely free trade increases producer surplus by .347315. Losses due to tariff revenue reduction hardly had an effect on the net gains made from increasing liberalization, as trade creation was evident through further reductions in tariffs.

II. Literature Review

The EC-M is not an isolated policy. Following a summit in Barcelona in 1995, the EC began negotiations with several Mediterranean nations regarding not only a possible economic free trade area, but regional security, human rights, and development aid agreements as well (Messerlin 2001). In the decade since the initiation of this dialogue between the EC and its Mediterranean neighbors, a sizeable literature has developed on the evaluation and assessment of these agreements. Many of these models have used a computable general equilibrium model (CGE), with Messerlin’s computable partial equilibrium (CPE) model being the exception. Much of the work also provides a qualitative analysis with regard to certain economic indicators, such as increased FDI flows or the level of standard harmonization.
Morocco and Tunisia were the only Mediterranean countries to have completed and signed EMPAs (in 1998 and 2000, respectively) until the recent agreements between the EC and Israel, the Palestinian Authority, Jordan, and Egypt (Messerlin 2001). Both the quantitative and qualitative literature focuses heavily on Morocco’s and Tunisia’s experiences due to the length of time that has elapsed since the agreements were signed. Prior to the signature of EC-M, Rutherford, Rutstrom, and Tarr (1993) completed a CGE study of over 39 sectors of the Moroccan economy that significantly influenced future literature on EMPAs. Gains from the EC-M were estimated at 1.5% of GDP, and 2.5% of GDP should Morocco liberalize trade with the rest of the world, illustrating that some degree of trade diversion resulted from the EMPA (Rutherford 1993). In addition, the study estimated the price distortion with regard to the agricultural protectionism in the agreement to be 8%, a figure cited throughout the subsequent literature (Rutherford 1993). The model evaluated several different cases (FTA, unilateral liberalization against the EC, increased market access for agriculture, etc.) with differing estimated elasticities of demand. Ultimately, the study concluded that if there was increased access for agricultural products, Moroccan welfare would increase by .25% of GDP (Rutherford 1993). Rutherford et al replicated this study in 2000, choosing to focus on an EMPA between the EC and a “representative Arab Mediterranean country.” Welfare increases were higher than in the Morocco-specific study, as the model predicted gains between 3-5% of GDP (Rutherford et al 2000).

Partial equilibrium models, focusing on a particular segment of an economy were not primarily used in the EMPA literature; however, Messerlin’s thorough analysis of EC protectionism across multiple sectors during the 1990s creates an accessible model that easily can be applied to such a narrow industry as Moroccan olive oil. In his book, *Measuring the Costs of Protectionism*, Messerlin calculates the overall level of protectionism in the EC to be 13-14% for the years 1990-1997, with protection in the agricultural sector averaging 34% over the same time period. His model calculates the “consumer-subsidy equivalent tariffs” (CSE-based tariff) for over 40 sectors in a partial equilibrium framework based on that used by Hufbauer and Elliot (1994). Specifically regarding the EMPAs, Messerlin concludes that the selective reduction in protectionism (such as in the agricultural sector) leads to increased effective rates of protection, and the complex and bureaucratic reputations of the Mediterranean governments provide little hope for dynamic effects.

### III. Empirical Specification

#### The Partial Equilibrium Model: Hypothesis

Using a computable partial equilibrium model developed by Hufbauer and Elliot (1994) (with applications to the specific EC case by Messerlin), we start by calculating the supply and demand conditions in both the import and domestic markets for olive oil, estimating the demand elasticity for domestic and imported olive oil, as well as the elasticity of supply for domestic olive oil, under three different price scenarios: Before the EMPA,
Under the EMPA, and Free Trade. Using the supply and demand schedules, we analyze the welfare effects of protectionism using the traditional cost-benefit analysis used in trade theory, looking at changes in consumer and producer surplus, as well as tariff revenue as further liberalization occurs. The null hypothesis is that there is no significant change in welfare effects between the different scenarios; the alternative being that the reduction, and eventual removal, of tariff barriers, causes a significant increase in consumer and foreign producer surplus.

**The Partial Equilibrium Model: Assumptions**

The model used in this analysis draws upon that of Messerlin (2001), and the Hufbauer and Elliot (1994) model, with slight modifications. In order to calculate the demand and supply schedules for the import and domestic markets, the following assumptions are enforced:

- Domestic goods and imports are perfect substitutes
- The import supply curve is perfectly elastic
- The domestic supply curve is upward sloping
- All markets are perfectly competitive
- Morocco will be considered a small open economy, and thus cannot affect world prices

The first assumption differs from that of Messerlin’s model; we have chosen to consider Moroccan olive oil as a perfect substitute for domestically produced olive oil. The EC, through its Common Market Organization (CMO) for olive oil, has set such rigorous quality standards, in addition to labeling requirements, that it would be difficult to calculate a significant difference between Moroccan olive oil and domestically produced olive oil, aside from differences due to the tariff weighted price (“Working Paper of the Directorate-General for Agriculture: The Olive Oil and Table Olives Sector”). Any difference in price (in a free trade scenario) would depend upon consumer preferences and the possibility of a home bias (as many olive-oil producing countries primarily consume their own products). While there is significant evidence that cultural and ethnic preferences may affect price, with the limitations in the length of this paper, it is difficult to consider complex tastes, and thus we assume that a home bias does not exist.

**The Partial Equilibrium Model: Data Specifications**

A significant challenge in developing this model has been the collection of accurate data that will accurately reflect the trade in olive oil between Morocco and the EC in order to extrapolate the costs of protectionism that result from the EC-M agreement. Much of the data needed to be transformed in order to make it acceptable for analysis, and we have made several assumptions, both of which are detailed in this section. While my data set is not extremely large, we were able to expand it quite late in the research in order to obtain significant results.

In order to replicate the import prices in the various scenarios, we used as a cost, insurance, and freight (CIF) unit price, calculated using information from the UN Comtrade database. Using data on EC member bilateral imports of olive oil from Morocco from 1962 onward, we calculated this price by dividing the value of total trade by the number
of units (originally in kilograms, but later converted to tons to match the tariff specification). This calculation yielded the CIF unit price per ton that was then scaled to reflect the various tariff scenarios. For each year, these prices varied across countries and in order to create a single price for each year, the various country prices were averaged, with those countries not importing from Morocco in that year removed from the data set. The prices were originally reported in US dollars; we converted them to “euros” using a combined Deutschmark/ECU/Euro exchange rate series (IMF). By using unit prices, some of the bias with regard to EC producer prices influenced by the CAP was avoided; however, the CIF unit price calculation was still somewhat of a crude measure of “world” prices. According to the International Olive Oil Council, European prices should be used as an index for world prices, so that even the value of the goods imported from other countries will reflect the CAP-supported price to some extent. In addition, the reduction of tariff barriers may also have an effect on Moroccan exporter prices which would not be captured using this type of analysis.

Calculating the tariff rates for the various scenarios was slightly more complex. The tariff rates are Consumer Subsidy Equivalent (CSE) based ad valorem tariff equivalents, calculated using the OECD method, so that \( t = T/Pm' \), where \( T \) is the tariff per ton and \( Pm' \) is the world price. According to the EC tariff schedule provided by the International Customs Tariff Bureau, the current MFN rate is €1,226 euros/ton, while a 10% reduction under EC-M would result in a tariff per ton of €1,103.4. Appendix IV calculates the CSE-based ad valorem tariff equivalents, which are then used to represent \( t \) and \( t' \) in the model. These ad valorem tariff equivalents were then applied to the world price in order to calculate the Before the EMPA price and the Under the EMPA price, where a 10% tariff reduction is taken into account.

Quantity demanded and imported also underwent a series of transformations and assumptions before being finalized in the model. For the domestic market analysis, data on EC olive oil consumption and quantity imported from 1979 to 2005 was extracted from the United States Department of Agriculture (USDA). In the import market analysis, we used the same bilateral trade flow data collected from the UN Comtrade database to structure the unit prices. Using data from 1962 to 2005 for each member state (inclusion in the data set began at the year of accession to the EC), total EC quantity imported from Morocco in kilograms was summated and then converted to tons. This data set particularly alleviated problems with the initial USDA olive oil imports data set that were due to double counting of imports between EC countries in early years of integration, as well as less accurate figures with regard to bilateral flows of trade between the EC and Morocco.

**The Partial Equilibrium Model: The Import Market**

The Import Market is defined such that the inelastic foreign supply curve is given by the prevailing price under the given scenario (Before EMPA, Under EMPA, or Free Trade), and the import demand curve is a function of both the domestic price and the import price.

The prices under the various scenarios are the following (where \( Pm' \) is the world price, \( t \) is the tariff rate before the EMPA and \( t' \) is the tariff rate under the EMPA):

- Price Before EMPA: \( Pm_B = Pm'(1+t) \)
• Price Under EMPA: \( P_{mEMPA} = P_m'(1+t') \)
• Free Trade Price: \( P_{mFT} = P_m' \)

Import demand is defined by the following equation:
\[
Q_m = \gamma P_m^{Emm}
\]

\( Emm \) is the elasticity of demand for imported olive oil and will be estimated as the parameter in the empirical analysis in order to discern the supply and demand schedules under the three different scenarios. However, the model will not test whether or not there is a significant change in either parameter between the different scenarios.

While the model is originally specified as a Cobb-Douglas demand function in terms of elasticity, it will be empirically tested in a logarithmic specification:
\[
\ln Q_m = \ln \gamma + Emm \ln P_m
\]

Using the estimates of \( \gamma \) and \( Emm \), we will be able to produce welfare analysis graphs for the import market. The initial liberation (decrease in \( t \) from \( t \) to \( t' \) under the EMPA agreement) will cause the price to fall from \( P_m' + t \) to \( P_m' + t' \), resulting in gains in consumer and producer surplus. Quantity imported increases from \( Q_m \) to \( Q_m' \). A consumer surplus gain accrues equal to the area \( abcd \). Similar calculations are performed for the movement from Under the EMPA to Free Trade.

This graph also includes additional foreign export supply curves to represent the border prices that foreign producers receive under the Before EMPA and Under EMPA scenarios. The foreign producer surplus gain from the lower tariff in the Under EMPA scenario is equal to \( eihg + \text{if} Q'MQM \cdot \text{ij}Q'MQ'M \). These welfare gains are especially important because they represent the gains to Moroccan producers from the increasing levels of liberalization present in the different scenarios. Losses in tariff revenue due to increasing liberalization can be quantified by \( \text{ij}Q'MQM \cdot \text{degh} \) when the agreement is instituted and \( \text{degh} \) when tariffs are completely liberalized. This loss occurs only in the EC, and is subtracted from the consumer and producer surplus gains in order to determine the net effect of liberalization.

**The Partial Equilibrium Model: The Domestic Market**

Within the domestic (EC) market for olive oil, the model is defined rather simply as a function of domestic demand and supply, which includes the supply of imports in addition to domestic suppliers. Because this is an equilibrium model, \( Q_d = Q_s \), which are both defined by the following equations:
\[
Q_s = \alpha P_m^{Es}
\]
\[
Q_d = \beta P_d^{Edm}
\]

The variables are the same as those included in the import market, with the exception of the parameter \( Es \) which represents the elasticity of domestic supply and \( Edm \), which represents the elasticity of demand. Like in the import market, the model will be estimated with a logarithmic specification and the elasticities will be estimated as the parameters of the model:
\[
\ln Q_d = \ln \alpha + Edm \ln P_m
\]
\[
\ln Q_s = \ln \beta + Es \ln P_d
\]

Calculating the estimates of \( Edm \) and \( Es \), we will be able to produce a welfare analy-
sis graph for the domestic market. Before the EMPA, the domestic price is \((Pm' + t)\) and domestic quantity demanded is equal to \(Qd\); liberalization to the lower tariff rate, \(t'\), results in a decline in the import price, which shifts the domestic demand curve inward, from \(Dd\) to \(Dd'\), which drops the quantity demanded to \(Qd'\) and the price falls to \((Pm' + t')\). A similar pattern occurs when the economy moves from the EMPA tariff schedule to a free trade scenario, as the tariff is removed completed and the domestic price equals the world price.

In the domestic market, there is a consumer surplus gain from lower prices equal to the area \(swyz\). This is exactly equal to the loss suffered by EC producers due to lower prices. Intuitively, this producer loss does not affect Moroccan producers because they are still receiving a higher border price in either the Under EMPA or Free Trade scenario, and thus only domestic producers face lower prices through the trade liberalization. Net welfare gains in the domestic market are equal to zero since the consumer gains and producer losses offset each other, and there is no tariff revenue loss to consider.

Once the elasticities of supply and demand were estimated, we then calculated the demand and supply functions for both the domestic and import market in order to assess the welfare gains from increasing liberalization (and conversely the costs of maintaining protectionism). After calculating the intersections, we developed the graphs included in Appendices VI and VII.

In performing the welfare analysis, we found that there were small gains had by increasing liberalization. Within the domestic market, the consumer surplus gain from instituting the EMPA totals 8,149.1905; however, completely liberalizing trade results in a consumer surplus gain of 46,976.13505. These gains in consumer surplus are canceled out by equal losses by domestic producers due to the price decrease. This results in a net welfare effect of zero for the domestic market.

Similar effects are found in the import market as the trade regime becomes more open – consumer surplus increases under the EMPA by .0002067, and .0921193 when trade becomes 100% free. In addition to the gains to consumers from increasing liberalization, producer surplus also rises from the increasing liberalization. Instituting the EMPA and receiving a 10% reduction in the tariff results in an average border price shift from \(€1,126.90\) per ton to \(€1,249.50\) per ton, and an increase in producer surplus of .0003186. However, the shift to free trade increases the average border price from \(€1,249.50\) per ton to \(€2,352.90\) per ton (the free trade/world price). The shift to free trade also results in a .3473153 increase in producer surplus. These increases in producer surplus are particularly significant because they illustrate the gains to Moroccan producers through increasing liberalization compared to the current tariff regime under the EMPA.

However, increasing liberalization does cause losses, specifically with regard to tariff revenue. Within the import market, instituting the EMPA will result in a loss of tariff revenue equal to .0000114; liberalizing the tariffs the remaining 90% results in a much larger tariff revenue loss of .16177. When either of these figures is subtracted from the consumer and producer surplus gains, the net welfare effect continues to be strongly positive: .0005139 when the EMPA is adopted and .27525 when free trade is initiated. Trade creation, originally defined by Viner as welfare gains exceeding net welfare losses on net,
occurs in this study – while the magnitude of the gains is small, the positive sign results prove that trade theory intuition is correct. The welfare analysis illustrates that any form of protectionism in the olive oil sector within the EC-M agreement results in negative welfare effects.

**IV. Policy Implications**

From the empirical analysis, the welfare analysis of the shifts from first instituting the EMPA and ultimately liberalizing to completely free trade in olive oil appear to follow the intuition behind trade theory – increasing liberalization will provide additional gains to consumers and foreign producers, while causing domestic producer losses. Much larger gains are obtained through shifting to free trade, however, rather than from instituting the EMPA. Also, the results appear to conform to my hypothesis – that the current EMPA between the EC and Morocco continues to distort trade in olive oil, and that a shift to completely free trade would result in benefits not only to EC consumers, but to Moroccan olive oil exporters.

Unfortunately, the time has passed to make a true policy argument against the agricultural protectionism found in the EC-M agreement. However, within the model there is the potential to calculate a “what-if” scenario to illustrate the gains that could have been actualized had the EC-M agreement eliminated all tariffs on olive oil. Performing the same welfare analysis as in Section IV, signing an EMPA agreement with completely free trade in olive oil would have resulted in a gain of .2596121. Interestingly, this is a smaller figure than the welfare gains that result from the shift from the EMPA to free trade. This is most likely due to the larger amount of tariff revenue lost in this shift from complete protection to free trade, while the gains due to increased flows and lower prices are not transmitted as strongly due to the lower elasticity of import demand. While these gains are foregone due to the actual nature of the EC-M, it is important to illustrate the costs of protectionism that result from signing an EMPA (or any preferential trading agreement for that matter) which includes less than full reduction in tariffs, even in only a few sectors.

Ultimately, trade creation does occur as tariffs are reduced – a conclusion that is clearly in sync with traditional trade theory, although the gains are ultimately rather small. What is troubling is the fact that despite these welfare losses from this remaining level of protection, tariffs in the olive oil sector remained. The political economy of tariff policy, especially in a supra-national institution such as the EC, is rather complex. Looking at the welfare analysis results, the domestic market findings may provide some insight. In this market, the domestic producer losses equal the domestic consumer gains, resulting in no net change in welfare effects due to the elimination of the tariff. This could show that the losses to domestic producers from free trade (and thus the gains had from protectionism) are so significant as to politically warrant a high level of protection. In the import market, the small amount of Moroccan imports of olive oil to the EC accounts for the small size of the welfare gains, which are not large enough to outweigh the political consequences of the domestic producer loss.
V. Sensitivity Analysis

Estimating the costs of protectionism is a challenging task – data is often difficult to obtain and transform, calculations often become tangled, and conclusions often differ from actual policy choices. However, further application of my adaptation of the Hufbauer and Elliot and Messerlin models could result in more robust results and stronger evidence against the inclusion of agricultural, or any form of, protectionism within EMPAs. Studying a larger sector or multiple countries, with a larger data set, would strengthen the analysis significantly.

VI. Conclusion

Although Doha may be dead, preferential trading agreements continue to remain extremely popular, especially along North-South lines. These agreements are negotiated with the intentions to not only promote increased trade but to foster economic and social development as well. Objectives such as these are difficult to achieve when the agreements, such as the EMPA between the EC and Morocco, are riddled with protectionist safeguards, especially in areas such as agriculture, where differentials in wages and resource endowments often favor production in developing nations.

The paper performed a case study of the agreement between the EC and Morocco, especially the provision which called for only a 10% reduction in the tariffs on olive oil imported from Morocco into the EC. Through an empirical welfare analysis, it was calculated that small welfare gains were found from increasing trade liberalization and that trade creation occurred, especially when trade was liberalized further than currently agreed upon in the EC-M. Rather than slapping high tariffs on developing country exports, such as Moroccan olive oil, developed countries, such as the EC, should take measures to reduce their own inherent agricultural protectionism. Doing so could lead not only to more beneficial preferential trade agreements, but to success on the multilateral level as well.
References


Pm' + t = 2,962.793
Pm' + t' = 2,840.193
QD = 63.9691
Q'D = 58.9680

\[ swyz = [(63.9691) (2,962.7931 - 2,840.193)] + [(1/2) (2,962.7931 – 2,840.193) (63.9691-58.9680)] \]
\[ swyz = [(63.9691)(122.6)] + [(1/2)(122.6)(5.00109)] \]
\[ swyz = 7,842.6233 + 306.36726 \]
\[ swyz = 8,149.1905 \]

Consumer surplus gain/Domestic producer surplus loss from shifting to the EMPA

Net Outcome: Zero welfare changes (neither trade creation nor trade diversion occurs)
FIGURE 2: DOMESTIC MARKET - UNDER EMPA TO FREE TRADE

\[
\begin{align*}
\text{Pm} + t' &= 2,840.193 \\
\text{Pm}' &= 1,736.793 \\
\text{QD} &= 58.96801 \\
\text{Q'D} &= 26.17996 \\
\text{swyz} &= [(26.17996)(1103.4)] + [(1/2)(1,103.4)(32.78805)] \\
\text{swyz} &= 28,886.9786 + (1/2)(36,178.33437) \\
\text{swyz} &= 46,976.13505 = \text{Consumer surplus gain/Domestic producer loss from shifting from the EMPA to Free Trade} \\
\text{Net Welfare Effect} &= \text{Zero welfare changes (neither trade creation nor trade diversion occurs)}
\end{align*}
\]
FRUITLESS ENDEAVORS

FIGURE 3: IMPORT MARKET - BEFORE THE EMPA TO UNDER THE EMPA

<table>
<thead>
<tr>
<th>Pm' + t</th>
<th>Pm' + t'</th>
<th>Pm' − t'</th>
<th>Pm' − t</th>
</tr>
</thead>
<tbody>
<tr>
<td>3,578.904409</td>
<td>3,456.304409</td>
<td>1,249.504409</td>
<td>1,126.904409</td>
</tr>
</tbody>
</table>

QM = 0.0000016481
Q'M = 0.0000017413

Consumer Surplus Gain = \(abcd\)

\[abcd = [(0.0000016481)(3,578.904409 - 3,456.304409)] + [(1/2)(3,578.904409 - 3,456.304409)(0.0000017413 - 0.0000016481)]\]

\[abcd = 0.00020675 = \text{Consumer Surplus gain from adopting the EMPA}\]

Producer Surplus Gain = \(eigh + ifQ'MQ'M\)

\[eigh = [(0.0000017413)(1,249.504409 - 1,126.904409)]\]
\[eigh = [(0.0000017413)(122.6)]\]
\[eigh = 0.00021349\]

\[ifQ'MQ'M = [(0.0000017413 - 0.0000016481)(1,126.904409)\]
\[ifQ'MQ'M = 0.000318609 = \text{Producer surplus from adopting the EMPA}\]
Tariff Revenue Loss = $ijh$

\[ ijh = [(0.000017413 - 0.000016481)(1,249.504409 - 1,126.904409)]\]

\[ ijh = 0.0000114355 \] Tariff revenue loss from adopting the EMPA

Net Welfare Changes = 0.00020675 + 0.000318609 - 0.0000114355 = 0.000513926

Adopting the EMPA results in a net welfare gain, and thus trade creation occurs.

**Figure 4: Import Market - Under EMPA to Free Trade**

Consumer Surplus Gain = abcde

\[ abcde = [(0.000017413)(3,578.904409 - 2,352.904409)] + [(1/2)(3,578.904409 - 2,352.904409)(0.000017413 - 0.00148436)]\]

\[ abcde = 0.0921193 \] Consumer Surplus gain from shifting to free trade

Producer Surplus Gain = cdgf + deQ'MQ_M

\[ cdgf = 0.00148436(2,352.904409 - 1,249.504409) \]
\[ cdgf = (.000148436)(1,103.4) \]
\[ cdgf = .1638946643 \]

\[ deQ'MQM = (.000148436 - .0000017413)(1,249.504409) \]
\[ deQ'MQM = .18342063 \]

\[ abcde + deQ'MQM = .347315297 = \text{Producer Surplus gain from shifting to free trade} \]

Tariff Revenue Loss = \( degh \)
\[ degh = [(0.000148436 - 0.0000017413)][(2,352.904409 – 1,249.504409)] \]
\[ degh = .16177 = \text{Tariff Revenue Loss from shifting to free trade} \]

Net Welfare Changes = \(.0921193 + .347315297 - .16177 = .27525 \)

Shifting to free trade causes positive welfare gains, and thus trade creation occurs
MIGRATION AND FDI: IS THERE A DYNAMIC BILATERAL COMPLEMENTARY RELATIONSHIP BETWEEN FOREIGN TERTIARY STUDENT ENROLLMENT AND FDI OUTFLOWS?

Jennifer Noh
Georgetown University

I. Introduction

Traditional migration theory stems from the Hecksher-Ohlin (1933) model and predicts that migration has a negative impact on the source country’s productivity as its skilled labor market is depleted. This concept, otherwise known as the “brain drain,” is widely accepted as a phenomenon related to the emigration of high-skilled workers. Studies on the interaction between the international labor and capital markets, markedly migration and FDI flows, have shown that emigration of skilled workers and FDI inflows are negatively correlated and act as substitutes. However, within the last two years, new literature has been published questioning whether FDI and skilled migration are complements or substitutes (Agost et al 2006), and the new school of thought advocates that while brain drain and FDI inflows are negatively correlated contemporaneously, skilled migration is associated with future increases in FDI inflows (Kugler and Rapaport 2005). Incited by the idea that migration acts as a catalyst for forging business networks between the host country and the source country, which effectively increases FDI outflows from the host country to the source country, a dynamic complementary relationship between migration and FDI is inferred. While there is a good amount of recent literature addressing the question of whether high-skilled migration inflows and FDI outflows are substitutes or complements or the effect of tertiary education as an institutional determinant of FDI inflows in a given country, there is a dearth of literature focusing specifically on the relationship between FDI outflows to the source country and foreign tertiary student enrollment in the host country. As the graduates of higher education are the future high-skilled workers, I would like to see if the relatively new theory of the dynamic complementary relationship between FDI and migration applies to this subsection of migrants.

II. Literature Review
The theory that I intend to test is whether or not there is a bilateral complementary relationship between tertiary student enrollment in the U.S. and U.S. FDI outflows to the source country. The link between migration from developing countries to developed countries and FDI outflows from developed countries to developing countries has not been widely analyzed in migration literature as it has predominantly focused on different push factors as main determinants of increasing world migration (Agost et al 2006).

In contrast to standard trade models’ prediction that FDI and migration act as substitutes, the newer literature questioning this conventional understanding of migration and FDI is spearheaded by scholars such as Rauch, Gao, Tong, Buch, Kleiner and Toubal, Kugler and Rapoport, Docquier and Lodigiani and Federici and Giannetti.

In “Diasporas and Development: Theory, Evidence, and Programmatic Implications,” Rauch (2003) emphasizes the importance of diasporas in promoting international trade and investments. By creating or substituting for trust in a weak international legal environment and by also providing market information or supplying referral services, Rauch argues that skilled migrants have catalyzed FDI flows back to their home countries (Docquier and Lodigiani 2006).

In “Ethnic Chinese Networks and International Investment: Evidence from Inward FDI in China” written by Gao (2003), the strong link between diasporas and inward FDI is further studied and confirmed in the specific case of China. Gao measures the magnitude of the diaspora externality by considering both the population share of ethnic Chinese and the log of the absolute population of ethnic Chinese in the source country. Also in 2003, Tong further contributes to this discussion in her paper, “Ethnic Networks in FDI and the Impact of Institutional Development.” She studies the role of ethnic Chinese in promoting bilateral investments by using the product of the numbers of ethnic Chinese in pairs of countries in 1990 in a gravity model framework.

In “Where Enterprises Lead, People Follow? Links between Migration and German FDI,” written in 2003, Buch, Kleiner and Toubal demonstrate that German FDI outflows and migration inflows are strongly complementary. By computing gross and net stocks of migrants in order to obtain proxies for the community of Germans living abroad and of foreigners living in Germany, they found that FDI are complements to migration. More specifically, the evidence for the complementary relationship between the stocks of German migrants and the stocks of German FDI abroad are much stronger than the evidence for the immigration of foreigners and FDI inflows.

It is not until “Skilled Emigration, Business Networks and Foreign Direct Investment,” written by Kugler and Rapoport in March 2005, is this acknowledgment of migration encouraging FDI inflows put in the context of the established “brain drain” theory. Kugler and Rapoport argue that while emigration of workers into a country may mitigate the incentives for FDI from the host to the origin country of migrants to some extent, “brain drain” acts as a catalyst for the creation of international business networks and is positively correlated with ensuing capital inflows. In other words, the relationship between migration and FDI is characterized by the offsetting effects of “contemporaneous substitutability” and “dynamic complementarity,” or past skilled migration associated with an increase
in current FDI inflows. They combined U.S. census data on immigration stocks by country of origin and education level in 1990 and 2000 with data from the U.S. Bureau of Economic Analysis on FDI outflows by destination country and sector. The dependent variable of the model is the change in each country’s sectoral capital stock that is attributed to accumulated FDI of U.S. origin with depreciation accounted for. They used GDP, GDP per capita, and regional dummies for Europe and Latin America as controls with interaction terms between regional dummies and migration variables to make allowance for regional determinants of migration and FDI patterns.

In “Skilled Migration and Business Networks,” written in October 2006, Docquier and Lodigiani estimate the magnitude of business network externalities in dynamic empirical models of FDI-funded capital accumulation through cross-sectional and panel frameworks. They use original data on capital and migration stocks rather than flows and isolate the skilled diaspora. In the cross-section model focusing on the period of 1990-2000, their sample of 114 countries reveals strong network externalities associated to the skilled diaspora. The effects were stronger in democracies and countries exhibiting intermediate corruption index. Furthermore, because the size of the diaspora matters, larger countries are more likely to benefit from skilled migration (Docquier and Lodigiani 2006).

In “Temporary Migration and Foreign Direct Investment,” written in December 2006, Federici and Giannetti provide the first dynamic continuous time model that analyzes the complementary relationship between migration and FDI and introduces a “revelator effect” variable, that represents the stock of information about the more or less favorable environment of FDI in the home country. They assume that migration is temporary, return migrant flow equals migrant outflow in each period, and investment is only due to FDI inflows coming from emigrants’ host countries. The continuous time model shows a convergence of the economy towards its partial equilibrium growth path, and numerical simulations confirm the complementary relationship. As a result, the policy suggestions involve encouraging labor mobility and increasing skill acquisition of migrants while they are in the host country.

Apart from the existing literature, another possible reason why skilled migration and FDI outflows are dynamically complementary is that even if the foreign graduates return to their home countries immediately following graduation, there is now a group of skilled workers in the home country that have an American education. As a diploma from a U.S. university serves as a signal that he or she has the highest level of education available internationally, a U.S. investor may see that as further incentive to invest abroad due to the abundance of reliable skilled labor.

In summary, the literature has evolved from the study of the determinants of skilled migration and FDI outflows independently to the fusion of the two. There is adequate evidence backing up the theory of a dynamic complementary relationship between skilled migration and FDI outflows in contemporary literature.

III. Data and Empirical Strategy
I have decided that the fixed effects panel data method is most appropriate for my thesis because it will allow me to analyze a cross-section of countries over a given time span. Not only will I be able to draw conclusions based on trends over several years, but I will be able to account for country-specific characteristics that may affect foreign direct investment. For instance, geographical, language, and cultural proximity, as well as existing bilateral trade relationships, likely affect FDI, and would be accounted for in my proposed model. Variables I am able to omit as a result are dummy variables for region and distance, dummy variables for whether or not English is an official language, and competitiveness rankings.

The countries that will make up my 25 observations are: India, China, Korea, Japan, Canada, Taiwan, Mexico, Turkey, Germany, Thailand, United Kingdom, Indonesia, Colombia, Brazil, Hong Kong, Kenya, France, Nigeria, Pakistan, Malaysia, Venezuela, Russia, Nepal, Jamaica, and Singapore. The time interval of time $t$ is between 2001-2005.

Because I have data on both the total FDI from the U.S. and the professional, scientific, and technical services FDI (PST FDI) from the U.S., I will run regressions on two different dependent variables. Since the business networks branching from the tertiary students in the U.S. would most likely be in the professional, scientific, and technical services sectors, it will be interesting to see how tertiary student migration is related to U.S. FDI outflows to high-skilled sectors in a given country compared to total U.S. FDI outflows. I determined my control variables by studying the relevant literature in order to find the other determinants of FDI. From my literature review, I found that Kugler and Rapoport used GDP, GDP per capita, and regional dummies for Europe and Latin America as controls with interaction terms between regional dummies and migration variables to make allowance for regional determinants of migration and FDI patterns in “Skilled Emigration, Business Networks and Foreign Direct Investment” (Kugler and Rapaport 2005). Thus, I include the control variable of GDP per capita, but omit GDP since I am including total population in year $t-1$ to account for the size of the country and market, and I would like to avoid the problem of endogeneity. The other control variable, income from FDI in year $t-1$, is included because it seems logical that return on past investment affects future investment.

Some other determinants of FDI that I considered and decided to exclude include regional and language dummies as well as Gini coefficients, the data of which was not available for every year. Also, I am assuming it will be accounted for by my panel data analysis. In addition, foreign exchange fluctuations and tax rates are also omitted since the literature showed them to have weak correlations with FDI outflows.

I obtained my data from several reliable sources. I used data from the U.S. Department of Commerce Bureau of Economic Analysis for the dependent variables, “Total FDI Outflows from the U.S. to country i (in $ mil.) in time $t$” and “Professional, scientific, and technical services FDI Outflows from the U.S. to country i in time $t-1$” as well as for the control variable, “Income from FDI in country i, time $t-1$.” The data for the variable of interest, “No. of tertiary students from country i in U.S. in time $t-1$,” is from the Institute of International Educations Network’s Report on International Educational Exchange,
Open Doors. According to Open Doors, international students in the U.S. are defined as students who are neither U.S. citizens, immigrants, nor refugees, thus excluding permanent residents. The data for “Real Per Capita GDP in country i, time t-1” and “Population in country i, time t-1” are from CountryWatch. The table below is a summary of the variables and their respective sources:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>fdi&lt;sub&gt;it&lt;/sub&gt;</td>
<td>Total FDI Outflows from the U.S. to country i (in $ mil.) in time t</td>
<td>U.S. Department of Commerce Bureau of Economic Analysis</td>
</tr>
<tr>
<td>student&lt;sub&gt;it-1&lt;/sub&gt;</td>
<td>No. of tertiary students from country i in U.S. in time t-1</td>
<td>Institute of International Educations Network’s Report on International Educational Exchange, Open Doors</td>
</tr>
<tr>
<td>student&lt;sub&gt;it-2&lt;/sub&gt; (student&lt;sub&gt;hist&lt;/sub&gt;)</td>
<td>No. of tertiary students from country i in U.S. in time t-2</td>
<td>Institute of International Educations Network’s Report on International Educational Exchange, Open Doors</td>
</tr>
<tr>
<td>fdi&lt;sub&gt;_inc&lt;/sub&gt;it-1</td>
<td>Income from FDI in country i, time t-1</td>
<td>U.S. Department of Commerce Bureau of Economic Analysis</td>
</tr>
<tr>
<td>percapgdp&lt;sub&gt;it-1&lt;/sub&gt;</td>
<td>Real Per Capita GDP in country i, time t-1</td>
<td>CountryWatch Data</td>
</tr>
<tr>
<td>pop&lt;sub&gt;it-1&lt;/sub&gt;</td>
<td>Population in country i, time t-1</td>
<td>CountryWatch Data</td>
</tr>
<tr>
<td>pst_fd&lt;sub&gt;i&lt;/sub&gt;it-1</td>
<td>Professional, scientific, and technical services FDI Outflows from the U.S. to country i in time t-1</td>
<td>U.S. Department of Commerce Bureau of Economic Analysis</td>
</tr>
</tbody>
</table>

Upon creating graphics of the relationship between the independent variables, ln_student and ln_student<sub>hist</sub>, and the dependent variables, ln_fdi and ln_pst_fd<sub>i</sub>, by country, I found that the lines resembled little correlation. Emerging markets such as India and China, the FDI inflow from the U.S. remained relatively constant over the period of five years and the student enrollment fluctuated, whereas in a country such as the United Kingdom, student enrollment has remained relatively constant over the period of five years and the level of FDI outflows has fluctuated more. While the correlation between FDI outflows and student enrollment appeared most positive in the lagged models, the relationship seemed almost quadratic in the ln_fdi and ln_student graphic.

**IV. The Model and Regression Results**

My preliminary empirical results consist of four fixed effects models based on panel data. They are as follows:
I will run the above mentioned regressions at the 95% confidence interval. Also, when
the term dynamic model is used, I am equating the dynamic model to the lagged model.
The term does not imply any other further calculations beyond the addition of a lagged
variable.

I first tested for heteroskedasticity by running a non-robust model and a robust model
and comparing the standard errors. I found that heteroskedasticity does in fact exist and
will therefore use the robust option.

In the robust fixed effects model using ln_fdi as the dependent variable, the independent
variable of interest, ln_student, has a coefficient of -1.54, a t-value of -1.68 and a p-
value of 0.103 at the 95% confidence interval, which is relatively reassuring. The variable,
ln_percapgdp, is the most statistically significant, with a t-value of 1.87 and a p-value of
0.070. The variables, ln_student and ln_fdi_inc, are both negative, while ln_percapgdp and
ln_pop have positive coefficients. These coefficients represent the elasticity of U.S. FDI
outflows with respect to the independent variable, with student enrollment and real per
capita GDP being the statistically significant variables of interest. It is surprising that the
coefficient of ln_fdi_inc is negative, but since it is such a small number, I will interpret
this as having no effect.

In the second model, I replaced student enrollment in time t-1 with student enrollment
in time t-2 in order to run a dynamic model. The coefficient of the variable of interest,
ln_student_hist, increased from the static model coefficient of -1.54 to -0.78. However, the
variable of interest is no longer statistically significant.

In the third model, I have U.S. FDI outflows into the professional, scientific, and technical
services sectors as the dependent variable. I again use the robust fixed effects model
as heteroskedasticity was found to be present as evidenced by the disparity in standard
errors. The results are rather surprising. The variable, ln_percapgdp is no longer statistically
significant, while ln_student and ln_fdi_inc are now very much statistically significant.
However, the elasticity of U.S. FDI outflows into the professional, scientific, and technical
services sectors with respect to tertiary student enrollment in the U.S. is even more negative.
There are also sign changes, as the elasticity of FDI into these particular sectors with respect to income from FDI in the previous year now being positive and statistically significant, and the elasticity of FDI in these sectors with respect to real per capita GDP now being negative but no longer statistically significant.

In the fourth model, I again replaced student enrollment in time t-1 with student enrollment
in time t-2 in order to run a dynamic model version of the third model. The coeffi-
cient of the variable of interest, ln_student_hist, increased from the static model to -7.06. However, the variable of interest is no longer statistically significant.

Breaking the analysis down by variables, the signs of the coefficients of the statistically significant variables make sense for the most part. The coefficients for student enrollment are negative for all four regressions, and student enrollment is more substitutable for PST FDI than FDI. The coefficients for income from the previous year’s FDI investments are positive, which is as expected since higher returns would indicate a healthy investment environment for future investment. I found that none of the coefficients for population were statistically significant nor economically significant. The puzzling outcome to decipher is that of the effect of per capita GDP on FDI outflows all else equal. For the FDI model, the coefficient is positive and statistically significant for the static model, but is no longer statistically significant and less positive in the lagged model. As for the PST FDI model, the coefficient in front of per capita GDP is negative and statistically insignificant in the static model, yet very much positive and statistically significant in the lagged model. Therefore, I am left concluding that the relationship between an increase in per capita GDP will result in an increase in total FDI outflows in the static model, and that an increase in per capita GDP will result in an increase in FDI into the professional, scientific, and technical service sectors in the dynamic model. This may indicate that PST FDI only makes up a small percentage of total FDI and thus it was a sound decision to include PST FDI as a dependent variable in my analysis.

The table below is a summary of the results, with the significant coefficients in highlighted:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model</th>
<th>Coefficient</th>
<th>Std. Err</th>
<th>T-Stat</th>
<th>P-Value</th>
<th>Stat. Signif?</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln_student</td>
<td>FDI Static</td>
<td>-1.542</td>
<td>0.919</td>
<td>-1.68</td>
<td>0.103</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>FDI Lag</td>
<td>-0.783</td>
<td>1.662</td>
<td>-0.47</td>
<td>0.644</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>PST FDI Static</td>
<td>-7.282</td>
<td>2.709</td>
<td>-2.69</td>
<td>0.013</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>PST FDI Lag</td>
<td>-7.061</td>
<td>2.499</td>
<td>-2.83</td>
<td>0.012</td>
<td>Yes</td>
</tr>
<tr>
<td>ln_fdi_inc</td>
<td>FDI Static</td>
<td>-0.027</td>
<td>0.356</td>
<td>-0.08</td>
<td>0.94</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>FDI Lag</td>
<td>0.666</td>
<td>0.561</td>
<td>1.19</td>
<td>0.251</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>PST FDI Static</td>
<td>1.63</td>
<td>0.531</td>
<td>3.08</td>
<td>0.005</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>PST FDI Lag</td>
<td>1.936</td>
<td>0.605</td>
<td>3.2</td>
<td>0.006</td>
<td>Yes</td>
</tr>
<tr>
<td>ln_percapgd</td>
<td>FDI Static</td>
<td>4.782</td>
<td>2.55</td>
<td>1.87</td>
<td>0.07</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>FDI Lag</td>
<td>4.654</td>
<td>4.21</td>
<td>0.39</td>
<td>0.699</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>PST FDI Static</td>
<td>3.345</td>
<td>4.691</td>
<td>-0.69</td>
<td>0.495</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>PST FDI Lag</td>
<td>3.919</td>
<td>3.081</td>
<td>1.92</td>
<td>0.073</td>
<td>Yes</td>
</tr>
<tr>
<td>ln_pop</td>
<td>FDI Static</td>
<td>0.009</td>
<td>0.145</td>
<td>0.06</td>
<td>0.953</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>FDI Lag</td>
<td>-0.073</td>
<td>0.207</td>
<td>-0.35</td>
<td>0.729</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>PST FDI Static</td>
<td>0.091</td>
<td>0.382</td>
<td>0.24</td>
<td>0.813</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>PST FDI Lag</td>
<td>0.023</td>
<td>0.5</td>
<td>0.14</td>
<td>0.888</td>
<td>No</td>
</tr>
</tbody>
</table>
V. Conclusion

From my preliminary empirical results, I have concluded that a change in a country’s real per capita GDP has a strong partial effect on the change in U.S. FDI outflows to that country and that an increase in income from FDI in the previous year has a strong partial effect on the change in U.S. FDI outflows into the professional sectors. However, the elasticity of U.S. FDI outflows with respect to tertiary student enrollment in the U.S. is -1.54, and the elasticity of U.S. FDI outflows into the professional, scientific, and technical services sectors with respect to tertiary student enrollment in the U.S. is even more negative at -7.28. Therefore, I fail to reject my null hypothesis that there is no positive static relationship between a change in tertiary student enrollment in the U.S. and change in U.S. FDI outflows to the source country. This is in line with the literature that there is a substitutional static relationship between skilled migration and FDI.

As for the dynamic models, the elasticity of U.S. FDI outflows with respect to tertiary student enrollment in the U.S. became less negative for both total FDI outflows and FDI outflows into the professional, scientific, and technical services sectors. Although the coefficients are still negative, due to the increase in the coefficient when going from the static to dynamic models, I conclude that theory of a dynamic bilateral complementary relationship between FDI and skilled migration somewhat applies to the subgroup of tertiary students. These results are in accordance to my expectation that an increase in the tertiary student enrollment in the U.S. would result in an increase in U.S. FDI outflows due to the establishment of business networks. However, it is surprising that the dynamic complementary relationship is stronger in the total FDI dynamic model than in the professional, scientific, and technical FDI dynamic model. This is because I would think that the foreign student graduates would go into the professional, scientific, and technical service sectors and the establishment of business relationships would result in a stronger complementary relationship.

VI. Areas of Improvement and Policy Implications

It would be interesting to further investigate longer lag times so as to allow for the tertiary students to be in the workforce longer in order to be in a position of influence before the effect of their business networks can be realized. However, this will not be pursued at the current moment due to the inaccessibility of additional data. I recognize that this is a major weakness of this paper since with only a lag time of t-2, it is difficult to assess whether or not the positive trend would continue over more time. Another potential weakness that my colleague, Tham, pointed out is that the forging of business networks is only applicable to students who pursue employment opportunities in the U.S., which is also dependent on the issuance of H1B visas to graduating foreign graduates each year (Tham). There also may be some discrepancy between countries such as Great Britain, which has a comparable education system to the U.S. as well as close relations, and Nigeria, whose education system is sub par to the U.S.’s. According to the scatter plots, there appears to
be a gap towards the median of the graphics. In general, the unavailability of data proved to be a major hindrance in proving a long-term relationship between foreign tertiary enrollment and FDI outflows. In addition to running the regressions with lagged variables going further back in time, it would also be interesting to obtain the foreign tertiary student enrollment from more countries. I would then run regressions after grouping the countries by geographically region or by OECD and non-OECD countries.

Another point of concern is the shock of September 11, 2001 and its effect on both U.S. FDI outflows and U.S. foreign tertiary enrollment. Although the study focuses on the years 2001-2005, the lag variables in time t-1 and t-2 utilize data from prior to the pivotal year of 2001. Because of the sharp decrease of FDI outflows and foreign tertiary student enrollment post September 11th, it is possible that this may distort the trend results over time. For future study, it would be interesting to fix any of the data in the years 1999-2000 to the 2001 numbers to see if this improves my regression results.

Another important note is the problem of double causality that may be found in this analysis. As with most regression models concerning FDI, there is the question of which direction the causality is going in. However, because I am focusing on the correlation of whether or not the relationship is that of substitutes or complements, this problem does not affect my paper as much as if I were focusing on causality.

If a dynamic complementary relationship does, in fact, exist between foreign tertiary student enrollment and FDI inflows into the source country, a significant policy implication results. If a country would like to attract more FDI, it could encourage study abroad in the countries it is looking to increase FDI inflows from. As for companies in the host country, if they are looking for FDI investment opportunities abroad, the employment of former foreign students may facilitate the flow of investment information and lead to more FDI outflows.

References

Agost, Solferino, Tria, (2006): The Migration and FDI Puzzle: Complements or Substitutes?
Docquier and Lodigiani (2006): Skilled Migration and Business Networks
Hecksher and Ohlin. (1933): Interregional and International Trade.
Kugler and Rapoport (2005): Skilled Emigration, Business Networks and Foreign Direct Investment.
THE DETERMINANTS OF INTERNATIONAL TERRORISM: ITS PERPETRATORS AND THEIR TARGETS

David Wolff
Dartmouth College

ABSTRACT

Several recent economic studies have failed to identify the hypothesized correlation between poverty and terror. In this paper, I will examine the results of those studies and offer a more precise analysis of the poverty/terror relationship. The main contribution of this paper is to disaggregate 'terrorism' into two different measures – the origin of terrorists as compared to the targets of terrorism – which should allow for a more precise identification of the factors influencing both measures. I show that the targets of terrorism are correlated with a country's economic performance, while political and demographic characteristics are more highly correlated with the origins of terrorists.

I. Introduction

Since the September 11th terrorist attacks against the United States, public officials such as George Bush and Al Gore, and academics like Joseph Nye, Dean of the Kennedy School and Laura Tyson, Dean of the London School of Economics, have argued that poverty causes terrorism. However, several recent economic studies have failed to identify this hypothesized correlation between poverty and terror (see for example Abadie 2006). In this paper, I will examine the results of those studies and offer a more precise analysis of the poverty/terror relationship. The main contribution of this paper is to disaggregate ‘terrorism’ into two different measures – the origin of terrorists as compared to the targets of terrorism – which should allow for a more precise identification of the factors influencing both measures. In the paper, I will test the same variables used by Abadie
David Wolff (2006) in order to show how the political and demographic variables are more highly correlated with the origins of terrorists, while the macroeconomic variables are more highly correlated with the targets of terrorism.

The definition of terrorism is disputed, but in this paper, I will be using the U.S. State Department’s definition: “the term ‘terrorism’ means premeditated, politically motivated violence perpetrated against noncombatant targets by subnational groups or clandestine agents, usually intended to influence an audience.”

One of the first analyses to address the link between poverty and terrorism was Alan Krueger’s 2003 article “Education, Poverty and Terrorism: Is there a Causal Connection?” Krueger finds that the effect of GDP/capita on terrorism is insignificant once levels of civil liberties have been controlled for. Alberto Abadie expands upon these findings in his 2006 article “Poverty, Political Freedom, and the Roots of Terrorism.” Abadie finds that the risk of a terrorist attack is not significantly higher for lower incomes once one has controlled for geography and political freedoms.

Abadie and Krueger arrive at the same conclusions, but they do so by testing two different variables. Krueger is testing GDP/capita on the likelihood of becoming a terrorist, while Abadie is testing it on the likelihood of incurring a terrorist attack. In this article, I will combine both of these methods in order to compare their conclusions. Using the State Department’s yearly international terrorist reports, I will develop a variable for the number of terrorist attacks against a country as well as the number of terrorists from a country, before analyzing the effect of economic indicators on both dependent variables. By drawing both measures from the same data set, I will be able to compare the results for both dependent variables and will therefore be able to identify how various factors affect each component of terrorism differently.

The rest of this paper will be organized as follows. I will summarize each of the variables, as well as present the hypotheses. I will present the multivariate analysis, and I will then offer a number of robustness tests for my findings. Finally, I will review my conclusions.

II. Multivariate Analysis

Table 1 reproduces the results of the basic regressions run to test the hypotheses of this paper. Columns 1 through 4 test the effect of the independent variables upon the number of terrorists from a country, while columns 5 through 8 test the effects of the same variables upon the likelihood of attacks happening against a country. The two dependent variables used in this paper are (1) “number of terrorists” and (2) “number of terrorist attacks.”

Column (1) illustrated the independent effects of three economic variables on the number of terrorists originating in a country. Neither GDP/capita nor trade is significant at conventional significance levels in this initial regression.

Column (2) adds measures for political rights, political rights squared, and education to the original regression. Political rights – coded so that higher numbers imply fewer rights - are positive and significant at the 5% level, showing that decreased amounts of
political rights are correlated with increases in the number of terrorists. The Political Rights squared variable is negative and significant at the 5% level, showing that there is a non-linear effect of political rights. As Abadie argues, highly authoritarian regimes are most likely able to use repression to limit the ability of individuals to become terrorists. The last variable, education, is insignificant in the regression, indicating that once economic size and political rights are controlled for, there is no significant correlation between years of schooling and terrorism.

Column (3) adds measures of religious, ethnic, and linguistic fractionalization to the regression run in column (2). In this regression, the political rights variable remains significant while the religious fractionalization measure is negative and significant, implying that increased levels of religious fractionalization reduce the number of terrorists produced. This is counter-intuitive as conventional wisdom would have us believe that increased fractionalization should increase the number of terrorists. One possible explanation is that this study is on the determinants of international terrorism, and increased religious fractionalization might be correlated with increases in domestic terrorism. An increasingly fractured populace would be more likely to produce terrorists at home than it would be to export them to other countries. Looking at the data, the three countries which produced the most terrorists have below average levels of religious fractionalization: Pakistan (.38), Palestine (.34), and Columbia (.15), vs. an average of .433.

Column (4) adds three additional geographic measures to the regressions from column (3). The geographic controls used - country area, average country elevation, and fraction of the country in the tropics - are all insignificant in the regression. Adding the controls does not qualitatively change the results. A lack of political rights is still significantly correlated with an increase in the number of terrorists, while increased religious fractionalization is correlated with a decrease in the number of terrorists. After including the controls, religious fractionalization becomes negative and significant.

Columns (5) through (8) test the effects of the same set of variables on the determinants of terrorist attacks. Column (5) shows the effects of the same two economic variables as were used in column (1). GDP/capita is positive and significant at the 5% level. This indicates that increased living standards are positively correlated with an increase in the number of terrorist attacks a country receives. This coincides with intuition, as the countries with the highest standards of living are typically the countries that terrorists target in an attempt to garner the most attention. The variable for trade is negative and significant, indicating that countries with a higher percentage of trade to exports receive fewer terrorist attacks. This is contrary to intuition as one would hypothesize that countries which are more open to trade are also countries which are the easiest targets for terrorists and are therefore going to be more likely to be attacked. However, when one goes back to look at the data, the countries that were targeted the most – India and the United States – have significantly lower trade ratios than the rest of the world: the international average is 40% while the U.S. and India are at 9% and 14.7% respectively. While nobody would argue that the low trade ratios of these countries is the reason they are being targeted, the coincidental correlation has such a large effect on the international average as to make the trade vari-
able significant and negative. Nevertheless, running the regression excluding India and the United States, the trade variable remains statistically significant and negative. Future research is needed to either refine the trade variable so that it captures the appropriate effect, or to explore and attempt to explain this negative correlation.

Columns (6), (7), and (8) all illustrate the same general picture. Adding control variables for fractionalization, political rights, schooling, and geography does not affect the significance of GDP/capita or trade. None of the control variables is significant at standard significance levels, and GDP/capita and trade remain significant at the 5% and 10% levels respectively.

As was predicted, the two different elements of terrorism – its perpetrators and its targets – are affected by different variables. The number of terrorists ‘produced’ by a country is affected by socio-cultural variables such as political rights and religious fractionalization, while the targets of terrorism are not statistically correlated with anything but economic variables.

III. Robustness Tests

In the previous section, initial correlations were illustrated between several independent variables and the numbers of terrorists and attacks. This section will serve to test these conclusions through several different robustness tests. Several robustness tests were run in the original paper, and are included in the attached Table 2. However, in the interest of space, only a few tests will be discussed. The first test was to exclude potential outliers from the dataset. Two cross-border rivalries between India-Pakistan, and Israel-Pakistan produce dramatically more terrorists and attacks than most other countries within the data. The results produced by excluding these 4 outliers are included in Columns (1) and (5) of Table 2 below. There are some minor changes, most notably political rights loses its significance, and GDP/capita is only weakly significant, but overall our general conclusions regarding the importance of socio-cultural variables for the generation of terrorists and the economic variables for terrorist targets still holds. A second test was run to account for the potential bias produced by an incremental dependent variable with a high percentage of zeros. Columns (2) & (6) produce the results of these negative binomial regressions, which remain roughly constant, with the largest change being the new significance of “trade” and “linguistic fractionalization.” A last robustness test was run which included real GDP in place of GDP/capita. Theoretically, market size might be more correlated with one of the elements of terrorism (targets) than living standards. Nevertheless, substituting GDP does not change the statistical significance in any of the previously reported regressions and therefore the results are not shown here.

One of the benefits of the dataset used in this paper is that it allows me to test if there is a difference in both aspects of terrorism before and after September 11. To test this, I created two versions of each dependent variable: one for 2000 to 2001 and one for 2002 to 2003. Columns (3) and (7) represent the results for 2000-2001 while columns (4) and (8) depict the results for 2002-2003.
As these columns show, the ‘production’ of terrorists and the targets of terrorism were affected very differently by the events of 9/11. The determinants of becoming a terrorist were quite different in the two time periods. Before 9/11, none of the socio-cultural or economic variables was significant, while after 9/11 political rights and religious fractionalization became significant at the 5% level. This could represent proof that the actions of 9/11 and the subsequent reaction of the U.S. government were counterproductive. U.S. counter-terrorism policy focused on politically repressive regimes with the belief that liberating these societies would generate fewer terrorists. However, it looks like after 9/11 politically-repressive regimes are actually generating more terrorists than they were beforehand. In contrast to the ‘production’ of terrorists, the targets of terrorism were affected only by economic variables both before and after 9/11. By looking at the pre- and post-9/11 data, we again see evidence that it is important to analyze both aspects of terrorism when discussing the effects of events upon ‘terrorism’.

IV. Conclusion

This paper has analyzed the determinants of international terrorism. Building off the work of Abadie (2006) and Krueger (2004), it developed two new variables – the number of terrorists originating in a country and the number of attacks against a country - in order to test how factors might affect each variable differently. This paper showed how socio-cultural variables such as the lack of political rights and decreases in religious fractionalization were correlated with increasing numbers of terrorists produced by a country. In contrast, economic, and not political or demographic, variables appeared to be the significant factors in determining which countries were the victims of terrorism. In general, higher levels of GDP/capita and lower levels of trade openness were correlated with higher levels of terrorist attacks.

This study’s main contribution was to introduce the concept of disaggregating the terrorism measure into two components and to simultaneously compare the effects of variables on each element. As this was the first use of these variables, there is a large amount of additional research that needs to be done in the future to determine their significance. Most importantly, the data behind each variable need to be checked and expanded: first, the data need to be corroborated by non-U.S. government sources, second, the number and origin of many of the terrorists need to further researched, third, the time period which the dataset covers needs to be expanded. On the independent variables, more needs to be done to expand the covered countries, and to expand the potential controls (current conflict levels, relative defense spending, unemployment rates, demographic rates, religion, etc) which are analyzed.
References


Table 1: Primary Regressions on Terrorists and Attacks

<table>
<thead>
<tr>
<th>Economic Variable</th>
<th>D.V. = Number of Terrorists</th>
<th>D.V. = Number of Attacks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td><strong>Log GDP/capita</strong></td>
<td>-1.22</td>
<td>1.365</td>
</tr>
<tr>
<td></td>
<td>(1.19)</td>
<td>(1.072)</td>
</tr>
<tr>
<td><strong>Trade</strong></td>
<td>-0.011</td>
<td>-0.072</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.044)</td>
</tr>
<tr>
<td><strong>Political &amp; Social</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of rights</td>
<td>4.07*</td>
<td>3.14**</td>
</tr>
<tr>
<td></td>
<td>(2.38)</td>
<td>(1.95)</td>
</tr>
<tr>
<td>Lack of rights ^ 2</td>
<td>-0.435</td>
<td>-0.293</td>
</tr>
<tr>
<td></td>
<td>(0.277)</td>
<td>(0.207)</td>
</tr>
<tr>
<td>School</td>
<td>-0.088</td>
<td>0.444</td>
</tr>
<tr>
<td></td>
<td>(0.368)</td>
<td>(1.26)</td>
</tr>
<tr>
<td><strong>Fractionalization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linguistic</td>
<td>3.59</td>
<td>3.68</td>
</tr>
<tr>
<td></td>
<td>(4.59)</td>
<td>(4.46)</td>
</tr>
<tr>
<td>Ethnic</td>
<td>1.22</td>
<td>2.99</td>
</tr>
<tr>
<td></td>
<td>(6.55)</td>
<td>(6.61)</td>
</tr>
<tr>
<td>Religious</td>
<td>-4.78*</td>
<td>-6.19**</td>
</tr>
<tr>
<td></td>
<td>(2.71)</td>
<td>(2.99)</td>
</tr>
<tr>
<td><strong>Geography</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country Area</td>
<td>8.99*</td>
<td>8.99*</td>
</tr>
<tr>
<td></td>
<td>(5.39)</td>
<td>(5.90)</td>
</tr>
<tr>
<td>Elevation</td>
<td>-0.013</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>(0.164)</td>
<td>(0.253)</td>
</tr>
<tr>
<td>Tropical Area</td>
<td>-3.69</td>
<td>-3.69</td>
</tr>
<tr>
<td></td>
<td>(4.40)</td>
<td>(4.81)</td>
</tr>
<tr>
<td>R-squared</td>
<td>.04</td>
<td>.15</td>
</tr>
<tr>
<td>Observation</td>
<td>132</td>
<td>99</td>
</tr>
</tbody>
</table>

Notes: All regressions in Abadie’s original table included regional dummies for North America and Western Europe, Latin America and the Caribbean, Middle East and North Africa, Eastern Europe and Central Asia, and the rest of Asia and the Pacific. I have attempted to recreate those dummies as best as possible. Heteroscedasticity-corrected standard errors are in parentheses.

* Statistical significance at the 10% level
** Statistical significance at the 5% level
Columns (1) and (5) test only economic variables
Columns (2) and (6) add in the political and educational variables to the previous regressions
Columns (3) and (7) add fractionalization measures to the previous regressions
Columns (4) and (8) add geography measures to the previous regressions.
### Table 2: Robustness Tests

<table>
<thead>
<tr>
<th>Economic Variable</th>
<th>D.V. = Number of Terrorists</th>
<th>D.V. = Number of Attacks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Log GDP/capita</td>
<td>1.59*</td>
<td>1.21**</td>
</tr>
<tr>
<td></td>
<td>(0.832)</td>
<td>(0.323)</td>
</tr>
<tr>
<td>Trade</td>
<td>0.008</td>
<td>-0.03**</td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Political &amp; Social</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of rights</td>
<td>2.22</td>
<td>2.33**</td>
</tr>
<tr>
<td></td>
<td>(1.62)</td>
<td>(0.596)</td>
</tr>
<tr>
<td>Lack of rights ^ 2</td>
<td>-0.198</td>
<td>-0.23**</td>
</tr>
<tr>
<td></td>
<td>(0.208)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>School</td>
<td>-0.088</td>
<td></td>
</tr>
<tr>
<td>Fractionalization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linguistic</td>
<td>0.394</td>
<td>4.79**</td>
</tr>
<tr>
<td></td>
<td>(3.28)</td>
<td>(1.25)</td>
</tr>
<tr>
<td>Ethnic</td>
<td>1.05</td>
<td>-0.712</td>
</tr>
<tr>
<td></td>
<td>(4.58)</td>
<td>(1.15)</td>
</tr>
<tr>
<td>Religious</td>
<td>-4.55*</td>
<td>-4.94**</td>
</tr>
<tr>
<td></td>
<td>(2.65)</td>
<td>(1.08)</td>
</tr>
<tr>
<td>Geography</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country Area</td>
<td>7.89</td>
<td>2.89**</td>
</tr>
<tr>
<td></td>
<td>(5.09)</td>
<td>(0.809)</td>
</tr>
<tr>
<td>Elevation</td>
<td>.152</td>
<td>0.118**</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>Tropical Area</td>
<td>2.40</td>
<td>-0.407</td>
</tr>
<tr>
<td></td>
<td>(2.04)</td>
<td>(0.621)</td>
</tr>
<tr>
<td>R-squared</td>
<td>.18</td>
<td>.08</td>
</tr>
<tr>
<td>Observation</td>
<td>128</td>
<td>132</td>
</tr>
</tbody>
</table>

Notes: All regressions in Abadie's original table included regional dummies for North America and Western Europe, Latin America and the Caribbean, Middle East and North Africa, Eastern Europe and Central Asia, and the rest of Asia and the Pacific. Heteroscedasticity-corrected standard errors are in parentheses.

* Statistical significance at the 10% level

** Statistical significance at the 5% level

Columns (1) and (5) exclude India, Pakistan, Israel, and Palestine
Columns (2) and (6) are negative binomial and not ordinary least squares regressions
Columns (3) and (7) limit the dependent variables to the years 2000 and 2001 (pre-9/11)
Columns (4) and (8) limit the dependent variables to the years 2002 and 2003 (post-9/11)
The Indirect Impact of Obesity on Economic Growth

Tanja Groth
University of Saint Andrews

The purpose of this paper was to review the relationship between obesity and productivity and to model the impact of obesity on economic growth in countries from the Organization for Economic Cooperation and Development (OECD). A standard economic growth model including a set of demographic variables was formulated using data from 22 countries of the OECD over the period 1978-2002. Both a cross-country data structure and a panel data structure were modeled.

Obesity in this paper is denoted as a Body Mass Index (BMI) of more than 30kg/m2, in accordance with World Health Organization (WHO) guidelines. BMI is measured by dividing weight in kilograms by squared height in meters (or by dividing weight in pounds by height in inches and multiplying by 703).

Medically, obesity is classified as an eating disorder much like anorexia nervosa or bulimia nervosa, but in contrast to these, the underlying causes of obesity are still relatively unclear. While the prevalence for anorexia and bulimia affect less than 1% of the total population in western countries, obesity prevalence ranges from 3.2% (Japan) to 34.94% (US) in the OECD countries.

Economically, obesity can be treated as a textbook example of an individual’s optimization problem; the tradeoff between time, money and energy expenditure to achieve a given level of health. Variables which affect the optimization trade-off include relative food prices, declining or stable physical activity levels in conjunction with increased energy intake, substitution towards ‘fast’ food and foods eaten away from home due to increases in female labor force participation, and the decline in smoking.

Obesity has a significant negative impact on health in terms of its etiologic role in a variety of diseases. Chief among these are diabetes mellitus (type 2 diabetes), coronary heart disease, hypertension and hypercholesterolemia. The consequences of an excessive BMI result in substantial direct and indirect costs to both the individual and to the society. Direct costs include costs of prevention, diagnostics and treatment, whilst indirect costs are primarily due to premature mortality, disability and lost economic productivity due to illness.

There is general agreement in the literature that the influence of health on productivity has significant impact on economic growth. A healthy workforce produces more, works
harder and saves more for future consumption – key for investment in physical capital. A
healthy individual, expecting a long lifetime, is more likely to invest in education in order
to increase his consumption/utility through achieving higher socioeconomic status (SES).
The more educated a society is, the better, as a high level of education enhances life
expectancy by making the society better at absorbing new advances in medicine and tech-
nology.

Any additional year of overall life expectancy results in a 1% increase in GDP 15 years
later. This corresponds with the rule of thumb for translating labor productivity growth into
potential output, which is between 1 and 0.75 %. Presumably, this would also hold true for
a reverse effect. Estimates of years of life lost from obesity for an otherwise healthy 40-
year old adult can range from 2-3 years to 6-7 years, for a BMI of 33 and 30 respectively.
Since, for example, the US has an obesity prevalence of app. 35%, it is not unlikely that
national life expectancy could be affected.

Obesity shows large persistence effects, which indicates that we have yet to experience
the full consequences of the boom in average BMI. The explosion in obesity prevalence
that has taken place over the last 25 years is exponential in its impact; many of the direct
costs in terms of hospitalization and medical treatment only reach their peak when the
patients are above 65 years of age. Children of obese or working mothers are much more
likely to be obese themselves, and thus an increase in adult obesity leads to an increase in
childhood obesity. Studies have shown that obese children do not do as well in school as
children with normal BMIs’, suffering from lack of energy and concentration, which
affects their education. Furthermore, nutrition disorders in children affect their later health,
for example through increasing the risk of chronic conditions such as coronary heart dis-
ease, diabetes mellitus and hypertension.

Sturm (2002) finds that the impact of obesity on chronic diseases is significantly lar-
ger than the effects of current or past smoking or alcoholism. He finds that the general
health impact of obesity is equivalent to that of 20 years’ worth of aging, or in the case of
physical health, to 30 years’ worth of aging.

Defining the optimal BMI range for highest life expectancy or most years of disabili-
ty-free life has proven problematic. For example, one US study reports that the optimal
BMI for non-smoking Caucasian males of age 50 or over lies in the range 23 – 28, where-
as others report increased health risks for a similar subject group with BMI values greater
ranges of ‘healthy’ BMI are not conducive to maximal health in China and Japan, and that
lower levels of BMI result in the same relative risk for some diseases associated with obe-
sity. Conversely, results from the U.S. indicate that African-Americans do not appear to
suffer the same incidence of chronic illnesses related to specific ranges of BMI; rather,
their optimum lies considerably above what is recommended for Caucasians.

This paper uses an OLS regression to model the impact of obesity on economic
growth. The standard regression equation is as follows:

\[ GRGDPCCH = \beta_0 + \beta_1 rgdpc + \beta_2 x + \beta_3 Z + \varepsilon \]  

(2)
where $GRGDPCH$ refers to the growth rate of income per capita, $rgdpch$ is the natural logarithm of the initial income per capita, $x$ is a vector containing the levels of factors which affect the initial state, and $Z$ represents the growth rates of the variables which affect the growth rate of income. The variables included in $x$ are inequality, fertility, education, life expectancy, urban density, density and the ratio of the working population to total population. $Z$ is comprised of the growth rates of the female participation rate, investment, openness, total population, working population and obesity over the period 1978-2002.

The bulk of the data on obesity prevalence was obtained from OECD and WHO tables. Due to significant deficiencies and variance in the data available, micro level observations were used to compare and in some cases replace existing data or the lack thereof. For data that was segregated into male and female percentages, World Bank data on the ratio of males to females was used to transform the data into an overall aggregate. Most likely the results would be more accurate if separate data for males and females was used, but as most of the data available was for a national average, this approach was not feasible. A table was then constructed with the data classified according to country and year. Not all data could be extrapolated from available sources, in which case some countries were excluded from the relevant analysis. This paper uses the prevalent measure of obesity as set down by the WHO.

Time dummy variables are included in one of the models to account for global economic shocks such as the worldwide recession of the 1980s and cyclical productivity effects. These are in units of 5-year periods due to the explanation set forward above. There exists some debate in the literature regarding the validity of using country-specific dummies. Most regard it as a necessity to capture country-specific influences that arise from differences in culture and attitudes, e.g. the attitudes towards diet. In contrast, Bloom, Canning & Malaney discourage the use of country-specific dummies because it would interfere with the necessary presence of the initial income level and thus result in biased coefficient estimates. As no other paper reports this difficulty, and as the dissenters themselves use region specific dummy variables, a set of country fixed effects has been included in one of the models.

**Cross-Country Results:** Available calories per capita, taken from the UN Food and Agriculture Organization (FAO), are used as an instrumental variable (IV) for the prevalence of obesity as no other suitable variable could be found. The principal justification for using IVs is to correct for possible endogeneity in the model specification; for this purpose the included IVs cannot be correlated with the error term. In terms of the obesity variable, the instrument relevance (correlation = -0.326) is not conclusive, although instrument exogeneity is satisfied (correlation = 0.006). The other IVs are all exogenous to the residual to 13 or 14 decimal points. The coefficients and significance of the IVs and the initial values included are significantly different from the results, which may indicate endogeneity in the original specification. However, as no satisfactory IV for obesity was obtainable, further specifications will be run without the use of IVs.

**Panel Data Results:** The lack of comprehensive data for obesity made it difficult to compile 5-yearly differences for all countries. 4 countries had to be excluded entirely as
there were only 1 or 2 observations available for each. The data used for the included countries was calculated with respect to the cross-country averages for each period and the cross period averages of each country. This data set is inaccurate for use as anything other than a robustness check for the cross-country results. It does serve as an indicator for the average growth of obesity over each 5-year period in the OECD, but as this varies greatly over each country, the coefficients in the table are unlikely to be representative of actual impact.

It is likely that the problems with these socioeconomic and demographic variables stem from their interaction with each other, and from possible interactions with the growth in aggregate income. This study was unable to solve the endogeneity problems implicit in the model specifications as to do so would have required instrumenting all the included variables with relevant IVs, something which was beyond the scope of this paper.

More than 50% of adults have been defined as overweight or obese in at least 10 OECD countries, a trend which has exhibited significant growth over the last 20 years. The unprecedented increase in obesity prevalence that the world is currently experiencing is a unique occurrence in the history of mankind. Obesity seems to follow the format of what some have dubbed the ‘technophysio evolution’; a form of human evolution that is ‘biological but not genetic, rapid, culturally transmitted and not necessarily stable.’ Biologically, the human body has been programmed over the centuries to store up body fat in times of abundance in expectation of future food scarcity. However, OECD countries do not experience food scarcity issues, indicating a fundamental shift from optimal behavior in the past to optimal behavior in the present.

Although the impact of obesity on productivity is undeniable, the direct impact it has on economic growth is more difficult to demonstrate. Aggregate modeling may simply be too crude a tool to capture the complexities of the impact of obesity on productivity and economic growth. A production function that can accurately model changes in factor productivity would be a better way to model the impact of obesity, if productivity in the service industry could be modeled effectively. As the OECD economic output is comprised in large part of the service industry where differences in productivity are difficult to measure, current productivity models may not be up to the task.

Despite proven correlations between health and GDP, measures of GDP fail to adequately represent movements in health. In some cases, growth in GDP has even been proven to be health-stunting, leading some to believe that an alternate measure of social welfare is needed. In the case of obesity, the high associated risks of morbidity and mortality may negate the potential social results gained from investments into training and education, resulting in a waste of resources. This is especially relevant for the OECD countries, where high aggregate levels of income would indicate more resources for investment.
References


Bhargava, Alok; Dean T. Jamison; Lawrence Lau; Christopher JL Murray; “Modeling the Effects of Health on Economic Growth”, GPE Discussion Paper Series: No. 33 Global Programme on Evidence for Health Policy.


**References**


Peeters, Anna; Luc Bonneux; Wilma Nusselder; Chris De Laet; Jan J. Barendregt; 2004, “Adult Obesity and the Burden of Disability throughout Life”, Obesity Research, Vol. 12, pp 1145-1151.


THE DETERMINANTS OF LATIN AMERICAN IMMIGRANTS’ SETTLEMENT LOCATIONS IN THE SOUTHWESTERN UNITED STATES

Ashley Halpin
Dartmouth College

ABSTRACT

In this paper, I test whether migrant income and migrant networks in U.S. counties influence the settlement patterns of Latin American immigrants in the Southwestern United States. I use a cross-section of 2000 and 2005 county-level data for California, Arizona, New Mexico and Texas and find that both factors are important. Latin American immigration flows to a particular county in 2005 are positively and significantly correlated with Hispanic per capita income in that county in 2000 and with the share of the population that is Latin American foreign born in that county in 2000. For the broader population of total immigrants to the Southwestern United States, I find that migrant networks influence settlement patterns but Hispanic per capita income does not.

I. Introduction

Immigration is a major political issue in the United States. There is concern that immigrants are having adverse effects on the economies in which they locate (see Murdock, Saenz, and Zhai, 1999). As a result, many cities and states are implementing policies that try to deter immigrants from settling there. Most immigration policies attempt to dissuade migration to the U.S. by decreasing the economic incentives of moving (Espinosa and Massey, 1997). These policies, however, seem to be ineffective because the foreign population of the U.S. is rising (U.S. Census Bureau). An alternative policy option would be to target specific locations that are attractive to immigrant settlers.

Few researchers have examined what determines where immigrants locate, and even fewer have linked that to migration theory and the U.S.-Mexico border relationship.
Acknowledging this void, my paper focuses on the determinants of Latin American immigrants’ settlement choices in the Southwestern U.S. In particular, I examine whether the neoclassical and social capital theories of migration determine the settlement locations of immigrants. Neoclassical theory suggests that immigrant flows increase with expected earnings. On the other hand, the social capital theory is explained by migrant networks; as the foreign born population of an area increases, the flow of immigration to that area will increase, as well. My hypothesis is that immigrant networks will determine Latin American immigrants’ location choices, but that income and wages will not. Information about expected earnings in a specific area of the U.S. may not be widely available to those considering emigrating. As a result, potential migrants rely on their connections in the U.S. to decrease the cost of migrating when determining where to locate.

To test this hypothesis, I use an OLS regression of a cross-section of county-level data for the four Southwestern U.S.-Mexico border states. My main independent variables, Hispanic per capita income and a ratio of the foreign born Latin American population to the total population, are from 2000, whereas my dependent variable, a ratio of the foreign born population who entered the U.S. since 2000 to the total population, is from 2005. This time differential allows me to test how county characteristics impact the flow of immigration to that county in a future time period. If my hypothesis is correct, I will find a positive and significant correlation between the Latin American foreign-born ratio and my dependent variable, while the Hispanic per capita income variable should be insignificant. The results I find show strong evidence supporting migrant networks in determining where immigrants locate; however, there is inconclusive evidence as to whether expected earnings play a role. Both Latin American migrant networks and Hispanic earnings are significant and positively related to the inflow of immigrants in the subsequent time period; however, when the variables change to reflect earnings and networks for a broader population of immigrants, only the network variable remains significant.

These findings have some implications for the policy issues that I previously mentioned. In learning what pulls immigrants to certain areas of the Southwest, government officials are able to better target their policies. If certain county characteristics are more likely to increase immigration in the next period, governments can predict their county demographics for the future. This could be used to better allocate government funds for the subsequent period. The U.S. Border Patrol could use this information to target areas where there may be many undocumented immigrants, whom many see to be a drain on the economy. In addition, many non-governmental organizations that provide immigration services could use location determinants to better target their services.

While there are clear benefits to exploring what factors pull immigrants to certain areas of the country, little research has been done in this area. Since the 1980s, there has been much research devoted to determining the causes of immigration to the United States; however, there have been few empirical studies comparing the major schools of thought on migration theory. Espinosa and Massey (1997) compared the five major theories of migration and found that the neoclassical economics theory is not a large factor in determining who decides to migrate; however, they did find evidence supporting the social cap-
ital theory. Most other studies have focused in on one specific factor that may cause someone to choose to emigrate, such as, income and wages (Hanson and Spilimbero, 1999; Borjas, 1999; Massey, 1987), migrant networks (Massey, 1987), and job networks (Munchi, 2003).

While a literature review finds evidence that neoclassical theory (earnings) and the social capital theory (networks) determine whether an individual migrates, they do not provide evidence of the relative importance of these theories in determining where migrants decide to settle. A few studies have looked at the determinants of where immigrants locate (Borjas, 1999; Newbold, 1999, 2002). They found economic factors to have a minimal impact in choosing a destination location. My paper combines ideas from the empirical work based on migration theory with the work looking at the determinants of migrants’ settlement locations. While this is most similar to Newbold’s studies (1999, 2002) because I look at county-level factors, as opposed to household-level data, it differs in that I focus on migration theory and the relationship of immigration from the U.S.-Mexico border with the Southwestern U.S.

II. Migration Theory

In this paper, I examine the neoclassical and the social capital theories of migration (Massey, 1999) in their effectiveness at determining where migrants choose to locate in the Southwestern United States. Neoclassical theory is derived from models used to describe internal labor migration, which are based on the geographic differences of labor supply and demand. The theory simply states that a labor-abundant economy will have a low equilibrium wage compared to a capital-abundant economy, which will have a high equilibrium wage. As a result of this wage differential, people will migrate to the labor-scarce, higher wage economy from the labor-abundant economy. Over time, the labor migration will yield an equilibrium wage, which reflects the costs of migration.

In addition to the macroeconomic theory, the neoclassical theory has a microeconomic model of individual choice. In this model, individuals migrate because they determine the difference between their expected present value income after migrating and their current income is greater than their expected costs of migrating. In summary, the neoclassical theory suggests that people migrate because they expect higher earnings in their new location and these higher earnings outweigh the financial and psychological costs of migration.

The social capital theory is also used to explain immigration patterns. According to Bourdieu and Loic Wacquant, as cited in Massey (1999, 43), “social capital is the sum of the resources, actual or virtual, that accrue to an individual or a group by virtue of possessing a durable network of more or less institutionalized relationships of mutual acquaintance and recognition.” Migrant networks are a type of social capital that is made up of the connections between “migrants, former migrants, and non-migrants in origin and destination areas through ties of kinship, friendship, and shared community origin” (Massey, 1999, 44). These networks lower the costs and risks of migration, because they give immi-
grants access to employment and higher wages in their new country and information while they are still in their home country. Lower costs from migration networks increase the returns to migration, which increases the incentives to migrate. As more people migrate to one area, those migrant networks grow; connections are enhanced so that future migrants are able to find higher paying jobs. As a result, I expect future migrants to locate in counties where there are currently high proportions of immigrants in the population, which suggests strong migrant networks.

I use the OLS method to test the neoclassical and social capital theories of migration in determining where Latin American immigrants choose to locate in the U.S. The effectiveness of these theories can be tested simultaneously using the following specification:

\[
\text{ImmigrantPopulationShare}_{i,2005} = \beta_1 \text{HispEarningsPerCapita}_{i,2000} + \beta_2 \text{LAImmigrantPopShare}_{i,2000} + \beta_3 X_{i,2000} + \beta_i,2005 + \varepsilon_{i,2005}
\]

The population share of new immigrants in county \(i\) at time \(t+1\) is regressed on measures of the average earnings of Latinos in county \(i\) at time \(t\), the share of the population of immigrants from Latin America in county \(i\) at time \(t\), and a vector of control variables that are relevant in determining county \(i\)'s population share of immigrants in the subsequent time period. I use two consecutive Census time periods to account for the time lapse between when a person decides to immigrate based on information gained through his cost-benefit analysis, and when he actually arrives in his final destination. The information used to decide where one immigrates is from time period \(t\), but the immigrant will not arrive to his destination until time \(t+1\) due to preparation and travel time. In this specification, if \(\beta_1\) is significant, there is evidence supporting the neoclassical theory, and if \(\beta_2\) is significant, there is evidence supporting the social capital theory. My hypothesis states that \(\beta_1\) will be significant, but that \(\beta_2\) will not.

**III. Data and Variables Descriptions**

The data set used in this paper was compiled from the 2000 U.S. Census and the 2005 American Community Survey. I use a cross-section of county-level data for the four U.S. southwestern states that share a border with Mexico: Arizona, California, New Mexico, and Texas. The data is limited to counties larger than 65,000 people, which amounts to 88 counties. The independent variables are from the years 1999 and 2000 and the dependent variable is from 2005. Through these two periods of data, I show that the county characteristics in one time period affect the flow of immigration to that county in the subsequent time period.

To proxy for the dependent variable, recent immigrant population of county \(i\) in 2005, I use a measure of the ratio of the foreign born population in the county that immigrated to the U.S. in 2000 or later to the total population of the county (\(\text{ForBorn05enter00}\)). This measure includes only immigrants who have arrived to the U.S. since 2000. I assume no
internal migration, so an immigrant’s location is assumed to be his first post-migration settlement destination in the U.S. I also assume that this location was determined by information that the immigrants received during the 1999-2000 period when they resided in their home country. It is also important to note that this measure includes U.S. citizens of the U.S. and undocumented immigrants. The data is self-reported by household, so it most likely does not include seasonal migrants. This measure also includes all nationalities of immigrants that arrived to the U.S. since 2000. In trying to determine where Latin American immigrants locate, I would ideally use a measure that only included recent Latin American immigrants; however, this data is only widely collected by the decennial census. Since the majority of immigrants to the Southwestern U.S. are from Latin America, the measure I use will still be heavily weighted towards where Latin American immigrants locate. As a result, I find it an appropriate measure to indicate the location patterns of Latin American immigrants.

To proxy for the neoclassical theory, I use a measure of the Hispanic and Latino per capita income in 1999 (HispIncome99). In a cost-benefit analysis of migration, a higher expected income in the new location would increase the net yield of migrating. As a result, migrants will move to counties with higher expected incomes in order to increase the benefits of migration. This effect will be shown as a positive correlation between the Hispanic per capita income in 1999 and the ratio of foreign-born population to total population in 2005.

For a proxy of migrant networks, I use the ratio of the foreign born population from Latin America in 2000 to the total population of the county (ForBornLA00). A greater percentage of a population that is Latin American signifies greater migration networks because the foreign-born will be involved in a larger capacity in more diverse sectors of society due to their higher percentage of the population. Being involved in diverse areas creates connections across a county, which can be used to help new migrants to the area find better jobs. A larger population of foreign-born people in a county also signifies a higher probability that a migrant has family in that area, which is a further incentive to locate there. As a result, I expect the ratio of the population that is foreign-born Latin American in a specific county to be positively correlated with the flow of immigrants to that county in the following time period. There is also the possibility that migrant networks can become saturated, and new migrants to the area will not be able to find jobs because they were already filled by the previous period of migrants. I test this saturation effect by including the square of the 2000 foreign-born Latin American ratio. I expect the coefficient of this measure to be negative, signifying the diminishing returns of migrant networks.

In addition to these two theories, I control for other variables that might influence where a migrant decides to locate when emigrating from Latin America. Because the Southwestern United States share a border with Mexico and the majority of immigrants to those states are from Latin America, I assume that most migrants enter the U.S. via the U.S.-Mexico border. To control for the costs of locating in one county as opposed to another county that is farther from the border, I constructed a distance variable (Distance).
The variable uses the longitude and latitude of the center of each county and the same for nine major U.S.-Mexico border entry points. I use the shortest distance from the county to one of the nine border towns to represent the distance from the county to the border. Distance proxies for migration costs because traveling farther distances take more time and resources, and risk associated with traveling from one’s home country increases with distance due to information asymmetries. As a result, I expect the distance from a county to the U.S.-Mexico border to be negatively correlated with the foreign born population in that county.

I also control for the ratio of population that lives in a rural area to the total population (Rural). Between 1942 and 1964, over 4 million Mexicans came to the U.S. through the Bracero Program to temporarily work on ranches and farms (Espinosa, 1999). After this program ended, many immigrants continued coming illegally to the U.S. to fill this need for labor. This implies a demand for Mexican workers in rural sectors, which would be attractive places to locate because of job security. However, rural sectors also lack factors that immigrants may want in choosing where to locate such as, English classes, opportunities for night jobs, or public transportation. If a county has a larger population, it is more likely to have greater infrastructure, more commerce, and more social services; the cost per capita of these resources decreases as the population increases, so the government is able to provide more. The combination of these factors and others would be an incentive to move to a larger city. As a result, it is unclear whether the rural population variable will be positively or negatively correlated with the recent immigrant population ratio.

Finally, I control for state-specific policies that may affect where immigrants locate between the four states by using state dummies. For example, California’s Proposition 187, a largely anti-immigrant proposition (ACLU, 1999), may deter migrants from settling in California. The state dummies will control for state policies and sentiments that are pro-immigration or anti-immigration, which may influence where a migrant settles.

IV. Results

Table 1 reports the results for the basic OLS regression specified in equation 1. The dependent variable, the ratio of the county population that is foreign-born in 2005 and entered the U.S. in 2000 or later, is regressed on four different specifications of independent variables. Note that the income and distance variables are logged. In addition, all reported estimates are corrected for heteroskedasticity.

As expected, my measure for the neoclassical theory, Hispanic per capita income, is positively correlated with the recent immigrant population ratio in the following time period. In fact, the coefficient is significant at the 1% level for all four specifications. This gives strong evidence that migrants locate in areas where they expect to earn higher wages.

There is also strong evidence supporting the social network theory. As expected, the foreign-born Latin American population ratio in 2000 is positively correlated with the ratio of recent immigrant population to total population in 2005 and is significant at the 1% level. There is, however, only weak evidence of a saturation effect in these Latin
American migrant networks when examined across all specifications. This suggests that immigrants do tend to locate in areas with strong Latin American migrant networks; however, these networks may lose their effect as they can become saturated with new immigrants looking for jobs.

In addition to the two theories being tested, some of the control variables included are significant in determining where Latin American migrants locate in the U.S. I find that the rural population variable is negative and significant. This confirms that immigrants tend to move to more populated, urban locations, and also implies that many of the migrants working in the rural sectors may be seasonal migrants, and are not picked up in the U.S. Census data. Finally, the log distance variable is positive and significant at the 5% level, which is not what I would expect because it proxies for cost. There are more migrants settling far from their point of entry than near. Undocumented immigrants may move farther from the border to avoid apprehension and the risk of being sent back to their home country. Legal immigrants, however, may be repelled from the border due to the high Border Patrol surveillance and racial profiling, which I believe to be a psychological cost.

V. Alternative Specifications

To test the robustness of my Latin American specification, I move to a more general model, which accounts for all nationalities of immigrants in the independent variables. For this, I replace two Latino-specific independent variables, the measures of Hispanic income per capita and the ratio of foreign born Latin American population to total population. Instead, to proxy for the neoclassical theory, I use the per capita income in 1999 for the entire population. For the social network theory, I use the ratio of the population of all foreign-born residents to the total population of the county. As in the previous specification, I expect these variables both to be positively correlated with the dependent variable. While these two variables change, each of the four regressions in this specification includes the same variables as the previous specification. This specification will reflect immigration patterns for all nationalities of immigrants to the Southwestern U.S.

For the results of this alternative specification, see Table 2. This set of regressions is notable because income is insignificant for all four regressions. This suggests that in the previous specification, the significance of Hispanic income per capita may have been reflecting an effect that is specific to the Latin American bias in my sample. This measure could also more accurately reflect the earnings an immigrant expects when deciding where to locate than the overall population’s average earnings. Further research is needed to determine whether the neoclassical economics theory is relevant in determining where immigrants locate.

In this specification, there is evidence supporting the social capital theory with a saturation effect. The percent foreign-born population in 2000 and its square are significant at the 1% level for all four regressions. This suggests that for the immigrant population in general, migrant networks are a strong pull factor in determining where to locate. At a certain point, though, these networks become saturated and do not provide as strong of incen-
tives for locating in the area as previously.

VI. Conclusions

Using a cross-section of 88 Southwestern U.S. counties, I find strong evidence of the social capital theory in determining where migrants locate. This applies specifically to Latin American migrant networks, as well as to more general foreign-born networks. An increase in the foreign born Latin American population of a county leads to an increase in the flow of immigrants to that county in the subsequent time period. On the other hand, there is inconclusive evidence as to whether the neoclassical theory explains where immigrants locate. The Hispanic per capita income explained Latin American immigration patterns, but the more general per capita income measure did not. In further work, I would like to explore other measures to represent an immigrant’s expected wage; these may provide further evidence to determine the applicability of the neoclassical economics theory in determining settlement patterns. These results suggest that policy makers should reconsider their stance on immigration. This study shows that earnings may not be pull factor in determining where migrants locate, so decreasing economic incentives, as was previously done, may not be effective in deterring migrants from certain locations. Instead, governments should work with the migrant networks already in place in their counties to solve some of the economic problems associated with large populations of immigrants. Migrant networks are self-perpetuating, so it is important for governments to face these problems straight on, while the networks are at their smallest.

While the results found in my paper have potential implications for policymakers, they must be taken with caution. My research is limited to a small sample of data and imperfect variables. In further work, I would like to expand my data set and research the lag between when a person decides to immigrate and when he arrives in his final destination. This would allow for a more accurate representation of the effects of county specific characteristics on the subsequent period of immigration.

References


DETERMINANTS OF SOUTH AMERICAN IMMIGRANTS’ LOCATION


TABLE 1: LOCATION DETERMINANTS OF FOREIGN BORN IN U.S. IN 2005: LATIN AMERICA FOCUS

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Hispanic Per Capita Income 1999</td>
<td>0.099***</td>
<td>0.095***</td>
<td>0.079***</td>
<td>0.078***</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.015)</td>
<td>(0.018)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Ratio Pop. Born Latin Amer. 2000</td>
<td>0.212***</td>
<td>0.361***</td>
<td>0.306***</td>
<td>0.288***</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.061)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>Ratio Born Latin Amer. 2000 sq.</td>
<td>-0.547***</td>
<td>-0.440**</td>
<td>-0.193</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.203)</td>
<td>(0.202)</td>
<td>(0.239)</td>
<td></td>
</tr>
<tr>
<td>Ratio Rural Population</td>
<td></td>
<td>-0.023**</td>
<td>-0.030***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.010)</td>
<td>(0.009)</td>
<td></td>
</tr>
<tr>
<td>Log Shortest Distance to Border</td>
<td></td>
<td></td>
<td>0.011**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>AZ</td>
<td>0.011*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TX</td>
<td>0.003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.55</td>
<td>0.57</td>
<td>0.59</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Dependent variable is the ratio of the population which is foreign born in 2005 and entered the U.S. since 2000 to the total county population. All ratio variables are the stated variable to the total county population. AZ, NM, and TX are state dummy variables which equal 1 if the county lies within that state. The omitted state variable is CA.
Data Sources: 2000 U.S. Census, 2005 American Community Survey
### Table 2: Location Determinants of Foreign Born in U.S. in 2005

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Per Capita Income 1999</td>
<td>0.015</td>
<td>0.016</td>
<td>0.014</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Ratio Foreign Born Pop. 2000</td>
<td>0.168***</td>
<td>0.297***</td>
<td>0.281***</td>
<td>0.348***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.041)</td>
<td>(0.043)</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Ratio Foreign Born Pop. 2000 sq.</td>
<td></td>
<td>-0.368***</td>
<td>-0.342***</td>
<td>-0.415***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.118)</td>
<td>(0.119)</td>
<td>(0.132)</td>
</tr>
<tr>
<td>Ratio Rural Population</td>
<td></td>
<td></td>
<td>-0.007</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Log Shortest Distance to Border</td>
<td></td>
<td></td>
<td></td>
<td>0.007**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>AZ</td>
<td></td>
<td></td>
<td></td>
<td>0.013***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.005)</td>
</tr>
<tr>
<td>NM</td>
<td></td>
<td></td>
<td></td>
<td>0.013**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.006)</td>
</tr>
<tr>
<td>TX</td>
<td></td>
<td></td>
<td></td>
<td>0.009***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>Observations</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.65</td>
<td>0.68</td>
<td>0.67</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Dependent variable is the ratio of the population which is foreign born in 2005 and entered the U.S. since 2000 to the total county population. All ratio variables are the stated variable to the total county population. AZ, NM, and TX are state dummy variables which equal 1 if the county lies within that state. The omitted state variable is CA.

Data Sources: 2000 U.S. Census, 2005 American Community Survey
THE EFFECT OF MICROFINANCE INSTITUTIONS ON INCOME INEQUALITY: EVIDENCE FROM GHANA

R. Priya Mathew
Washington University

I. Introduction

Microfinance Institutions (MFIs) are small savings and loans programs often started by non-profit organizations as a tool to help the poor. There are over 10,000 MFIs around the world lending to over 30 million people (UNDCF, 2006). Can these banks live up to their lofty goals of reducing poverty? Many studies have attempted to answer this question by analyzing whether these programs reach “the poorest of the poor”. This research has found that although MFIs may help the participants of the programs, they have failed to reach the very poor (Morduch 1998, Zaman 2000). A flaw in these studies is that poverty as a concept is difficult to define, quantify, and test. These studies emphasize that relative to the other participants, the impacts for the very poor participants are quite small. Although this is likely true, the small impact for the poor may be highly valued. The more important issue to address is whether the poor participate at all in these programs, and if they do participate what are the effects of the programs on the community. Research on this subject has found that those that participate in the programs are more likely to be those who already have the drive, motivation or income advantage and are among the most successful in their communities (Pretes 2002, Coleman 2002). The majority of these studies do not include the possible spillover effect of MFIs from the participants to non-participants in their analysis. MFIs could be indirectly helping the very poor in this way (Zohir and Matin, 2000).

No study has analyzed the question of who microfinance helps by studying the income distribution effects of introducing MFIs into a community. The question of who MFIs help is important to understand as these programs become more and more popular. The programs could affect income inequality in three major ways depending on whom the programs reach. The MFI could increase inequality, decrease inequality, or not cause any change in inequality.

If previous research is right and MFIs mostly reach the already successful and impact is very small for the poor who do participate, then the community level inequality would
increase as a result of the MFI. The poor would be relatively worse off. Another possibility is that MFIs could decrease inequality. If the claims of microfinance proponents are true and MFIs mostly reach the poorest of their communities, then one would expect a decrease in inequality. A third possibility is that MFIs do not change the inequality of their communities because there is equal participation in the programs from the members of the community.

To provide a more comprehensive analysis than previous studies, I also attempt to explain the MFIs impact on income inequality by analyzing the participation and spillover effects. The participation effect is important because inequality would only be affected by MFIs if the institutions increase the incomes of participants. Otherwise, there would be no change in inequality. The spillover effect is important because, it is possible that MFIs affect inequality, but the spillover effect from participants to non-participants contributes to the magnitude of this effect. For example, if MFIs mainly help the already successful, motivated entrepreneurs, there may not be as great of a shift upward in inequality because there are spillover effects to the non-participants from the improved business climate.

In this paper I do not make a normative argument as to whether it is good or bad for MFIs to change income inequality in a community. I use inequality as a tool to analyze whether MFIs have been successful in reaching the poor as this is both their major claim for support and their major source of criticism. My null hypothesis is that after introducing the MFI there will be no change in income inequality. To further explain this result, I test the null hypotheses that MFIs do not increase the incomes of participants, and that MFIs do not have a spillover effect from participants to non-participants. I use a dataset from Freedom from Hunger’s (FFH) Credit with Education program in Ghana. I test the inequality null hypothesis by analyzing the change in Gini Coefficient in the program group versus control group. I use a regression model to test for the participation and spillover effect. My results suggest that there was no increase in inequality after the MFI was introduced, and there may have been a slight decrease. Data limitations make it difficult to determine the significance of the decrease in Gini Coefficient. Cross-country Gini Coefficient comparisons indicate that it is possible that the .06 difference in Gini Coefficients between the Program and Control groups after the MFI is introduced may be a large and substantial amount. Regression analysis provided strong evidence that the income of participants increased after the MFIs were introduced. Participation was correlated with about a 12% increase in income, and there was no evidence of a correlation between future participation and income. Thus, there was also no evidence of a self-selection bias based on income. The results also found a significant negative coefficient for the Program variable in the 1993 baseline regression indicating that the program communities were poorer than the control communities before the MFIs were introduced. In the 1996 regression, the Program variable coefficient was positive and no longer significant. These results are not conclusive, but they do suggest there may be a spillover effect.

The data from developing countries is infamously difficult for research purposes. However, given the limitations, my study has raised important issues about the current analysis of the impact of microfinance on the poor. Although the evidence is mostly sug-
gestive rather than conclusive, it provides some evidence that income inequality does not increase and may decrease after MFIs are introduced. This result indicates that it is possible that microfinance is reaching more of the poor than its critics have argued. The suggestive evidence of a spillover effect and the strong evidence that participation is correlated with higher incomes, gives hope to microfinance proponents that the programs may be working in the communities into which they are introduced. The most important contribution of this paper is that it provides a comprehensive analysis of the effect of MFIs on income inequality, taking the participation and spillover effect into account. This analysis has shown that these effects are important, and researchers and microfinance proponents should take this into account in their future work.

II. Literature Review

Although the poor could benefit greatly from savings, investment and loans, they often have no access to financial markets because they lack credit and collateral. De Soto (2000) argues that the poor are asset rich, but lack the means to convert these assets into investment and capital because of their inability to access the credit market. The problem may be that although the poor demand financial goods, there is a lack of supply for them (Hulme, 2000). MFIs propose to bridge the gap and provide the poor with small loans and savings. In 2006, Muhammad Yunus won the Nobel Peace Prize jointly with the Grameen Bank which is widely recognized as one of the first MFIs established (NYT, 2006). The Norwegian Nobel Committee (2006) believed that “Yunus and Grameen Bank have shown that even the poorest of the poor can work to bring about their own development”. The United Nations Capital Development Fund (2006) supported this claim and identified microfinance as a means to meet the Millennium Development Goals, the first of which is to halve poverty and hunger. Many resources have been channeled on the hope that microfinance can live up to its high expectations. This confidence that MFIs can drastically reduce poverty by reaching the very poor is often based on anecdotal evidence which describes stories of the poor turning into budding entrepreneurs starting businesses and lifting themselves out of poverty as a result of microfinance loans (Business Week, 2006). However, anecdotes highlight the most successful stories of participants and cannot give an overall picture of the program’s impact (Armendariz and Morduch, 2005). This is related to an incentive problem, because those providing the impact assessment often have a vested interest in the success of the MFI. It is important to gain unbiased information about the impacts of microfinance considering the great amount of money and hope that depends on the success of these programs.

The major criticism of microfinance has been that it has not reached the very poor, and that the participants of the programs who benefit the most are those who already were well-off or had the entrepreneurial spirit to succeed without the MFI loans. Hulme and Mosley (1998) found that there is an “impact frontier” where MFIs must choose between targeting the poorest of the poor and achieving a very small impact or targeting those closer to the poverty line and seeing a big impact. Zaman (2000) reports a similar result in his
study where participants that have above a threshold of loans over 10,000 taka, about 100 dollars, felt the greatest impact. Since the very poor are not able to reach this threshold they do not see as large of a positive impact as other participants in the program. Although the results of these studies are important, they all find that there are some gains for the majority of participants in the programs. A more important concern is the criticism that the very poor do not participate in these programs at all. Coleman (2002) found that the wealthiest villagers were twice as likely to participate in MFIs. In village banking systems they were more likely to get controlling program committee positions. Managers of MFIs may also have incentives to avoid the very poor since their goals are to be sustainable, and the poor are more costly to serve than those nearer to the poverty line (Murdoch and Armendariaz 2005).

The majority of these studies ignore spillover effects from participants to non-participants that could also help the poor. Spillover effects are an indirect way for MFIs to reach the very poor. Participants could spend, hire new employees, and open new businesses which would help the community in general. Zohir and Matin (2004) developed a theory that microfinance can have spillover effects to many aspects of the community such as transportation and health. Mushtaque, Chowdhury and Mosley (2004) suggest that MFIs have broad impacts like stabilizing volatile financial sectors, using derived demand to increase employment, and providing institutional inspiration.

This paper examines the issues brought up by these theories through the context of income inequality. Proponents of microfinance would like to see inequality decrease. The view of critics is that it is likely to lead to an increase in inequality because it helps the already successful. The spillover effect could affect the magnitude of the MFI’s impact on inequality by positively helping the income of both participants and non-participants.

III. Data

For my analysis, I use a dataset from Freedom from Hunger (FFH) which was conducted in Ghana during the period of 1993-1996. The MFIs introduced were based on a Credit with Education village banking model. Participants attended weekly meetings where they received business and nutritional education, paid back part of their principal and interest on loans, and deposited savings. The average loan was $60.39 and the average savings was $7.38. There were 17 communities chosen from coastal Ghana to participate in the study. The communities were then randomly assigned to be either program communities or control communities. Control communities never received MFIs during the time of research. MFIs were started in program communities as soon as possible after the 1993 baseline survey. Individuals in program communities could then choose whether to participate and take out a loan from the MFI or not. A baseline survey of 308 people was conducted in 1993, before introducing MFIs into the communities, and a follow up survey of 306 different people was conducted in 1996, after introducing the MFIs into the program communities. The women from the control and non-participant communities were randomly selected to be surveyed. In total from both surveys, there were 614 women in the sample.
The Freedom from Hunger program required different women to be used in the follow up year than were surveyed in the baseline year. Although naturally one would think that a longitudinal study surveying the same women in 1993 and 1996 would provide the best analysis, in practice, participant dropouts have proved to be a large problem in evaluating microfinance programs. This is especially relevant since many of these dropouts are the poor or unsuccessful (Hulme 2000). This dataset employs a second best strategy, and reclassifies the 1993 baseline program survey group as “future participants” and “future non-participants” based on whether they chose to participate after the MFI was established in their community. This allows researchers to compare the “future participant” group and the “participant” group. Although they are not the same people they are likely to have similar levels of entrepreneurial drive and talent.

The survey instrument asked over 300 questions on a variety of issues including business, health, attitude, and demographics. In an impact study of this data, MkNelly (1998) uses a comparison of means test to show that participants have higher incomes, better business knowledge, and better health practices than non-participants or control groups. In my analysis, I use this data to test the question of whether MFIs affect income inequality in program communities versus control communities.

IV. Methodology

My null hypothesis is that there will be no change in income inequality after the MFIs are introduced in program communities versus control communities. To further explain MFIs effect on inequality, I test two null hypotheses regarding the participation and spillover effect. The first null hypothesis is that there is no positive correlation between participation and income. The second null hypothesis which tests the spillover effect is that there is no correlation between being in a program community with an MFI available and income. To test my inequality null hypothesis, I analyze the change in Gini Coefficient, which is a popular measure of inequality that measures the distance between the Lorenz curve and the 45 degree line. A Gini Coefficient of 1 represents perfect inequality, where one person has all the income in an economy while a Gini Coefficient of 0 represents perfect equality, where each person in the economy has equal income. To reject my null hypothesis, I would expect to find significant change in Gini Coefficient in the Program group versus the Control group.

I developed a community level fixed effects linear regression model to test for both participation and the spillover effect. This model is based on the Coleman (2002) model testing participation and outcome and the Morduch (1998) model testing participation, eligibility and outcome. The following is my linear regression model:

\[ \ln (Y_{ij}) = X_{ij} \alpha + V_j \beta + M_{ij} \gamma + P_{ij} \delta + e_{ij} \]  

where \( i \) is an individual in community \( j \), \( Y \) is the outcome measured in income ($) and \( X_{ij} \) is a vector of individual level characteristics. The individual level characteristics of age
and literacy are standard to include in microfinance impact literature. The total number of live children was included in the model because this characteristic was found to be significantly different between groups in the group comparison of means test (see Table 4 and 5).

The regression is run twice, once for 1993 baseline survey data and once for 1996 follow up survey data. For the participation coefficient, I analyze the relationship between both future participation and participation on income to test for a possible self-selection bias in my data. If in the 1993 baseline regression, future participation is correlated with a large increase in income, then it is possible that those who participated in the programs were those who were already motivated or successful. If I find that the 1993 regression shows no significant relationship between future participation and income, but the 1996 regression shows a significant relationship between participation and income, then this strong evidence that it was the program that increased incomes of participants. This result would allow me to reject my null hypothesis that participation is not positively correlated with higher incomes. It would also show that there is no evidence of a self-selection bias. I test the spillover effect explanation by using a dummy variable for program availability—whether the individual was in a program community or not. The 1993 regression coefficient for the Program dummy variable tests the randomness of the community selection. The 1996 regression coefficient for the Program dummy variable tests for the spillover effect. If there is a significant positive correlation between just having an MFI available and income, this would indicate there was a spillover effect from participants to non-participants in the program community. This result would allow me to reject my null hypothesis that there is no spillover effect from the MFI.

<table>
<thead>
<tr>
<th>Definition of Variable</th>
<th>1993</th>
<th>1996</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ, Program Dummy Variable</td>
<td>Not significant</td>
<td>+, significant</td>
</tr>
<tr>
<td>γ, Participation Dummy Variable</td>
<td>Not significant</td>
<td>+, significant</td>
</tr>
<tr>
<td>β, Community Fixed Effects</td>
<td>Not significant</td>
<td>Not significant</td>
</tr>
<tr>
<td>α, Individual Characteristics (Age, Literacy, Total Number of Kids)</td>
<td>+, significant</td>
<td>+, significant</td>
</tr>
</tbody>
</table>
V. Results

The results indicate there is no increase in income inequality and possibly a slight decrease. In the 1993 survey the Gini Coefficient was the same for the Program and Control group. In 1996, there is a .06 smaller Gini Coefficient for the Program group vs. the Control group. Although this data indicates a decrease, due to data limitations it is difficult to determine whether this is a large and significant magnitude. Cross country comparison of Gini Coefficients can help provide more information on whether this change is substantial. Mukhopadhaya (2004) divides countries by the extent of inequality in the following five ranges: less than .3, .3 to .4, .4 to .45, .45 to .5, and above .5. The ranges of .4 to .45 and .45 to 5 are particularly interesting because they indicate that .05 is a significant difference in Gini Coefficient. However, it is also important that the highest range is anything above .5. Thus it is possible that on the higher and lower ends of the scale, a .05 difference in Gini Coefficient may not be as significant as in the middle of the scale. It is also important to consider how much the magnitude of a country’s Gini Coefficient changes over time. Beck, Demirgüç-Kunt and Levine (2004) created a variable for growth in Gini Coefficient in the period of 1960-1999. The countries with the largest magnitude of change are Finland with a -.018 decrease and Ecuador with a .011 increase. Thus the .03 magnitude of change found in my analysis may be large and substantial. However, because the time period of the Freedom from Hunger research was only 3 years, it is unclear whether this effect would hold in a long run analysis. Thus I am not able to reject my null hypothesis, but are suggestive that that there is no increase in inequality and possibly a decrease.

My regression analysis rejected the null hypothesis that participation was not correlated with income. I found that the participation coefficient, ã, was positive and significant in the 1996 follow-up survey, but not significant in 1993 baseline survey. This indicates that participation has a positive correlation with income, and there is no relationship between future participation and income. Thus, there is no evidence of self-selection based on income. The 1996 follow-up regression results show that participation is correlated with a .11 increase in log income, which translates into about a 12% increase in income for participants. The results of the participation coefficient are consistent with the results which suggest that income inequality slightly decreased after the MFI was introduced. The Program coefficient measuring the spillover effect was negative and significant in 1993, indicating that the Program communities were poorer than the Control communities. In 1996, the Program coefficient was positive, although not significant. Although I cannot reject my null hypothesis that there was no spillover effect, there is weak evidence that a spillover effect may have occurred which is also consistent income inequality analysis.

<table>
<thead>
<tr>
<th>TABLE 2: LINEAR REGRESSION COEFFICIENTS FOR LOG INCOME MODEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline 1993 Survey</td>
</tr>
<tr>
<td>Coefficient</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Program,</td>
</tr>
<tr>
<td>Participation</td>
</tr>
</tbody>
</table>

*Coefficient is significant at the 10% level  ** Coefficient is significant at the 1% level
VI. Conclusion

This study provides evidence that introducing a MFI into a community can increase incomes of its participants without increasing the level of income inequality within the community. There may even be a slight decrease in inequality after the MFI is introduced and positive spillover effects on non-participants. Thus, critics may have overstated the claim that microfinance is only helping the already successful and motivated in the communities. However, this is only one study with limited data and suggestive results. There are clear policy and academic implications from this research. The current focus of impact analysis from both microfinance proponents and academics is on the issue of poverty alleviation, but often fails to take into account self-selection bias, spillover effects, and inequality. This study shows the importance of these measures in providing a comprehensive analysis and explaining the impact of microfinance.

References


### Table 3: Linear Regression Coefficient for Log Income

<table>
<thead>
<tr>
<th>Community</th>
<th>Baseline 1993 Survey Coefficient</th>
<th>t</th>
<th>Follow-up 1996 Survey Coefficient</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.54**</td>
<td>31.48</td>
<td>1.37**</td>
<td>14.92</td>
</tr>
<tr>
<td>Program Participation</td>
<td>-.101*</td>
<td>-1.91</td>
<td>.038</td>
<td>0.42</td>
</tr>
<tr>
<td>Age</td>
<td>-.001</td>
<td>-.73</td>
<td>.006**</td>
<td>2.28</td>
</tr>
<tr>
<td>Live Children</td>
<td>.006</td>
<td>1.29</td>
<td>.005</td>
<td>0.49</td>
</tr>
<tr>
<td>Literacy</td>
<td>.044**</td>
<td>2.95</td>
<td>.071**</td>
<td>2.56</td>
</tr>
<tr>
<td>Community 2</td>
<td>.076</td>
<td>1.47</td>
<td>-.009</td>
<td>-0.09</td>
</tr>
<tr>
<td>Community 3</td>
<td>-.099</td>
<td>-1.89</td>
<td>.060</td>
<td>0.61</td>
</tr>
<tr>
<td>Community 4</td>
<td>.081</td>
<td>1.58</td>
<td>-.031</td>
<td>-0.40</td>
</tr>
<tr>
<td>Community 5</td>
<td>.019</td>
<td>.43</td>
<td>.049</td>
<td>0.63</td>
</tr>
<tr>
<td>Community 6</td>
<td>.026</td>
<td>.50</td>
<td>-.009</td>
<td>-0.11</td>
</tr>
<tr>
<td>Community 7</td>
<td>.074</td>
<td>1.74</td>
<td>-.014</td>
<td>0.21</td>
</tr>
<tr>
<td>Community 8</td>
<td>.052</td>
<td>1.01</td>
<td>.180</td>
<td>1.91</td>
</tr>
<tr>
<td>Community 9</td>
<td>.051</td>
<td>1.43</td>
<td>-.128</td>
<td>-1.33</td>
</tr>
<tr>
<td>Community 10</td>
<td>.086</td>
<td>1.68</td>
<td>.086</td>
<td>1.10</td>
</tr>
<tr>
<td>Community 11 (dropped)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community 12</td>
<td>-.044</td>
<td>-.86</td>
<td>.020</td>
<td>0.21</td>
</tr>
<tr>
<td>Community 13</td>
<td>.118**</td>
<td>2.72</td>
<td>.051</td>
<td>0.77</td>
</tr>
<tr>
<td>Community 14</td>
<td>.093*</td>
<td>2.16</td>
<td>-.023</td>
<td>-0.30</td>
</tr>
<tr>
<td>Community 15</td>
<td>-.003</td>
<td>-.08</td>
<td>.072</td>
<td>0.91</td>
</tr>
<tr>
<td>Community 18</td>
<td>.084*</td>
<td>1.96</td>
<td>.042</td>
<td>0.50</td>
</tr>
<tr>
<td>Community 19</td>
<td>.080</td>
<td>1.84</td>
<td>.013</td>
<td>0.17</td>
</tr>
</tbody>
</table>

*significant at 10% level **significant at 1%
### Table 4: Group Mean Comparison T-Tests:
Baseline 1993, Mean (and Standard Deviation)

<table>
<thead>
<tr>
<th></th>
<th>Future Participants N = 48</th>
<th>Future Non-Participants N = 161</th>
<th>Control N = 99</th>
<th>Significance (p&lt;.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Family Members</td>
<td>6.4 (2.2)</td>
<td>6.3 (2.4)</td>
<td>6.7 (2.9)</td>
<td>None</td>
</tr>
<tr>
<td>Dollar Value of Selected Assets</td>
<td>105.1 (214.0)</td>
<td>177.4 (719.1)</td>
<td>230.2 (740.6)</td>
<td>None</td>
</tr>
<tr>
<td>Individual’s Age</td>
<td>29.9 (7.5)</td>
<td>27.7 (6.4)</td>
<td>27.7 (6.1)</td>
<td>Future Participants vs. Control, Future Participants vs. Future Non-Participants</td>
</tr>
<tr>
<td>Individual’s Height</td>
<td>156.2 (4.3)</td>
<td>156.1 (5.2)</td>
<td>156.0 (5.2)</td>
<td>None</td>
</tr>
<tr>
<td>Dollar Value of Monthly Profit</td>
<td>5.9 (7.3)</td>
<td>4.2 (7.2)</td>
<td>5.3 (7.3)</td>
<td>None</td>
</tr>
<tr>
<td>Total of Live Children</td>
<td>3.5 (2.3)</td>
<td>3.0 (1.8)</td>
<td>2.9 (1.7)</td>
<td>Future Participants vs. Control, Future Participants vs. Future Non-Participants</td>
</tr>
<tr>
<td>% Married</td>
<td>85%</td>
<td>88%</td>
<td>81%</td>
<td>None</td>
</tr>
<tr>
<td>% Having Over Primary Education</td>
<td>48%</td>
<td>40%</td>
<td>34%</td>
<td>None</td>
</tr>
<tr>
<td>% Literate</td>
<td>42%</td>
<td>31%</td>
<td>32%</td>
<td>None</td>
</tr>
</tbody>
</table>

### Table 5: Group Mean Comparison T-Tests:
Follow-up 1996, Mean (and Standard Deviation)

<table>
<thead>
<tr>
<th></th>
<th>Future Participants N = 48</th>
<th>Future Non-Participants N = 161</th>
<th>Control N = 99</th>
<th>Significance (p&lt;.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Family Members</td>
<td>6.4 (2.2)</td>
<td>6.3 (2.4)</td>
<td>6.7 (2.9)</td>
<td>None</td>
</tr>
<tr>
<td>Dollar Value of Selected Assets</td>
<td>105.1 (214.0)</td>
<td>177.4 (719.1)</td>
<td>230.2 (740.6)</td>
<td>None</td>
</tr>
<tr>
<td>Individual’s Age</td>
<td>29.9 (7.5)</td>
<td>27.7 (6.4)</td>
<td>27.7 (6.1)</td>
<td>Participants vs. Control, Participants vs. Non-Participants</td>
</tr>
<tr>
<td>Individual’s Height</td>
<td>156.2 (4.3)</td>
<td>156.1 (5.2)</td>
<td>156.0 (5.2)</td>
<td>None</td>
</tr>
<tr>
<td>Dollar Value of Monthly Profit</td>
<td>5.9 (7.3)</td>
<td>4.2 (7.2)</td>
<td>5.3 (7.3)</td>
<td>None</td>
</tr>
<tr>
<td>Total of Live Children</td>
<td>3.5 (2.3)</td>
<td>3.0 (1.8)</td>
<td>2.9 (1.7)</td>
<td>Participants vs. Control, Participants vs. Non-Participants</td>
</tr>
<tr>
<td>% Married</td>
<td>85%</td>
<td>88%</td>
<td>81%</td>
<td>None</td>
</tr>
<tr>
<td>% Having Over Primary Education</td>
<td>48%</td>
<td>40%</td>
<td>34%</td>
<td>None</td>
</tr>
<tr>
<td>% Literate</td>
<td>42%</td>
<td>31%</td>
<td>32%</td>
<td>None</td>
</tr>
</tbody>
</table>
WHAT IS THE EFFECT OF THE ABUNDANCE OF NATURAL RESOURCES ON INCOME DISTRIBUTION?

Andrew O’Brien-Penney
Georgetown University

I. Introduction

Why have resource-poor developing countries like the newly industrializing economies (NIEs) of East and Southeast Asia, which had similar GDPs per capita to that of developing countries in Sub-Saharan Africa and South America fifty years ago, dramatically outperformed their resource-rich counterparts? How have these countries also been able to achieve more equitable distributions of income than their resource-abundant counterparts? For the sake of slowly developing countries and their economic well-being, we must ask whether natural resources play a role in determining the distribution of income within a country and subsequently the GDP growth of the country.

II. Data Sources, Model, and Strategy

The dependent variable of this study, income distribution, will be denoted by the Gini Coefficient in my main regression and by the percentage of income earned by the poorest 10%, poorest 20%, richest 20%, and richest 10% of the population in each of the auxiliary regressions. I will include data for all countries (for which data is available) except developed countries (i.e. the US, Canada, Australia, New Zealand, Japan, and Western Europe). The reason why such countries are not included is because I seek to show the correlation between resource abundance and income distribution only in developing countries. The data for income distribution is the most recent available according to the 2006 United Nations Human Development Report, based on World Bank figures.

Much of the data for each of the independent variables exist for a long historical time series; however, the data pertaining to income distribution does not. Since countries conduct income surveys at varying times I will assume that the distribution of income did not change substantially from the survey date to the date for all data for the explanatory variables. The income survey information falls within the 10-year range of 1993-2003. To achieve a proper time correlation between my regressors and regressands, I have chosen to use data for the independent variables from the year 2000, because this is the weighted
average of the income survey years for the 100 remaining observations. Through this achieved time correlation, my data set takes on a cross-sectional aspect, allowing me to regress our determinants of income distribution against the measurements thereof. By using such cross-sectional data, I use the ordinary least squares (OLS) method, as described by (Odedokun 2001). Also, because of the likelihood of heteroskedasticity existing in such sample data, I will employ the use of a robust coefficient-covariance matrix (White 1980).

I will regress these dependent variables against variables others have used to explain income distribution, specifically GDP per capita (a proxy for labor productivity), GDP growth, the ratio of government expenditures to GDP, and the ratio of government transfers to GDP.

The data for GDP per capita (pc) are based on data from the World Bank World Development Indicators Database (WDI). Since this variable is not on a percentage basis, I must use the log of GDP per capita (g). Some may believe that because of the possible endogenous effect of GDP growth on income distribution and since GDP growth should influence income distribution in the next period, one should use lagged data (in my case 1999 data for GDP growth). However, because GDP growth is by definition the change in GDP from one period to the next, the lagged effect is already present in the data. Also, my decision to use 2000 data for GDP growth instead of that from 1999 is supported by other researchers using income distribution to determine GDP growth for the same period, as in Odedokun and Round (2001). The ratio of government expenditures to GDP (G) is a proxy for the size of the government as discussed in Odedokun and Round (2001). Also, Odedokun and Round (2001) used the variables for both the ratio of government subsidies to total government expenditures (s) and the ratio of government subsidies to GDP (S). Data for these two variables are scarce for the year 2000 specifically, so I have used data from both the year 2000 and for the closest year to 2000 for which data are available.

Along with the above variables, I will include several independent variables pertinent to the discussion of the impact of natural resources on income distribution. First, when considering the influence of agricultural capacity on income distribution, I will first use a variable for the amount of arable land per person (a), measured in the log of hectares per person (obtained from WDI). The log is necessary to allow for a direct correlation between hectares per person, which is not in percentage format, with the measures of income distribution, which are. I will also employ a variable for the ratio of agricultural raw materials exports to total merchandise exports (ae), and a variable for the ratio of agricultural raw materials exports to GDP (AE), both based on figures from the United Nations Conference on Trade and Development (UNCTAD) Handbook of Statistics.

Next, we must account for the impact, if any exists, of the abundance of point natural resources on income distribution. Since such resources utilize little area for extraction, a variable denoting the amount of land available for such enterprises would be irrelevant. I will employ two variables: first, a variable for the ratio of point resource exports (all fuels, ores, and metals) to total merchandise exports (pe) and a variable for the ratio of point
Thus, the main equation to examine the effects that the proposed determinants of income distribution have on income distribution is as follows:

\[ \text{Gini} = \beta_0 + \beta_1 \log(pc) + \beta_3 g + \beta_4 G + \beta_5 s + \beta_7 \log(a) + \beta_8 aE + \beta_9 AE + \beta_{10} pe + \beta_{11} PE + e \]

where the Gini coefficient is the dependent variable in this main regression. I will also run auxiliary regressions in which the independent variables will remain the same and the dependent variable will be the percentage of income earned by the poorest 10%, the poorest 20%, the richest 20% and the richest 10%.

The hypotheses for each of these regressions will be that each of the determinants of income distribution has a statistically significant impact on income distribution. My null hypothesis for each of the determinants is that each of their coefficients will be 0. I will determine whether I can reject the null in favor of the alternative that each coefficient is significant. It is of particular importance to this paper is to show the statistical significance or insignificance of each of the natural resource abundance determinants of income distribution.

I will regress all of the independent variables against each of the metrics of income distribution. I will also regress all of the natural resource determinants of income distribution against each of the metrics of income distribution. Also, I will reduce the number of independent variables in each regression to the point at which all of them are statistically significant at or near the 5% level; finally, I will perform a robustness test in which I compare the overall significance and explanatory power (through the use of identical observations) of 1) the regression with only those regressors which are statistically significant individually to 2) the regression which includes all of the possible regressors

\[ III. \text{ Results} \]

First, I regressed all of the determinants of income distribution against the Gini coefficient, examined the magnitude and sign of the coefficients of the regressors, checked each of their p-values (based on robust standard errors) to determine each of their individual significance, and checked the overall significance of the regression. The results, show that \( \log(pc) \), \( g \), \( G \), and \( S \) are all statistically significant at or near the 5% level; whereas, none of the natural resource determinants of income distribution are individually statistically significant. The F statistic of 4.40 and the Prob > F statistic of 0.0002 indicate that this equation is significant overall. Finally, the R-squared of the equation shows strong explanatory power—the regressors account for 38.25% of the variation in the Gini coefficient.

I also wanted to determine the joint significance of several subsets of the regressors. I performed a Wald test on all of the income resource abundance determinants of income distribution and noticed that little explanatory power exists for this set of regressors as a whole (Wald statistic=0.73, prob > F of 0.6014). I also performed a Wald test on the each
of two sets of natural resource determinants of income distribution since the two variables ae and AE and the two variables pe and PE may exhibit multicollinearity. I performed a Wald test on each of these two sets of variables separately, including the other metric of natural resource abundance (a) in each test. I found that there exists no joint significance for the three variables a, ae, and pe (Wald statistic=0.51, prob > F of 0.6786). I also found that even weaker evidence exists for the joint significance of a, AE, and PE (Wald statistic=0.38, prob > F of 0.7709).

Then, I regressed only the proposed natural resource abundance determinants of income distribution against the Gini coefficient. I found an overall F-statistic of only 0.62, an equally statistically insignificant prob > F of 0.6825, and almost no explanatory power in the R-squared of 0.0224. Also, none of the regressors have any individual statistical significance.

I returned to my main regression, determined the regressor with the coefficient with the highest corresponding p-value, eliminated that regressor, and repeated this process until all of the remaining variables were statistically significant at or near the 5% level. Through this iterated deletion of statistically insignificant variables, I determined an equation with the remaining regressors: log(pc), g, and S. All of these variables have been shown to be significant beyond even the 1% level. Observing the coefficients of the variables in the last regression, all of the remaining coefficients have similar magnitudes and the same sign as in the main regression.

I continued in the same manner, regressing my independent variables against each income bracket (poorest 10%, poorest 20%, richest 20%, richest 10%). For the poorest 10%, I obtained similar results to those for the Gini coefficient. However, since a decrease in the Gini coefficient would imply an increase in the income share of the poorest 10%, we would expect to have coefficients of the variables with opposite signs to those in the equation against the Gini coefficient. The results show that I obtained such signs on the coefficients but that many of the variables in the main regression are statistically insignificant, just like in the equation with the Gini coefficient. Thus, I performed the same three Wald tests on the metrics of natural resource abundance, and noticed that none of the three sets of variables are jointly significant. To show any direct correlation between the metrics of natural resources and the income share of the poorest 10%, I regressed these and discover that there is no overall significance to this regression. I also performed the Wald tests on the two sets of natural resource abundance metrics and determined that neither had joint significance. Finally, I reduced the number of regressors until all of the remaining regressors were statistically significant, and discovered that four of the regressors in the main regression, log(pc), g, S, and ae, remained.

Examining the main regression of the independent variables against the income share of the poorest 20%, I found that the correlations between the determinants of income distribution and income distribution to be the same in sign and of comparable magnitude to those of the poorest 10%. Again, all of the measures of natural resource abundance are statistically insignificant along with s. Thus, I performed the same three Wald tests on the metrics of natural resource abundance, and noticed that none of the three sets of variables
were jointly significant. To show any direct correlation between the metrics of natural resources and the income share of the poorest 20%, I regressed these and discover that there is no overall significance to this regression. I also performed Wald tests on this regression to determine whether the two sets of natural resource abundance determinants of income distribution (a, ae, and pe vs. a, AE, and PE) were jointly significant in this auxiliary regression, which they were not. Finally, I reduced the number of regressors of the main equation until all of the remaining regressors were statistically significant, and discovered that four of the regressors in the main regression, log(pc), g, S, and ae remained. These four regressors are exactly the same as those which remained when examining the income share of the poorest 10%. Also, the coefficients of these regressors are similar in magnitude and same in sign as their counterparts in Equation 1.

When observing the coefficients of the regressors of the main regression in the regression of income share of the richest 20%, I noticed that their signs are opposite those in the previous two sets of regressions, except for the variable pe. This fits the logic that if some variable is associated with an increase in income share of a lower income bracket, the same variable will be correlated with a decrease in income share of a higher income bracket. Again, all of the measures of natural resource abundance were statistically insignificant along with s and now log(pc). I performed the three Wald tests on the metrics of natural resource abundance, and noticed that none of the three sets of variables were jointly significant. To show any direct correlation between the metrics of natural resources and the income share of the richest 20%, I regressed these and discovered that there is no overall significance to this regression. I also performed Wald tests on this regression to determine whether the two sets of natural resource abundance determinants of income distribution were jointly significant in this auxiliary regression, which they were not. Finally, I reduced the number of regressors of the main equation until all of the remaining regressors were statistically significant, and discovered that only two of the original regressors remain: G and S.

Finally, I regressed the determinants of income distribution against the income share of the richest 10%. I noticed that the signs of the coefficients were the same as those when I regressed against the richest 20% except for the variables AE and pe (they were also not statistically significant). Again, all of the measures of natural resource abundance are statistically insignificant, but now log(pc), g, and s are also insignificant. Then, I performed the three Wald tests on the metrics of natural resource abundance, and noticed that none of the three sets of variables were jointly significant. To show any direct correlation between the metrics of natural resources and the income share of the richest 20%, I regressed these and discovered that there is no overall significance to this regression. I also performed Wald tests on this regression to determine whether the two sets of natural resource abundance determinants of income distribution were jointly significant in this auxiliary regression, which they were not. Finally, I reduced the number of regressors of the main equation until all of the remaining regressors were statistically significant, and discovered that only two of the original regressors remain: G and S, just like when I examined the income share of the richest 20%.
IV. Conclusions

We observe four different outcomes from the performed regressions. First, that the determinants of income distribution vary depending on the metric or income bracket used to examine income distribution. Second, that $S$ is statistically significant in all regressions and correlates logically with each measurement of income distribution. Third, that $\log(pc)$ and $ae$ are inversely correlated and $g$ is directly correlated with the income share of the poorest income brackets; however, none of these three variables affect the income share of the richest income brackets, but $G$ does and is directly correlated. Fourth, that the many metrics of natural resource abundance ($\log(a)$, $AE$, $pe$, and $PE$) have no statistically significant effect on income distribution. In sum, the statistically significant variables are as follows:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect on Poorest 10 &amp; 20%</th>
<th>Effect on Richest 10 &amp; 20%</th>
<th>Effect on Gini Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\log(pc)$</td>
<td>Decrease</td>
<td>No Effect</td>
<td>Increase</td>
</tr>
<tr>
<td>$g$</td>
<td>Increase</td>
<td>No Effect</td>
<td>Decrease</td>
</tr>
<tr>
<td>$G$</td>
<td>No Effect</td>
<td>Increase</td>
<td>No Effect</td>
</tr>
<tr>
<td>$S$</td>
<td>Increase</td>
<td>Decrease</td>
<td>Decrease</td>
</tr>
<tr>
<td>$ae$</td>
<td>Decrease</td>
<td>No Effect</td>
<td>No Effect</td>
</tr>
</tbody>
</table>

This paper demonstrates the overall lack of correlation between natural resource abundance and income distribution. This result flies in the face of many other papers which have tried to prove that such correlation exists. Many previous papers have focused on regions or specific countries, whereas this paper tackles all developing countries and demonstrates that no correlation exists among any of the metrics of natural resource abundance and the metrics of income distribution (except for the relationship between $ae$ and the income share of the poorest brackets).
THE CONFLICTING EFFECTS OF FISCAL AND POLITICAL DECENTRALIZATION ON CORRUPTION: A CROSS COUNTRY STUDY

Jennifer Xi
Dartmouth College

ABSTRACT

Previous theoretical research and empirical studies make ambiguous predictions about the relationship between decentralization and corruption. In part, the uncertainty lies in the fundamentally different concepts of fiscal and political decentralization. This paper disaggregates the effects of fiscal and political decentralization on corruption using new measures of political decentralization. OLS estimations in a panel study of 44 nations across the period 1990-2000 provide evidence that political and fiscal decentralization have contrasting effects on corruption. Political decentralization, as manifested in state-level elections, increases corruption in a given country. Conversely, expenditure and revenue decentralization decrease corruption. The results are highly significant and robust to a variety of controls. The findings, however, are far from conclusive because the mechanisms that contribute to these outcomes are theoretically unclear. As a result, this paper refrains from offering decisive policy prescriptions.

I. Introduction

The merits of decentralization have been contested fiercely in recent decades. Those in favor of decentralization argue that local provision of public goods better fulfill diverse individual preferences, competition between localities makes provision of services more efficient, and reduced distance between the government and the governed increases accountability. Other theorists claim decentralization worsens corruption because local
governments are more susceptible to capture by local elites, and competition between
regions can generate negative externalities. Thus, the effect of decentralization on corrup-
tion is ambiguous from a theoretical standpoint. Empirical results from cross-country
econometric studies also exhibit mixed findings.

This ambiguity is due in part to the separation between fiscal and political decentral-
ization in current literature. Fiscal decentralization refers to the devolution of financial
responsibility e.g. tax collection and public service expenditure. In contrast, political
decentralization pertains to the election of local officials by local residents and increased
authority for local governments (Bardhan and Mookherjee, 2005). Most empirical studies
focus on fiscal decentralization because it is more straightforward to measure.

I attempt to remedy this gap in the literature with an empirical study on the effects of
fiscal and political decentralization on corruption. Using data across 44 countries from
1990-2000, I find that fiscal decentralization decreases corruption whereas political decen-
tralization increases corruption. In Section II, I present a general overview of the theory
and past studies. Section III describes the data set, and Section IV reports estimation
results and robustness tests. I conclude in Section V with some policy implications and
warnings as well as questions for further research.

II. Fiscal and Political Decentralization: Theory and Overview

The distinctive feature of a decentralized government is the division of responsibilities
between a central government and lower unit governments. Corruption is the abuse of
entrusted power for private gain (Shordt, 2006). The theoretical relationship between
decentralization and corruption is complex and inconclusive because decentralization
affects corruption through the two accountability mechanisms of interjurisdictional com-
petition (exit) and local democracy (voice) (Bardhan and Mookherjee, 2005).

Because factors of production are mobile between localities, jurisdictions compete to
attract investment with lower levels of corruption. If a region is corrupt or does not pro-
vide adequate infrastructure, a firm can “exit,” or relocate, to a different region. By con-
trast, central governments enjoy monopoly power over regulations and bribes, which
increase corruption because “exit” entails relocation to foreign destinations (Bardhan and
Mookherjee, 2005). Theory also predicts that decentralization increases corruption.

Local governments are also accountable to immobile constituents who hold local gov-
ernments accountable through “voice,” or local elections. The smaller the jurisdiction, the
more closely citizens can monitor elected officials. This theory predicts political decen-
tralization through local democracy will reduce corruption. On the other hand, small
groups have greater organizational power and are more prone to advance the interests of
the majority at the expense of the minority. Similarly, local governments may be more sus-
ceptible to capture by local elites and initial rent-holders. Political decentralization may also create vertical tiers of power; each tier can extract rents in a manner similar to the markup above monopoly price in a system of successive monopolies (Fisman and Gatti, 2002). The effects of political decentralization on corruption are thus ambiguous as well.

Fisman and Gatti (2002) find that fiscal decentralization has a strong negative relationship with corruption. In the following sections, I use the Fisman and Gatti (2002) test as a baseline for my analysis.

III. Data Description

The data set used in this paper is from 44 countries across the period 1990-2000. The measure of corruption is the International Country Risk Guide’s (ICRG) corruption index. The ICRG codes the index from 0-6 with 0 as the most corrupt. I recode the index for 0 to be the least corrupt and 6 to be the most corrupt.

A common measure of fiscal decentralization is the sub-national share of total government spending as used by Fisman and Gatti (2000). This measure of fiscal decentralization (Decentral_Exp) is the amount of sub-national expenditure over total government expenditure. As stated above, theoretical predictions regarding the effect of fiscal decentralization on corruption are ambiguous.

Expenditure decentralization does not account for the source of government revenue. Expense funds can be produced from local tax collection or from transfers from the central government. As a result, I use sub-national revenue (Decentral_Rev) as an alternative measure of fiscal decentralization. Sub-national revenue excludes grants from the central government and only measures locally generated funds as a percent of total government revenue. Revenue decentralization should be inversely related to corruption because governments are subject to more constituent oversight for taxes and revenues generated locally than from the central government. Revenue decentralization overlaps with political decentralization because the electorate usually awards powers of taxation at local levels of government. Therefore, I use expenditure and not revenue decentralization as the primary measure of fiscal decentralization. I do not use the two variables together because a high correlation exists between the two, as shown in Table 1. Revenue decentralization is used in additional tests for robustness.

I use a set of political decentralization variables from Phillip Keefer’s Database of Political Institutions (DPI). The first measure of political decentralization in the DPI is a contiguous autonomous region (Autonomous). A contiguous autonomous region assumes these regions are more likely to provide a check on the central government because they are close to the home county. A second measure of political decentralization is the occurrence of municipal government elections (Municipal Elections). This scale ranges from 0 to 2, coded as 2 if both the municipal executive and legislature are elected, 1 if the legislature is elected but the executive is appointed, and 0 if neither executive nor legislature is elected. The DPI codes the third measure, state government elections (State Elections), in the same manner as municipal elections from 0 to 2. The fourth measure (Authority) is a
dummy variable equal to 1 if regions have authority over taxation, expenditure, or legislation, and 0 otherwise. I do not use the fifth measure of political decentralization, senate constituencies, because the data is only available for a few countries.

A correlation matrix between the six measures of decentralization in Table 1 reveals that Decentral_Exp and Decentral_Rev are highly correlated as expected. Autonomous, Municipal, State, and Authority are also highly correlated. Because State Elections has the most observations (504 observations), I drop Municipal and Authority (291 and 249 observations, respectively) and only use Autonomous and State in the reported regressions.

I include a number of additional controls to minimize omitted variable bias. The regressions account for country and government size using log of country GDP, log of population, and total government expenditure as a percent of GDP. I use aggregate GDP in constant 2000 US dollars in the base specification in order to account for country and economy size rather than the level of development. Large countries can exploit economies of scale in the provision of public goods, which leaves fewer public offices per capita (Fisman and Gatti, 2002). Large countries are also more likely to decentralize to better meet the needs of diverse constituents (Tiebout, 1956). The predicted effects of the country size controls are ambiguous.

I also include trade openness (Openness), measured as exports and imports as a fraction of total GDP, and GDP growth as a percentage. Theory predicts a negative relationship between these two controls and corruption. Trade-dependent and high growth regions are less likely to protect local elites and insolvent firms because the opportunity cost of doing so is higher (Bardhan and Mookherjee, 2005). Although not used in previous papers, I also include GDP per capita in a robustness test to see if any of the variables are simply a proxy for wealthy countries, which tend to have less corruption.

To account for the gap between local expenditure and local revenue generation I created the variable Local Gap. This variable is measured as local expenditure minus local revenue divided by GDP. Local Gap measures the amount the national government grants or transfers to local governments. Because local governments are not accountable to their constituents with grant revenue in the same manner as tax revenue, theory predicts a positive relationship between Local Gap and corruption.

Ethnolinguistic fractionalization (ELF) is a time-invariant variable commonly used in the economics literature. ELF increases the probability of decentralization, as different ethnolinguistic groups are likely to have different preferences. In most specifications, I include regional and legal origin dummies to account for additional time invariant characteristics.

**IV. Empirical Results**

**OLS Estimation**

The basic specification is

\[
\text{Corruption}_{i,t} = \alpha + \beta_1 \text{Decentral}_\text{Exp}_{i,t} + \beta_2 \text{Autonomous}_i + \beta_3 \text{State}_{i,t} + \beta_4 \ln(\text{GDP})_{i,t} + \epsilon
\]
$\beta_5 \ln(\text{Population})_{i,t} + \beta_6 \text{Gov}\_\text{Exp}_{i,t} + \epsilon_{i,t}$

Where $\alpha$ is the constant and $\epsilon_{i,t}$ is the error term. The first specification in Table 2 reports the coefficients from the basic specification without regional or legal origin dummies. From this rough estimation, the results show a strong negative relationship between decentralized expenditure and corruption. Conversely, the results show a strong positive relationship between state elections and corruption. The regression also estimates a highly significant relationship between corruption and log GDP and percentage of government expenditure, which suggests that countries with higher incomes and larger governments have lower levels of corruption. Larger populations are correlated with higher levels of corruption. These relationships are tenuous at best; GDP, government expenditure, and population all lose significance in later specifications when additional control and dummy variables are added.

Columns (2) through (5) report regression results with regional and legal origin dummies. The inclusion of dummies and additional variables reduces the number of observations to 250. However, the explanatory power of the regression rises to over 70 percent as demonstrated by the increase in $R^2$.

Regression results in column (2) through (5) reveal that expenditure decentralization is negative and highly significant even after controlling for a wide range of dummy variables. The state election variable is positive and highly significant across specifications as well. The correlations between decentralization and corruption are not spurious; the magnitudes of the decentralization coefficients remain fairly constant throughout the specifications. Congruent with theoretical predictions, the net effect of decentralization on corruption is ambiguous. This estimation builds on theory by decomposing the effects of fiscal and political decentralization. These empirical results show that fiscal decentralization reduces corruption whereas political decentralization increases corruption.

The gap between local government expenditure and revenue generation is positive and highly significant in specification (5), which is consistent with theory. With fewer local taxes to pay, constituents monitor their representatives less carefully than they would if local tax revenues paid for public services. Local governments with money from the central government are also more likely to attract investment through corrupt means, such as tax evasion, because they are not dependent on local tax revenue. Regulatory evasion results in lower overall tax revenue and welfare. By extension, decentralized revenue generation would, in theory, lower corruption. I address this possibility in the next section.

The presence of an autonomous region does not have an effect on corruption in any regression. An autonomous region is not directly linked to the central government, so its effects on corruption, if any, likely work through different mechanisms than those between federalism and corruption. Country size variables (log GDP, log population, and government expenditure) all lose statistical significance once GDP growth is added to the regression in specification (3). GDP growth is marginally significant in specifications (3) and (4), but loses significance once ELF is added in regression (5). The variable ELF itself is not significant in any regression as a time invariant variable. These results show that gov-
government structure is more determinative of a country’s level of corruption than country size, growth, or openness.

**Robustness Tests**

To account for the effects of revenue decentralization on corruption, I replace Decentral_Exp with Decentral_Rev in the first two columns of Table 3. Once dummy variables are included, only revenue decentralization and state elections remain significant. State drops down to a 6% significance level in column (2). This likely is because state elections and revenue decentralization both account for authority at the local level. Revenue generation comes with powers of taxation, which are approved by the electorate. Notably, Local Gap loses significance once Decentral_Rev is added to the regression. This result suggests that Local Gap was a proxy for revenue decentralization in previous regressions. Although State decreases in significance, Decentral_Rev remains statistically significant at the 1% level and in fact increases in magnitude, further confirming the hypothesis that fiscal decentralization reduces corruption.

Another problem is measurement error due to the subjectivity of the corruption index. To control for this possibility, I add GDP per capita in Column (3) of Table 4. Country income is not significant, but expenditure decentralization, state elections, and local revenue gap all remain highly significant. I do not include GDP per capita in the base specifications of Table 3 due to potential redundancy between log GDP, log population, and GDP per capita.

Despite corrections for serial autocorrelation in the previous regressions using country clusters, autocorrelation may still be a problem. To further correct for autocorrelation, I use averages from the period 1990-1995 and 1996-2000. The results in the last two columns of Table 4 show that fiscal decentralization and state elections remain highly significant.

**V. Conclusions**

Using a panel data set of 44 countries from 1990-2000, I identify two separate and contrasting effects of decentralization on corruption. First, expenditure and revenue decentralization decrease the level of corruption in a given country. Second, local election of state legislatures and executives increase corruption. These results hold through the addition of several control variables and robustness tests.

The findings indicate that interjurisdictional competition, measured by fiscal decentralization, limits bureaucratic ability to extract rents in exchange for services and induces governments to reduce corruption in order to attract investment. The net effect of “exit” on corruption therefore appears to be negative. The net effect of “voice,” or political decentralization, on corruption is positive. The creation of a local or state government may introduce an additional bureaucratic layer individuals must bribe to receive services. Corruption also may increase because local governments to bail out pre-existing firms in order to salvage jobs, and by extension, votes.

These results, however, are only an initial assessment of the relationship between fis-
cal and political decentralization and corruption. This study is constrained by several factors. The limited data set excludes a number of significant countries such as China and Russia. As a result, the regressions do not include valuable data on countries currently in the process of decentralizing their fiscal, political and judicial systems.

An additional problem is the prospect of endogeneity. Corrupt officials may curtail decentralization in order to preserve their ability to extract rents at the central level. Fisman and Gatti (2002) attempt to instrument decentralization with legal origin, but this instrument is highly suspect as legal origin likely influences both decentralization and corruption through other direct avenues. This panel data set corrects some problems associated with endogeneity. Nonetheless, these results do not suggest a causal relationship; they simply suggest that a relationship exists.

In spite of these shortcomings, this study contributes to past empirical studies on the effect of decentralization on corruption by disaggregating the effects of fiscal and political decentralization. The results provide caution for policymakers who solely advocate for decentralization and a federal form of government.

References


## Table 1: Correlations Among Decentralization Variables

<table>
<thead>
<tr>
<th></th>
<th>Decentral_Exp</th>
<th>Decentral_Rev</th>
<th>Autonomous</th>
<th>Municipal</th>
<th>State</th>
<th>Authority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decentral_Exp</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decentral_Rev</td>
<td>0.9753</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autonomous</td>
<td>-0.2640</td>
<td>-0.1067</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Municipal</td>
<td>0.1996</td>
<td>0.2932</td>
<td>0.1762</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State</td>
<td>0.4871</td>
<td>0.5437</td>
<td>0.3051</td>
<td>0.6180</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>Authority</td>
<td>0.4815</td>
<td>0.5357</td>
<td>0.3070</td>
<td>0.4894</td>
<td>0.7197</td>
<td></td>
</tr>
</tbody>
</table>

## Table 2: OLS Estimation

<table>
<thead>
<tr>
<th>Corruption</th>
<th>No Dummies</th>
<th>With Legal Origin and Regional Dummies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Decentral_Exp</td>
<td>-2.134***</td>
<td>-1.239**</td>
</tr>
<tr>
<td></td>
<td>(0.615)</td>
<td>(0.636)</td>
</tr>
<tr>
<td>Autonomous</td>
<td>-0.194</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td>(0.330)</td>
<td>(0.335)</td>
</tr>
<tr>
<td>State</td>
<td>0.330***</td>
<td>0.348***</td>
</tr>
<tr>
<td></td>
<td>(0.125)</td>
<td>(0.100)</td>
</tr>
<tr>
<td>Log GDP</td>
<td>-0.480***</td>
<td>-0.220</td>
</tr>
<tr>
<td></td>
<td>(0.078)</td>
<td>(0.219)</td>
</tr>
<tr>
<td>Log Population</td>
<td>0.617***</td>
<td>0.341*</td>
</tr>
<tr>
<td></td>
<td>(0.0643)</td>
<td>(0.194)</td>
</tr>
<tr>
<td>% Gov_Exp</td>
<td>-0.047***</td>
<td>-0.043**</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>0.054*</td>
<td>0.057*</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Openness</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Local Gap</td>
<td>32040.39</td>
<td>124918.6***</td>
</tr>
<tr>
<td></td>
<td>(35917.03)</td>
<td>(43145.62)</td>
</tr>
<tr>
<td>ELF</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>349</td>
<td>334</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.529</td>
<td>0.6255</td>
</tr>
</tbody>
</table>

Robust standard errors are in parentheses, adjusted for clustering by country. *** denotes significance at the 1 percent level; ** denotes significance at the 5 percent level; * denotes significance at the 10 percent level. Regional and legal origin dummy coefficients are omitted because none are significant.
<table>
<thead>
<tr>
<th>Corruption</th>
<th>No Dummies (1)</th>
<th>With Dummies (2)</th>
<th>5-year Averages with Dummies (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decentral_Rev</td>
<td>-2.148***</td>
<td>-2.468***</td>
<td>-2.368***</td>
</tr>
<tr>
<td></td>
<td>(0.697)</td>
<td>(0.728)</td>
<td>(1.073)</td>
</tr>
<tr>
<td>Decentral_Exp</td>
<td></td>
<td>-2.452***</td>
<td>-2.836***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.755)</td>
<td>(1.104)</td>
</tr>
<tr>
<td>Autonomous</td>
<td>-0.189</td>
<td>0.184</td>
<td>0.216</td>
</tr>
<tr>
<td></td>
<td>(0.288)</td>
<td>(0.339)</td>
<td>(0.370)</td>
</tr>
<tr>
<td>State</td>
<td>0.301***</td>
<td>0.236*</td>
<td>0.320**</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.125)</td>
<td>(0.135)</td>
</tr>
<tr>
<td>Log GDP</td>
<td>-0.452***</td>
<td>0.086</td>
<td>-0.134</td>
</tr>
<tr>
<td></td>
<td>(0.123)</td>
<td>(0.290)</td>
<td>(0.389)</td>
</tr>
<tr>
<td>Log Population</td>
<td>0.551***</td>
<td>0.002</td>
<td>0.277</td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
<td>(0.307)</td>
<td>(0.410)</td>
</tr>
<tr>
<td>% Gov_Exp</td>
<td>-0.062****</td>
<td>-0.011</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.036)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>0.056*</td>
<td>0.038</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.030)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Openness</td>
<td>-0.002</td>
<td>-0.003</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Local Gap</td>
<td>-13807.49</td>
<td>23429.64</td>
<td>130089.9***</td>
</tr>
<tr>
<td></td>
<td>(40285.2)</td>
<td>(23039.63)</td>
<td>(133411.1)</td>
</tr>
<tr>
<td>ELF</td>
<td>-0.000</td>
<td>0.007</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>GDP per Capita</td>
<td>1.85e-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.31e-05)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>262</td>
<td>262</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.658</td>
<td>0.734</td>
<td>0.734</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.746</td>
<td>0.736</td>
</tr>
</tbody>
</table>

Robust standard errors are in parentheses, adjusted for clustering by country. *** denotes significance at the 1 percent level; ** denotes significance at the 5 percent level; * denotes significance at the 10 percent level. Regional and legal origin dummy coefficients are omitted because none are significant.
EXCHANGE RATE POLICY AND ECONOMIC GROWTH

Yi Sun
Georgetown University

ABSTRACT

The choice of foreign exchange rate regime has long been a debate in international economics, especially now when the U.S. strongly encourages China to revalue its currency and to adopt a more flexible exchange rate regime. Previous studies used traditional tripartite classification (pegged, intermediate, and flexible) to characterize exchange rate regime. However, the pegged regime characterizes both the exchange rate regime and monetary policy framework, while the other two characterize only the exchange rate regime, without a specific monetary policy. This study will account for this discrepancy, and try to answer whether a particular exchange rate regime promote economic growth. The paper will use fixed effect estimation with a panel data of 40 countries (20 developed countries and 20 developing countries) over 30 years (from 1973 to 2002) to examine the impact of exchange rate arrangements on economic growth. The paper finds that any exchange rate regime, no matter pegged, intermediate, or flexible, with a monetary policy anchor, has a positive effect on economic growth; otherwise, they do not matter. The results suggest that a strong monetary policy framework, rather than the type of exchange rate regime, promotes economic growth.

I. Introduction

Few questions in international economics have aroused more debate than the choice of exchange rate regime. The debate has been renewed in recent years as a result of a series of economic crises around the world. Should a country fix the exchange rate or allow it to float with the market? A variety of models, theories and propositions have been produced
in economic literature, yet little consensus has emerged. This paper seeks to answer an important question in this debate: does the nature of exchange rate regime influences a country’s economic growth?

II. Literature Review

In the debate of what type of exchange rate regime is preferred, one popular belief is the bipolar, or hollowing-out, or the two-corners point of view. They argue that only polar regimes, i.e. hard pegs or pure floats are likely to be sustainable. However, this notion has not been universally accepted. Among the few empirical studies that have examined this issue in a cross-country context, scholars have reached different conclusions.

The traditional classification of exchange rate regimes used in these studies is based on the degree of flexibility of the exchange rate. The typical classification scheme is tri-partite: pegged, intermediate and flexible. This classification however, does not distinguish exchange rate policy and monetary policy framework. The pegged exchange rate regime characterizes both, but the intermediate and flexible regimes capture only the exchange rate regime type, and these two types can have weak or strong monetary policy frameworks which have different implications for economic growth.

This study will follow the IMF’s effort to take account for different monetary policy frameworks, and classifies the arrangements as either with an explicit monetary policy anchor or not. By definition, all pegged exchange rate regimes have a nominal anchor (the exchange rate). In the case of intermediate or flexible exchange rate regimes, the author uses the IMF’s classification method. If the country has a monetary aggregate anchor or an inflation targeting framework, it is classified as with anchor; otherwise, the country does not have an anchor.

The author uses a panel-data set of 40 countries—20 developed countries and 20 developing countries—over a period of 30 years—from 1973 to 2002 to estimate the impact on medium-term growth of exchange rate regime by adopting a cross-country growth framework, which controls for other variables that might influence growth and accounts for country-specific effects. Estimations are carried out using a dynamic fixed effect model that would avoid two common econometric problems—potential endogeneity of the dependent variables and correlation between the unobserved country-specific effects and the dependent variables. This study is built upon previous studies and makes contribution in three areas: first, the time period spans from 1973 to 2002 which is the most recent data; second, this study distinguishes between each country’s central banks’ declared official exchange rate regime and the de facto exchange rate regime in practice; Third and most importantly, the paper extinguishes exchange rate regime and monetary policy arrangements, because the pegged regime characterizes both the exchange rate regime and monetary policy framework, while other regimes characterize only the exchange rate regime, without a specific monetary policy framework.

III. Theoretical Foundation
The theoretical foundation of this study comes from two areas: the economic growth literature and how exchange rate regime may influence growth. The general framework in growth theory usually consists of two types of variables, as explained by Barro and Sala-i-Martin (1995). One group of variables accounts for the initial position of the economy, and the other group captures differences across countries.

However, there could be many factors that influence one country’s economic growth, and this model does not provide a clear guide as to which control variables are most important in the growth process. These explanatory variables could range from rates of savings and investment to rates of fertility and mortality, from political regime type and government spending to religion and legal framework. Given the impossibility of including every single potential factor into the model, the author in this paper will draw on literature on economic growth and try to select the most common variables used in this field. Also, given the nature of this study—the impact of exchange rate regime on economic growth, two variables will be considered as determinants by default—exchange rate regime type and monetary policy arrangements.

The literature on exchange rate regime suggests that the type of exchange rate regime can influence economic growth directly and indirectly. The exchange regime can affect economic growth directly by its effects on economic shocks (Broda, 2002). The direct impact of a flexible exchange rate regime on economic growth is summarized in the graph below. It provides an inconclusive answer to whether a flexible or a pegged regime promotes economic growth.

**Figure 1: Direct Impact of Flexible Exchange Rate Regime on Economic Growth**

Exchange rate policy can also affect economic growth indirectly through its influence
on other determinants of economic growth, for example, investment level, openness to international trade and capital flows, and development of the financial sector. Some scholars believe that a fixed exchange rate provides more certainty and thus increases investment (Aizenman, 1994), but others argue that fixed regime would result in misalignments that distort the efficient allocation of investment across sectors (Bohm and Funke, 2001). Empirical studies present mixed results.

IV. Classification of Exchange Rate Regimes and Monetary Policy Arrangements

As mentioned earlier, it is widely acknowledged that the official, de jure exchange rate regime might differ from the actual, de facto regime. It is therefore, important to report both regimes in this study.

The official exchange rate classification is published annually in IMF’s Annual Report on Exchange Arrangements and Exchange Restrictions. Ghosh et al. (1997) uses the IMF’s data to develop a tripartite classification scheme. The author of this paper will adopt their tripartite scheme as the official classification and extend their data set through 2002 using the IMF’s annual publications.

The author examines the actual regime type identified by IMF staff, and groups them into a tripartite scheme. This scheme ranks exchange rate arrangements on the basis of their degree of flexibility and the existence of formal or informal commitments to exchange rate paths.

The IMF also classifies monetary policy framework into different groups. Some countries have an exchange rate anchor, some have a monetary aggregate anchor, and some other countries may have an inflation targeting framework. In this paper, countries with one or more of the above characteristics are considered to have a monetary policy anchor; otherwise, they are without an anchor. The figure below shows the classification scheme in this paper.

<table>
<thead>
<tr>
<th>Exchange Rate Regime and Monetary Policy Arrangements</th>
<th>De jure</th>
<th>De facto</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pegged</td>
<td>(1) Currency boards (2) Single-currency pegs (3) Basket pegs</td>
<td>(1) Currency boards (2) Conventional fixed peg (3) Pegged with horizontal bands</td>
</tr>
<tr>
<td>Intermediate</td>
<td>(4) Crawling pegs (5) Target zones</td>
<td>(4) Crawling pegs (5) Exchange rates within crawling bands</td>
</tr>
<tr>
<td>Flexible</td>
<td>(6) Floats with some intervention (7) Pure floats</td>
<td>(6) Managed floating with no predetermined path for the exchange rate (7) Independently floating</td>
</tr>
<tr>
<td>With Anchor</td>
<td>(1) Exchange rate anchor (2) Monetary aggregate anchor (3) Inflation targeting framework</td>
<td></td>
</tr>
</tbody>
</table>
V. Data and Methodology

This study uses a dynamic fixed effects econometric model, which addresses two common econometric problems: endogeneity of the explanatory variables and correlation between the unobserved country-specific effects and the explanatory variables.

The general framework comes from the growth theory, in which there are two groups of variables—the initial state variables, and the variables that capture differences across countries. The econometric equation is:

\[
\text{GrowthRate}_{i,t} = \alpha_i + n_t + V_{i,t}\delta + X_{i,t}\beta + u_{i,t}
\]  

where \(\text{GrowthRate}_{i,t}\) is the growth rate of real per-capita GDP in country \(i\) and period \(t\), \(\alpha_i\) is the country-specific effect, \(n_t\) is a time dummy, \(V_{i,t}\) is a row vector of growth determinants measured at the beginning of period \(t\), \(X_{i,t}\) is also a row vector of growth determinants, but measure as averages over period \(t\), and \(u_{i,t}\) is an error term where \(t = 5\) years.

The country-specific effect, \(\alpha_i\), captures the determinants of a country’s growth rate that are not already controlled for by other explanatory variables. It accounts for unobservable characteristics that vary across countries by not over time. The time dummy, \(n_t\), captures the effects of global shocks on economic growth.

In addition, the author has selected some other explanatory variables to make sure that the estimated coefficients on the exchange rate regime variables capture the effects of the exchange rate regime on growth and not the influence of other variables. Two of the variables capture the initial state of a country’s economy. One of them is the initial real per-capital GDP. The other variable that accounts for the initial state is the education level, measured by literacy rate in this study. Other explanatory variables include the growth rate of real per-capital GDP, monetary supply (M2), ratio of government consumption to real GDP, ratio of trade to real GDP, ratio of private sector credit to real GDP, ratio of gross private capital flows to real GDP, and ratio of real investment to real GDP.

The main purpose of this paper is to examine whether a particular exchange rate regime promotes growth, thus the variable of exchange rate regime, both the traditional tripartite classification scheme (pegged, intermediate, and flexible) and the one that distinguishes exchange rate regime and monetary policy arrangements will be used. In each case, the exchange rate regime is classified according to both de jure and de facto scheme.

VI. Results and Explanations

Table 1 shows the estimation results with different types of exchange rate regimes.

When under the de facto classification, all three exchange rate regimes have an impact on economic growth. More specifically, the sign on pegged regime is positive at the 1% level and the sign on the intermediate regime and the flexible regime are both negative and
significant. This indicates that when only controlling for different exchange rate regime types, without controlling for monetary policy arrangements, pegged exchange rate regime promotes economic growth, while intermediate and flexible regimes do not.

Table 2 presents the estimation results with both exchange rate regime types and monetary policy arrangements, which are quite different from the results of previous estimations. These estimation results suggest that pegged exchange rate regime is positively associated with economic growth, and intermediate regime without an anchor is negatively linked to economic growth, and all other regimes are statistically insignificant to economic growth. These results further suggest that once monetary policy is accounted for, the choice of exchange rate policy does not have a clear relationship with economic growth. In other words, no particular exchange rate regime promotes economic growth.

Finally, Table 3 presents the estimation results that only account for monetary policy arrangements, regardless of the exchange rate regime type. The sign and significance on the coefficients of most explanatory variables remain the same as in previous estimations. One interesting result is that the sign on the coefficient of monetary policy anchor is positive and significant at 1% level. This suggests that a strong monetary policy, which has either an exchange rate anchor, or a monetary aggregate anchor, or inflation targeting framework, will have a positive impact on economic growth.

VII. Drawbacks and Future Work

The author made an attempt to examine the effect of choice of exchange rate regime on one country’s economic growth, but with certain drawbacks. First, the number of observations is small. Second, although the author draws upon a great variety of literature on both economic growth theory and exchange rate regime, the selection of explanatory variables is not without question. Also, there are two puzzling results that need to be investigated. One is the sign on the coefficient of private credit. In this study, it turns out to be negative; however, in most literature, it is expected to be positive. The other is the insignificance of investment on economic growth. This is counterintuitive, and almost all literature concludes that there should be a positive relationship.

Third, this paper adopts the method used by IMF to classify exchange rate arrangements. The IMF uses exchange rate volatility and international reserve volatility as conditioning information for a de facto classification of exchange rate arrangements. This method is promising. There are, however, two main drawbacks to this approach. External shocks can lead to errors in interpreting the true nature of the regime, as noted by Hausmann, Panizza and Stein (2001). Furthermore, it is not apparent how one can control for the higher volatility that is associated with fixed regimes undergoing revaluations (Bailiu et al., 2002). To account for these two problems, scholars need to find an alternative way to classify exchange rate regimes.

Fourth, the paper reveals that countries typically declare a more flexible exchange rate regime than they actually have; however, the paper fails to identify the reasons. Do countries deliberately try to deceive the IMF, or investors? If not, why countries tend to be more
fixed in practice? These questions could be explored in future studies related to this topic.

VIII. Political Implications and Conclusion

By using a panel data set of 40 countries from 1973 to 2002, this study finds evidence that exchange rate regimes characterized by a monetary policy anchor, whether they are pegged, intermediate, or flexible, exert a positive impact on economic growth. Moreover, the study shows that intermediate and flexible exchange rate regimes without a monetary policy anchor are detrimental for growth. These results suggest that it is the presence of a strong monetary policy framework, whether it is an exchange rate anchor, a monetary aggregate anchor, or an inflation targeting framework, is important to growth. However, it is not the type of exchange rate regime per se that matters for economic growth. A strong monetary policy can occur under different types of exchange rate regimes. This study also emphasizes the importance of considering monetary policy framework that accompanies the exchange rate regime when assessing the macroeconomic performance of alternative exchange rate regime.

Therefore, a country should not preoccupy itself with the choice of exchange rate regime when considering economic growth. Rather, a country should focus on designing and implementing an overall sound and sustainable monetary policy. International organizations, such as the IMF and World Bank, and developed countries, should not advise developing countries or emerging markets to adopt a particular type of exchange rate regime; rather, they should help those less-developed countries to adopt well-developed monetary policy arrangements. A flexible exchange rate regime is somehow perceived as better than others; however, this study shows that a flexible regime could be detrimental when the country has a weak financial sector and an under-developed monetary framework.

The debate of the choice of exchange rate regime should pause for now. Economists and government officials should focus on developing a long-term monetary strategy which controls for inflation, interest rate, money supply, and foreign exchange reserves, among other.

References


EXCHANGE RATE POLICY AND ECONOMIC GROWTH


**Table 1: Fixed Effect Estimation with Different Types of Exchange Rate Regime**

<table>
<thead>
<tr>
<th>Variable</th>
<th>De jure</th>
<th>De facto</th>
<th>De jure</th>
<th>De facto</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial real per capita GDP</strong></td>
<td>-0.0879*** (0.001)</td>
<td>-0.0831*** (0.000)</td>
<td>-0.0935*** (0.000)</td>
<td>-0.0957*** (0.000)</td>
</tr>
<tr>
<td><strong>Money supply (M2)</strong></td>
<td>-0.0345** (0.039)</td>
<td>-0.0327** (0.054)</td>
<td>-0.0285** (0.038)</td>
<td>-0.0298** (0.043)</td>
</tr>
<tr>
<td><strong>Government consumption/GDP</strong></td>
<td>-0.1435*** (0.008)</td>
<td>-0.1335*** (0.006)</td>
<td>-0.1446*** (0.005)</td>
<td>-0.1513*** (0.001)</td>
</tr>
<tr>
<td><strong>Trade/GDP</strong></td>
<td>0.0647*** (0.002)</td>
<td>0.0743*** (0.000)</td>
<td>0.0850*** (0.000)</td>
<td>0.1063*** (0.000)</td>
</tr>
<tr>
<td><strong>Literacy</strong></td>
<td>0.022*** (0.000)</td>
<td>0.0246*** (0.000)</td>
<td>0.0231*** (0.000)</td>
<td>0.0269*** (0.000)</td>
</tr>
<tr>
<td><strong>Private credit/GDP</strong></td>
<td>-0.0075 (0.357)</td>
<td>-0.0123*** (0.007)</td>
<td>-0.0051 (0.538)</td>
<td>-0.0116*** (0.006)</td>
</tr>
<tr>
<td><strong>Private capital flows/GDP</strong></td>
<td>0.1148*** (0.002)</td>
<td>0.1169*** (0.001)</td>
<td>0.1165*** (0.001)</td>
<td>0.1273*** (0.001)</td>
</tr>
<tr>
<td><strong>Investment</strong></td>
<td>0.1439 (0.325)</td>
<td>0.1533 (0.196)</td>
<td>0.1637 (0.548)</td>
<td>0.2034 (0.166)</td>
</tr>
<tr>
<td><strong>Pegged</strong></td>
<td></td>
<td></td>
<td>0.0059** (0.036)</td>
<td>0.0112*** (0.003)</td>
</tr>
<tr>
<td><strong>Intermediate</strong></td>
<td>-0.0038 (0.257)</td>
<td>-0.0027* (0.079)</td>
<td>-0.0146*** (0.000)</td>
<td>-0.0071*** (0.002)</td>
</tr>
<tr>
<td><strong>Flexible</strong></td>
<td>-0.0013 (0.568)</td>
<td>-0.0071** (0.036)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(1) The figures in parentheses are p-values. This holds for all estimation results.
(2) *, ** and *** indicate statistical significance at 10, 5, 1 percent levels, respectively.
### Table 2: Fixed Effect Estimation with Different Types of Exchange Rate Regime and Monetary Policy Arrangements

<table>
<thead>
<tr>
<th>Variable</th>
<th>De jure</th>
<th>De facto</th>
<th>De facto</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial real per capita GDP</td>
<td>-0.0735***</td>
<td>-0.0949***</td>
<td>-0.0846***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Money supply (M2)</td>
<td>-0.0275*</td>
<td>-0.0306**</td>
<td>-0.0298**</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.041)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>Government consumption/GDP</td>
<td>-0.1007*</td>
<td>-0.1128**</td>
<td>-0.0859***</td>
</tr>
<tr>
<td></td>
<td>(0.075)</td>
<td>(0.034)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Trade/GDP</td>
<td>0.0872***</td>
<td>0.01097***</td>
<td>0.0143***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.008)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Literacy</td>
<td>0.0169***</td>
<td>0.0215***</td>
<td>0.0231***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Private credit/GDP</td>
<td>0.0064</td>
<td>0.0024</td>
<td>0.0038*</td>
</tr>
<tr>
<td></td>
<td>(0.308)</td>
<td>(0.548)</td>
<td>(0.0695)</td>
</tr>
<tr>
<td>Private capital flows/GDP</td>
<td>0.0619***</td>
<td>0.0796***</td>
<td>0.0927***</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Investment</td>
<td>0.1818</td>
<td>0.1427</td>
<td>0.1579</td>
</tr>
<tr>
<td></td>
<td>(0.209)</td>
<td>(0.572)</td>
<td>(0.638)</td>
</tr>
<tr>
<td>Pegged</td>
<td></td>
<td></td>
<td>0.0028**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.032)</td>
</tr>
<tr>
<td>Intermediate with anchor</td>
<td>0.0069</td>
<td>-0.0037</td>
<td>-0.0012</td>
</tr>
<tr>
<td></td>
<td>(0.486)</td>
<td>(0.565)</td>
<td>(0.671)</td>
</tr>
<tr>
<td>Intermediate without anchor</td>
<td>0.0012</td>
<td>-0.0089***</td>
<td>-0.0067***</td>
</tr>
<tr>
<td></td>
<td>(0.754)</td>
<td>(0.003)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Flexible with anchor</td>
<td>0.0008</td>
<td>-0.0018</td>
<td>0.0025</td>
</tr>
<tr>
<td></td>
<td>(0.836)</td>
<td>(0.674)</td>
<td>(0.626)</td>
</tr>
<tr>
<td>Flexible without anchor</td>
<td>0.0003</td>
<td>0.0024</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.894)</td>
<td>(0.435)</td>
<td></td>
</tr>
</tbody>
</table>

(1) The figures in parentheses are p-values. This holds for all estimation results.
(2) *, ** and *** indicate statistical significance at 10, 5, 1 percent levels, respectively.
TABLE 3: FIXED EFFECT ESTIMATION WITH DIFFERENT MONETARY ARRANGEMENTS

<table>
<thead>
<tr>
<th>Variable</th>
<th>De jure</th>
<th>De facto</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial real per capita GDP</td>
<td>-0.0978***</td>
<td>-0.0725***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Money supply (M2)</td>
<td>-0.0437**</td>
<td>-0.0391**</td>
</tr>
<tr>
<td></td>
<td>(0.0406)</td>
<td>(0.0370)</td>
</tr>
<tr>
<td>Government consumption/GDP</td>
<td>-0.1224***</td>
<td>-0.1165***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.0002)</td>
</tr>
<tr>
<td>Trade/GDP</td>
<td>0.0773***</td>
<td>0.0899***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Literacy</td>
<td>0.0135***</td>
<td>0.0131***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Private credit/GDP</td>
<td>-0.0232</td>
<td>-0.0316*</td>
</tr>
<tr>
<td></td>
<td>(0.137)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>Private capital flows/GDP</td>
<td>0.0978***</td>
<td>0.0936***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.0004)</td>
</tr>
<tr>
<td>Investment</td>
<td>0.1293</td>
<td>0.1344*</td>
</tr>
<tr>
<td></td>
<td>(0.484)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>Monetary policy anchor</td>
<td>0.0047***</td>
<td>0.0082***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

(1) The figures in parentheses are p-values. This holds for all estimation results.
(2) *, ** and *** indicate statistical significance at 10, 5, 1 percent levels, respectively.
JOINING THE EUROPEAN UNION: THE REAL EFFECT OF FOREIGN DIRECT INVESTMENT

Lucia Franzese
Georgetown University

I. Introduction

Within the framework of an increasingly globalizing world, many countries are following a new trend of integrating large economic areas. Several regions throughout the world have formed and joined common markets with the short-term intentions of stimulating intra-regional trade and investment and the long-run expectations that the combination of larger markets, tougher competition, more efficient resource allocation, and other various positive externalities will raise the growth rates of participating economies. Ultimately, these countries aspire to accelerate the process of economic development.

One crucial incentive of joining the European Union is the expectation of increased Foreign Direct Investment (FDI). FDI is highly desirable as it plays a key role as a source of financing and tax income, and generates new job opportunities and growth-enhancing technology spillovers. Furthermore, FDI has the potential ability to create new industries and shift production toward higher value-added activities (Medvedev, 2006). In recent years, the newest accession states in the EU have had specific advantages in terms of attracting FDI over other low-wage countries, such as China and India, as they have higher productivity, reduced tariffs and border problems, lower shipping costs, and a more familiar business environment.

Unfortunately, it is possible that this anticipation of net FDI flow increases is slightly optimistic and perhaps, joining the EU, or any regional integration, does not always deliver the promises of surging FDI for countries with less developed economies. For example, a paper by Antonios Georgopoulos and Heinz Gert Preusse, which concentrates on FDI in the European Union and Greece in particular, examines the strategies of Trans-national Corporations (TNCs) in the host economy that have taken place in the context of the reduction of EU market fragmentation and barriers to entry. The empirical results suggest that, in general, the participation of Greece in the EU has not increased the attractiveness of the country as a production base for TNCs. In fact, Greece neither managed to attract considerable amounts of export oriented foreign investment nor was it efficient in seeking FDI (Georgopoulos, 2006). The overall picture appears to be that, while the membership
in the European Union can have a considerable impact on the pattern of trade in Europe in general, some particular countries have not reaped all the expected benefits of increased foreign investment.

**Objective**

The main objective of this paper is to predict the effect of a country’s entrance into the European Union on its levels of net Foreign Direct Investment (FDI) flows. This paper contributes to the literature focusing on FDI and Regional Economic Integrations by adding to a small, but lively, debate questioning whether joining a preferential trade agreement does indeed increase net FDI flows into a member country, using the European Union as a case study.

**II. Literature Review**

As Foreign Direct Investment (FDI) is a complex and multi-dimensional subject, it has been difficult for economists to develop one single explanatory theory or even pinpoint all the determinants. However, the literature does indicate three main factors affecting FDI: trade, market size, and growth. Furthermore, when looking at the determinants of FDI for a member of a regional economic integration, such as the European Union, two additional factors can be identified as affecting FDI: investment climate, and investment provisions, which involves the elimination of regulatory and legal obstacles to international capital flows (Medvedev, 2006).

There does not yet appear to be consensus on the key determinants of FDI based on the evidence presented in the empirical literature. In part, this is because there are different types of FDI, which, in turn, are affected by different factors. The evidence seems to suggest that most of the new FDI flowing into members of these regional economic integrations is due to tariff-jumping and/or attraction to larger markets. However, it is unclear what overall impact regional economic integration, such as the European Union, will have on FDI. A general consensus seems to be that investment can be expected to cluster in those parts of the Regional Integration Areas where the investment climate is most favorable, and some countries may therefore be left with less FDI than before, Blomström and Kokko (1997). However, while the majority of studies tend to confirm a positive effect of EU-expansion on inward FDI, the evidence is not beyond all doubt. It has been argued by Eckert et al. (2005) that expansion might also have negative effects on FDI, which have not been considered sufficiently. This paper will specifically speak to this small but critical debate in the literature.

**III. Model for Empirical Analysis and Estimation Results**

**Model**

This paper will employ the initial model from Medvedev (2006) to use empirical data to develop a model to predict the net FDI inflows of a country after joining the European Union. The countries observed will be the EU 25, as of December 31, 2006 from 1970 to
The dependent variable is the average net FDI flows as a ratio to GDP. FDI is defined as the net inflows of investment to acquire a lasting management interest (10 percent or more of voting stock) in an enterprise operating in an economy other than that of the investor. In order to smooth out large one-time investments, the dependent variable, FDI, is defined as a moving centered average of three years. For example, for the data points in the year 1970, the FDI is the average corresponding to 1969, 1970, and 1971.

The independent variables can be divided into three categories: local effects, global effects, and the effect of EU membership. First, the local effects are taken into account through the level of GDP and its growth rate, the outward orientation or openness of an economy, labor costs in a particular country measured by the per capita income relative to that of the U.S., and macroeconomic stability that is gauged using inflation. Second, the global effects are measured through world GDP growth and world FDI scaled to world GDP. Third, the effects of EU membership are gauged using two dummy variables representing expected EU membership (with a value of one for the two years previous to entrance into the European Union) and actual EU membership (with a value of one for every year the country has been in the European Union), and the size of the extended common market (measured as the aggregate EU GDP scaled to the total European population, including non-member countries).

The model, defined as a panel data set, incorporating fixed country and time effects, uses the following equation:

$$ FDI_{it} = \alpha + \gamma_i + \beta_1 GDPREL_i(t-2) + \beta_2 OPEN_i(t-2) + \beta_3 GNIREL_{it} + \beta_4 GDPGRO_i(t-1) + \beta_5 CPIGRO_{it} + \beta_6 WLDFDI_{it} + \beta_7 WLDGRO_{it} + \beta_8 EUGDP_{it} + \beta_9 EUEXP_{it} + \beta_{10} EUMEM_{it} + \epsilon_{it} $$

(1)

where

- $ FDI_{it} $ is the ratio of the net FDI inflows of country $ i $ averaged over time periods $(t-1), t, \text{and} (t+1)$ to the average GDP of country $ i $ for the same time periods

- $ \alpha, \gamma_i $ are the constant and country fixed effects

- $ \beta_1 GDPREL_i(t-2) $ is the ratio of GDP of country $ i $ to GDP of the US in $(t-2)$

- $ \beta_2 OPEN_i(t-2) $ is the trade-to-GDP ratio of country $ i $ in $(t-2)$

- $ \beta_3 GNIREL_{it} $ is the ratio of per capita GNI of country $ i $ to per capita GNI of the US in $ t $

- $ \beta_4 GDPGRO_i(t-1) $ is the GDP growth rate of $ i $ between $(t-1)$ and $(t-2)$

- $ \beta_5 CPIGRO_{it} $ is the rate of inflation in $ i $ between $ t $ and $(t-1)$

- $ \beta_6 WLDFDI_{it} $ is the ratio of net world FDI inflows to world GDP (excluding country $ i $) in $ t $

- $ \beta_7 WLDGRO_{it} $ is the growth rate of world GDP between $(t-1)$ and $(t-2)$

- $ \beta_8 EUGDP_{it} $ is the log of the sum of GDP of EU partners of country $ i $ at
time $t$ (excluding country $i$) as a ratio to total EU population of all 25 countries

$\beta_9 \text{EUEXP}_{it}$ is equal to 1 for two years prior to country $i$ joining the EU

$\beta_{10} \text{EUMEM}_{it}$ is equal to 1 for every year country $i$ is part of the EU

$\varepsilon_{it}$ is the error term

The model tests the validity of the null hypothesis, which states that that joining the EU has no impact on net FDI flows to accession countries. Conversely, the alternative hypothesis states that joining the EU has an impact on net FDI flows accession countries.

**IV. Regression Results**

Equation (1) was run as a Fixed Effects model to estimate the regression coefficients, which are shown in Table 1. Theoretically speaking, the Fixed Effects model appears superior to the random effects model for this particular analysis because it takes into account differences among the countries that remain constant throughout the years. For example, Ireland’s use of English as its spoken language may be an added incentive for American firms looking to invest abroad, as they have a common language; the fixed effects coefficients separate this effect from other factors or variables. However, the Hausman test was still performed to determine whether the model should be run as a Fixed Effects or Random Effects model by testing the null hypothesis that the coefficients estimated by the efficient Random Effects estimator are the same as the ones estimated by the consistent Fixed Effects estimator. As the test came out with a significant P-value, the Fixed Effects model is considered to be a better model for this situation.

Overall, the equation (1) appears to do a satisfactory job explaining the variation in the dependent variable, with an $R^2$ of 33 percent, which is relatively good for panel data, and a statistical significance for the F statistic at the one percent level. Furthermore, the coefficients on all of the model’s control variables are signed in accordance with the expectations and are statistically significant at least at the 10 percent level.

**Local Effects**

Regression results indicate that local conditions are important factors determining FDI flows. In particular, the importance of a country’s market size as reflected by its GDP is shown by the regression coefficient of GDPREL, which is positive and quite large with a statistical significance at the 1 percent level. As in the previous literature, this suggests that a larger country in terms of GDP will attract more FDI because of the size of the potential market for the investing companies.

The opposing effects of trade on net FDI are captured by the OPEN variable. The OPEN coefficient is close to zero and statistically significant at the 1 percent level. This finding is particularly interesting, as it suggests that the substitution and complementary effects of trade seem to cancel each other out, and ultimately, trade neither increases nor decreases the total levels of FDI into a particular country.
GNIREL, which is the measure of the Gross National Income per capita of a particular country relative to the United States’ GNI per capita in the same year, is intended to capture differences in relative labor costs. The negative coefficient shows that higher labor costs are a deterrent for FDI inflows. The CPIGRO variable accounts for inflation, which does not appear to have an effect on FDI, as its regression coefficient is close to zero. This is similar to Medvedev’s findings.

Finally, the relationship between a country’s GDP growth and FDI is shown by the GDPGRO variable, which indicates a slight positive relationship with net FDI inflows with a 10 percent level of significance.

**Global Effects**

As shown by the regression results, the performance of the global economy has an important effect on FDI flows. Specifically, the coefficient for the WLDFDI variable is extremely high and statistically significant at the 1 percent level. This result is as expected, indicating that as the world FDI flows increase, so do the FDI flows into Europe.

The world GDP growth is accounted by the WLDGRO variable, whose coefficient is somewhat negative and statistically significant at the 10 percent level. The relationship between world GDP growth and FDI into the European Union is difficult to anticipate. The negative correlation could suggest that, during period of global economic expansion, other countries from around the world—such as emerging economies—become more competitive in terms of attracting FDI.

**Effect of EU Membership**

Regression results point to an important, albeit partly unexpected, relationship between EU membership and FDI flows. In particular, the regression coefficient for EUGDP, which indicates the effect of the entire European Union market size, is found to be positive, at the 1 percent level of statistical significance. This positive relationship is to

---

**Table 1: Fixed Effects Estimation Results: Net FDI Inflows to GDP**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDPREL</td>
<td>127.473**</td>
<td>27.410</td>
</tr>
<tr>
<td>OPEN</td>
<td>0.046**</td>
<td>0.014</td>
</tr>
<tr>
<td>GNIREL</td>
<td>-3.742†</td>
<td>1.919</td>
</tr>
<tr>
<td>GDPGRO</td>
<td>0.114*</td>
<td>0.045</td>
</tr>
<tr>
<td>CPIGRO</td>
<td>0.021*</td>
<td>0.009</td>
</tr>
<tr>
<td>WLDFDI</td>
<td>198.522**</td>
<td>43.970</td>
</tr>
<tr>
<td>WLDGRO</td>
<td>-0.223†</td>
<td>0.115</td>
</tr>
<tr>
<td>EUGDP</td>
<td>3.481**</td>
<td>1.216</td>
</tr>
<tr>
<td>EUMEM</td>
<td>-1.923**</td>
<td>0.720</td>
</tr>
<tr>
<td>EUEXP</td>
<td>-0.216</td>
<td>0.356</td>
</tr>
</tbody>
</table>

N: 516
R²: 0.333

Significance levels: † 10% * 5% **1%

Note: Standard errors are robust.
be expected, as companies originating from outside the European Union are not only attracted to a large individual country market size, but also the potential market of the entire integrated region.

While anticipated EU membership has no apparent effect on FDI, actual EU membership exhibits a negative effect on FDI. Specifically, as shown on Table 1, the EUMEM coefficient that represents actual EU membership exhibits a negative sign, at the 1 percent level of statistical significance. This result suggests that, once a country joins the European Union, it is likely that the net FDI to that country will decrease. This is contrary to the expected sign, as the overall effect on net FDI inflows of joining the European Union is generally thought to be positive.

There are several reasons that can explain this negative correlation. One possibility is that rather than adding a competitive edge in FDI attraction to each country that joins, the European Union causes the countries to compete against each other for FDI. Once a firm is investing in one country within the European Union, it no longer has to deal with tariffs or other border problems. Thus, when making a decision of where to invest abroad within a particular economically integrated region, a Multinational Corporation will simply revert back to the typical determinants for FDI location, such as factor costs. As the membership in the European Union grows larger, there are more countries that have equal characteristics in several aspects of the decision factors and must compete in a few specific categories.

Another explanation could be that, as a country joins the European Union, its economy begins to develop, ushering in higher standards of living and therefore, higher costs of production and inputs, such as labor, which ultimately causes a country to lose its competitiveness and threatens the movement of foreign operations to a lower cost country (the so-called “flying geese” effect). These two explanations can go hand-in-hand.

An alternative explanation offered by Eckert et al. focuses on FDI flows seeking psychic market proximity as the reason for reductions in FDI when a country joins the European Union. The authors explain that by the time a country is allowed to enter into the European Union, there no longer exists such insurmountable cultural differences. From the viewpoint of Western firms, there is a collective acculturation taking place in the countries before they join the European Union, which may lead to a rapid increase in psychic market proximity. A consequence of this development might be that FDI is no longer necessary to achieve the required degree of psychic market proximity. Thus, a decrease in market seeking FDI may decrease in the new areas, Eckert et al. (2005).

More likely, the negative relationship shown in the study is the result of a combination of the three alternative explanations. Furthermore, it is necessary to note that this paper is analyzing the conditional effect of EU membership; thus, membership could still cause FDI to increase indirectly through an increase in growth, for example. It is also possible that the positive relationship is present but the model in equation (1) is not capable of correctly capturing it due to several potential problems, including model misspecification and the use of inappropriate estimation methods.
V. Policy Implications

As the model indicates, the effect of joining the European Union can be detrimental to the attractiveness of a member country for FDI, particularly in the long run. This finding, along with the fact that EU nations are losing the race for FDI against non-EU states to nations from other parts of the world—see Oxlheim et al. (2004)—might require remedial actions from policy makers, both at a national and European scale.

As stated by Oxlheim et al. (2004), in a world where all macro policies (monetary, fiscal or industrial policies) are becoming more or less the same, the government and its agents in an individual member country will search for new ways to stand out with an FDI-seducing profile. One could expect the policy makers of individual countries in the EU to implement incentives to increase attractiveness for FDI, which would produce unfair and race-to-the-bottom competition for FDI within the EU.

A more effective approach would be perhaps for policy makers, both national and European, to strengthen the overall competitiveness of the EU for FDI, as compared to other nations and regions in the world (Delios, 2004). One possibility would be to implement incentives on a European-scale in order to slow or reverse the trend of losses by the EU in the world-wide competition for FDI inflows. Furthermore, such a policy would likely stop individual countries from implementing their own policies for FDI attraction. If European-wide policies were implemented, this would leave room for regions, as opposed to countries, to compete for FDI.

Another potential action would be for European policy makers to promote intra-EU FDI, thus not only promoting the development of European multinationals, but also providing the same benefits to countries that before joining the European Union may have been benefiting from extra-EU FDI.

Regardless of the measures taken by the European policy makers to alleviate the losses of FDI, it is crucial for those countries that have yet to join the European Union to take into consideration the possibility of a decline in FDI.

VI. Conclusions

This paper has outlined determinants of FDI, with a focus on the effects of joining the European Union. It found that the relationship between membership of a regional economic integration and FDI is complex, as multiple factors determine the flows of FDI. While it is generally believed that the positive effects would outweigh the negative consequences and ultimately increase the inflows of FDI into member countries, the findings from this paper suggest that this conclusion is not necessarily true. In fact, results show that actual EU membership have a negative effect on FDI. Three reasons are given to explain this phenomenon: i) as many countries in the EU have similar macro policies, the individual European countries have to actively compete for FDI against one another; ii) EU membership is expected to promote overall development, thus resulting in a higher standard of living, which raises factor prices, such as labor and thereby decreases its attrac-
tiveness for FDI; and iii) by the time a country joins the European Union, it is close enough to the other countries in terms of culture and way of doing business, so that psychic market proximity is no longer a concern for Multinational Corporations and thus decreases the need for multiple investments within the European Union.

The analysis also provides additional insights into the relationship between EU membership and FDI inflows, suggesting that it is the length of membership rather than the initial impact of EU membership that results in a decrease in FDI flows. Moreover, the analysis also suggests that lower income countries, particularly those with low factor prices, are more likely to benefit from increased FDI, at least initially, when joining the EU.

These findings have significant policy implications. Primarily, this paper puts into question the assumption that joining the European Union will improve the attractiveness of a member state for FDI, which is presumably a major incentive for joining the economic region. Furthermore, it raises the concern of whether there is competition among EU countries to attract FDI in their respective countries and the damaging implications of this competition.

There are several areas for further research that are necessary to understand these complexities of changes in FDI in the context of regional economic integration. Firstly, identifying the effects of joining the EU on FDI by sector would be essential in order to formulate policies specifically targeted at certain industries. It also would be interesting to complete a study similar to the one presented in this paper, but focusing on sub-states (i.e., regions within each nation), as other results may be masked by country effects. Lastly, further exploration into intra- and extra-regional FDI would be important to separate the differing effects, particularly to better generate policies that will have constructive impact on FDI. While the paper briefly speaks to the role of policy makers in the European Union, there is a need for further research in various directions to shed light on the actual effects of joining the European Union and construct more efficient policies.

References


Determinants of Sovereign Risk Premia in Commodity-Dependent Latin American Countries

Stephen Brinkmann
Georgetown University

I. Introduction

In this study I conduct an empirical examination of the determinants of sovereign credit risk premia in commodity-dependent Latin American countries. Using realized trading data for long-term bond debt issued by five Latin American governments, I conduct time-series analysis to determine what effect (if any) several hypothesized factors have on a proxy for the country risk premia. Country-specific explanatory variables include several macroeconomic indicators typically used in credit risk analysis and special indicators which attempt to capture the country’s dependence on primary commodity exports. Results are mixed, and in some cases I am able to demonstrate a ceteris paribus link between these factors and sovereign credit risk premia (SCRP). In general, though, my findings indicate that most of the variation in sovereign credit risk premia can be attributed to changes in the global appetite for risk.

This study is motivated in part by the need of a new analytical framework to value debt issued by sovereign entities, as these securities are influenced by a myriad of factors not affecting technically similar instruments like corporate bonds. Macroeconomic trends and perturbations, for instance, are likely to have a much greater impact on the solvency of sovereign issuers, and the unique importance of the issuer’s willingness to pay merits particular consideration. While traders and investors in the financial centers of the “north” certainly need such a framework to price emerging market sovereign debt, governments in the developing countries of the “south” also need to understand the price drivers of the bonds that they have issued, because those same factors will determine the interest rate they pay on any debt they will issue in the future. I single out countries of the “south” here because the structure of their economies is so different from that of industrialized countries that an entirely different set of factors is thought to influence the price of their sovereign debt. In particular, countries whose economies are heavily dependent on exports of primary commodities deserve special attention. Because these governments derive a significant portion of their tax revenue from primary commodities, which have proven to be some of the most volatile markets in the world, they should expect to pay a penalty when
borrowing capital in the international debt markets in the form of higher interest rates. I hope to test this hypothesis (albeit indirectly) in this study.

There has been a fair amount of empirical work relevant to the search for determinants of SCRP. Since a country’s SCRP is not a variable that is directly observable from market activity, researchers have chosen to proxy SCRP in different ways. I chose to model the credit spread of sovereign bonds that are actively traded in international markets against a number of hypothesized influencing variables, employing the “balance sheet” approach to sovereign credit analysis. I will also take the step of moving from the typical method of analysis (cross-country) to a time-series framework. Time series analysis is important in this field because while a cross-country model can do a good job of explaining why Colombia pays a credit spread of 700 basis points over U.S. Treasuries and France only pays 10, these models cannot really say why Colombia’s credit spread would go from 700bps to around 200bps in the space of a few years. For reasons I’ll discuss in the next section, time-series models of credit spread determinants seem much harder to specify than cross-sectional models, so there is not a great deal of research in that vein. My study also builds on the current literature in that I seek to parse out the impact of the global appetite for risk (or “price” of risk). Since it appears that a great deal of what we label “credit risk” (as proxied by the credit spread on a bond) is actually systematic market risk, we can improve the explanatory power of other variables by controlling for this world appetite for risk.

Last, but certainly not least, my model will differ from the current literature with the inclusion of commodity price volatility as an explanatory variable. Thus, there is a connection between my work and the “resource curse” literature, which gives me theoretical reason to believe that commodity price volatility might have explanatory power for the credit risk premia paid by resource-dependent countries. The volatilities of different commodities are thought to be more-or-less independent, and only certain commodities will be relevant for certain countries, so panel data is not conducive to this particular study. Instead, I will build multiple time-series models to evaluate determinants of credit risk premia for particular countries.

II. Empirical Strategy and Data

I design the empirical strategy with the goal of explaining variation in sovereign credit spreads over time, which is quite distinct from measuring variation in spreads across countries. Researchers must parse out strong time trends both in spreads and explanatory variables, and this creates fundamental “spurious regression” problems. We examine variables that are trending together, but the extent to which that relationship makes economic sense versus being coincidental is hard to determine. Over my sample period (2004-2007) we saw a significant decline in emerging market credit spreads and a commodities boom, but if we construct a time-series model in levels, can we really say that one impacts the other? With these ideas in mind, I chose to construct a fairly “rigid” model, in the sense that it will be difficult to show significant statistical relationships between variables, but the
results are less-likely to be biased by trends. The principle time-series regressions are specified as follows:

\[
\sum_{i=5}^{n} \Delta\text{SCRPT}_t = \beta_0 + \beta_1 \Delta\text{INFLATION}_t + \beta_2 \Delta\%\text{EXPORTS}_t + \beta_3 \Delta\%\text{DEBT}_t + \beta_4 \Delta\text{EMBI}_t +
[\beta_{i1}\sigma_{\text{commodity}_{i,t}} + \beta_{i2}\Delta\%P_{\text{commodity}_{i,t}}] + \beta_{n+1} t + \varepsilon_t
\]

Where:

- \( t \) = a monthly time step. \( t \) is also included as an explanatory variable to capture any time-trend in the variation of credit spreads not parsed out by specifying the model in first differences.
- \( \Delta\text{SCRPT}_t \) = the change in credit spread of the issuer over month \( t \). This variable was constructed by calculating yield-to-maturity of one of the sovereign bonds using trading price data and subtracting the risk-free rate of interest and cost of financing (as measured by the level of the interest rate swaps curve (or LIBOR) at time \( t \)) to provide a pure measure of sovereign credit risk.
- \( \Delta\text{INFLATION}_t \) = the change in the issuing country’s annualized inflation rate over month \( t \). The expected sign is positive, suggesting that increases in inflation cause an increase in credit risk premia.
- \( \Delta\%\text{EXPORTS}_t \) = the percentage change in the issuing country’s level of exports (in USD) over month \( t \). The expected sign is negative, suggesting that an increase in exports (a source of tax revenue and foreign exchange for the government) will improve credit quality.
- \( \Delta\%\text{DEBT}_t \) = the percentage change in the issuing country’s level of sovereign external debt (denominated in all currencies, but measured in USD) over month \( t \). The expected sign is positive, because a higher debt burden is thought to diminish creditworthiness (and increase risk premia).
- \( \Delta\text{EMBI}_t \) = the change in the JPMorgan Emerging Market Bond Index (Latin America Strip Spreads) over month \( t \). This variable proxies the global appetite for risk in Latin American emerging markets. I expect the sign to be positive, as an increase in regional risk premia should result in an increase in individual country risk premia.
- \( \sigma_{\text{commodity}_{i,t}} \) = the estimated price volatility of commodity \( i \) at time \( t \), measured via either simple historical standard deviation or GARCH (1,1) model (Generalized Autoregressive Conditional Heteroskedasticity) depending on the strength of autocorrelation tendencies in volatility data.
- \( \mu_{\text{commodity}_{i,t}} \) = the mean monthly change in price of commodity \( i \) over the one-year period preceding month \( t \). A discussion of why volatility is expressed as \( \sigma/\mu \) (the coefficient of variation) also follows. I expect the volatility variables to be positively correlated with credit spreads.
DETERMINANTS OF SOVEREIGN RISK PREMIA

$\Delta \%P_{commodity_{i,t}}$ = the percentage change in price of commodity $i$ over month $t$. I predict a negative sign for commodity price variables, which suggests that higher commodity prices will result in lower credit spreads.

In each of the five time series regressions, I incorporate commodity price and estimated volatility as explanatory variables only for the commodities thought to be important to the economy of the country whose credit risk premium I am seeking to explain. I determined what commodities should be included for each country by first consulting the 2007 CIA World Factbook to get an idea of the countries’ primary exports, and then confirming these results with other sources of information on the issuer’s economy (including the prospectus documents of the bonds themselves). My study was restricted to five Latin American countries with the best available data: Chile, Colombia, Jamaica, Mexico, and Venezuela.

III. Empirical Results

As expected, it was somewhat difficult to find an abundance of significant relationships in the time-series modeling. I ran 5 OLS regressions of the model specification described in the previous section. While most variables were individually insignificant, an encouraging sign is the high explanatory power of the model as a whole. R-squared and F-statistic values show high joint significance of the explanatory variables tested.

As a robustness test for the results of the initial models, I begin by regressing SCR for each country against only the EMBI and the time trend. EMBI was the only variable that was consistently of very high individual significance and of the expected sign. I therefore sought to demonstrate how much of the variation in sovereign credit spreads could be attributed to the global appetite for risk alone. The explanatory power of EMBI varied considerably across countries. In the Jamaican case, as expected, SCR seems totally unrelated to EMBI, or any other tested explanatory variable. For Chile and Mexico, global appetite for risk seems to explain a substantial amount of the variation in credit spreads, but not all of it. The new regressions for these countries yielded R-squared values of 0.1368 and 0.2640, respectively, compared with 0.5834 and 0.5481 for the initial regression including all explanatory variables. In the Chilean case, we could not reject the null that EMBI was an individually insignificant determinant of SCR. For Colombian and Venezuelan external debt, however, it appears that variation in credit spreads can be attributed almost entirely to variation in global appetite for risk, as proxied by the EMBI. The truncated regression for Colombia yields an R-squared of 0.5916 and for Venezuela, 0.5008. The original time-series models experienced their greatest success in predicting credit spread changes for these two countries, and it appears that much of that success is due to the high significance of the EMBI. In these cases, the EMBI variable had a t-statistic of 3.99 and 4.72, respectively. The coefficients on the EMBI variable were 0.935 for Colombia and 0.835 for Venezuela, meaning that for every 1 basis point increase in the EMBI, Colombian or Venezuelan credit spreads would increase by nearly as much. Additional robustness tests and tangential models were also performed.
IV. Conclusion

After conducting background research and empirical analysis on sovereign credit spreads in Latin America, I am unable to unequivocally demonstrate that country-specific factors, including commodity volatility, are determinants of a country’s sovereign credit risk premium. Through various specifications of a time-series model, I have shown that some of these country-specific variables do have some explanatory power in predicting the variation of credit spreads, and sometimes it is enough to show individual significance of variables at reasonable levels of statistical confidence. As mentioned earlier, though, time-series models of the variation in credit spreads are particularly difficult to specify and achieve meaningful results, so I take even the slightest sign of a relationship to be quite encouraging. During robustness testing, I was at least able to achieve significant joint-explanatory powers of the hypothesized determinants of SCRP.

Perhaps most importantly, this study contributes to the literature in that I am able to attribute most of the variation in SCRP to a single factor: global appetite for risk as proxied by the EMBI. By adding this variable, I am able to capture a great deal of the variation in SCRP that other time-series models fail to explain. Returning to the “resource curse” motivation for conducting this study, I am able to say that exposure to highly volatile commodity markets does have something to do with why resource-dependant countries pay higher interest rates on their external borrowing, but much less so than a primary factor that is completely out of the borrowing country’s hands: the overall strength or weakness of the global credit markets.

Implications for Market Participants and Hedging “Appetite Risk”

The results of this study have profound implications for participants in the sovereign credit markets, issuers and investors alike. Of greatest importance is the observation that a sovereign bond issuer’s credit spread, something typically attributed solely to country-specific factors driving that government’s creditworthiness, depends much more on global appetite for risk. In this final section, I demonstrate why market participants need to take this into account and suggest some strategies for hedging this risk.

In the first case, take the perspective of an investor in emerging-market government bonds. When a credit trader takes the view that the credit quality of an issuer will improve or worsen, a typical strategy is to purchase or sell that issuer’s bonds and then hedge out interest rate risk. For example, say that the trader predicts that Colombia’s creditworthiness will improve substantially over the next six months. He or she would purchase Colombian sovereign bonds, and then sell U.S. treasuries, initiate a floating-to-fixed interest rate swap contract (pay fixed), sell interest rate futures, or use some other hedging technique. The trader would then assume that, irrespective of yield curve behavior, the investment would have a positive return because the bonds would rally as Colombia’s credit spread compresses. However, I have demonstrated that some 60% of the variation in Colombian credit spreads is entirely dependent on the global appetite for risk, and actually has nothing to do with credit quality. Even if the trader is right about his or her predictions for Colombian creditworthiness, global appetite for risk could weaken over the next
six months, causing an opposite (widening) effect for Colombian credit spreads. If the weakening is strong enough, it could outweigh any creditworthiness gains and depress the price of the Colombian bonds.

Next, take the case of an issuing government. Say Colombia plans on issuing new fixed-rate sovereign bonds about six months into the future. Both corporate and sovereign borrowers planning on issuing bonds have good reason to hedge interest rate risk over the period before the bonds actually go to market. If interest rates were to rise over the next six months, Colombia would have to pay a correspondingly higher yield on its new bonds. Therefore, it would short U.S. Treasuries, sell interest rate futures, etc. to hedge that risk. But what if interest rates stayed unchanged, but the global appetite for emerging market debt deteriorated? Through no fault of its own, Colombia would be forced to offer a higher yield on its bonds in order to raise the capital it needs.

It should be obvious that both emerging-market debt issuers and credit investors need to think about hedging “appetite risk” – the possibility that the global appetite for risk will fluctuate over time. In today’s market, this could probably be best accomplished through buying or selling baskets of sovereign debt issued by the country in question’s “peers”. Of course, the issuer or investor would be concerned that variation in peer country-specific factors (i.e. macroeconomic or commodity factors) might inadvertently affect the behavior of their hedge, but if the basket of bonds traded was broad enough, these would likely cancel each other out. Transaction costs might be high for such a trade, though, so a much more targeted hedge would be more desirable if one could be found. Specifically, EMBI-linked derivatives would be very attractive to participants in the emerging-market sovereign credit markets. Based on the results of this study, I expect that the market for these derivative products will become much deeper in future years as issuers and investors continually seek to hedge their exposure to risks beyond their control or the specific motivation of their trade.

References


RETURNING MISPLACED MAIL:
UNDERLYING MOTIVES AND INFLUENTIAL FACTORS

Abdulla Humaidan
University of Warwick

I. Introduction

What do socially-conscious humans have in common with vampire bats? Fortunately, they do not share many qualities! Vampire bats provide an interesting example for reciprocal altruism, which is a vital theory behind the experiment in this paper. These bats feed at night by sucking blood from larger animals. Any bat failing to feed for more than two consecutive nights may starve to death. However, bats that have fed successfully donate blood to the hungry bat. In return, they may receive blood when they are hungry in the future. Altruism could be sacrificing anything from material to immaterial goods to improve others’ welfare. In the case of misplaced mail, one sacrifices time and effort as it involves walking to the mailbox and, in some cases, even buying stamps. This paper links theories of altruism, identifiability, and inter-ethnic reciprocity together by applying them to returning misplaced mail. By dropping envelopes and recording the response rates, I attempt to link between different theories. It would be interesting to explore the effect of inter-ethnic trust, identifiability, time of day, and location on the response rates. Reciprocal altruism suggests that, when returning misplaced mail, the principle would be to help everyone just in case you need a return favour in the future. This can be described as future empathy: one worrying that he is the one losing the envelope in the future.

II. Literature Review

Altruism can be defined as the unselfish concern for other individuals’ welfare. An altruistic act might be making donations to privately provided public goods (e.g. charity). Returning misplaced mail is another example of altruistic acts. However, previous literature argues that “pure altruism” might not be the only reason behind such deeds.

Andreoni’s (1989, 1990) theory of “impure altruism” argues that contributing to public goods generates satisfaction. This became known as the “warm-glow” effect. Building on this, Sugden (1999) provides two reasons why altruism generates the warm-glow. First, the increasing utility of those who contribute is a result of adjusting their behaviour to
match norms of social behaviour that reduce free-riding. Second, individuals gain utility as contributions to public goods allow them to think of themselves as “good people”.

Videras and Owen (2006) use data from 40 countries to show that life satisfaction and happiness are higher for individuals who contribute to environmental protection compared to those who do not. Perhaps more important, is their finding that utility does not increase proportionally with contributions. Hence, this is consistent with the warm-glow theory, as it implies that the act of giving is what generates utility. This relates to my experiment as it suggests that even returning one misplaced envelope only is sufficient to increase the individual’s utility.

Koch and Normann (2005) examine whether external forces are required to motivate dictator giving or if internal forces are sufficient. In order to guarantee experimenter anonymity, their experiment followed the Hoffman et al. (1994, 1996) procedure. They compare the outcomes from a standard double blind dictator game and a treatment where recipients do not even know that they participate in an experiment. They conclude that there are no significant differences between the two treatments. This suggests that dictators who give are purely internally motivated and are not influenced by external pressure. In other words, it does not matter to an individual how others view his actions.

This pure regard for others suggests that an individual will not care about how the “victim” perceives his actions. Consequently, building up on this outcome, I do not ask individuals to attach their details when returning an envelope. Having looked at previous literature describing the sources and motives behind altruism, it is also important to use other literature to hypothesise the outcome when testing different factors.

Schelling (1968) was one of the early economists to acknowledge that identifiable victims stimulate stronger emotional response than statistical victims do. This has since come to be known as the identifiable victim effect. Identifiable victims are specific victims of misfortune and they often draw extraordinary attention. While all “victims” in my experiment are “identifiable”, subjects are not permitted to communicate with them, unless they “choose” to send a note along with the returned envelope. Hoffman et al. (1996) concluded that allowing communication between individuals enhances other-regarding behaviour as it reduces the apparent social distance between them.

Loewenstein and Small (2003) used two studies to show that determining the victim increases caring. In a laboratory study, individuals were found to be more willing to compensate others who lost money when the “losers” were already determined. Additionally, in a field study, people contributed more to charities when their contributions would benefit families that were already selected from a list rather than when they were going to be selected from the same list.

Nevertheless, manipulating identifiability without changing other factors is difficult. The biggest challenge is that identifying victims usually includes revealing important information about them, such as gender or ethnicity. It is possible that people feel especially sympathetic to individuals with those characteristics. Consequently, it is possible that those characteristics rather than identifiability per se are responsible for the significant differences observed. Therefore, I attempt to address this limitation by testing whether
including a title (e.g. Dr.) makes a significant difference or not. While all four “victims” are male, I am also testing the effect of inter-ethnic reciprocity.

Trust plays a crucial part in society, especially in economic interactions. Bouckaert and Dhaene (2002) focus on ethnicity as a potential determinant of interpersonal trust and reciprocity. They adapt an experimental procedure, in which they investigate the average trust and reciprocity between small business entrepreneurs from Turkish and Belgian ethnic origins in Belgian City of Ghent. To avoid effects of sexual discrimination, all subjects were male. They conclude that both average trust and average reciprocity are independent of ethnic origins. However, as each “victim” in my experiment signals a certain ethnic origin, I will look for significant differences in results.

My field experiment attempts to link different theories of impure altruism, identifiability, and inter-ethnic reciprocity together by applying them to returning misplaced mail. Additionally, it will provide further empirical evidence to the theories mentioned above.

III. Variables and Expectations

It is essential to pose several questions about various empirical ideas that might influence one’s decision whether to return an envelope or not. For instance, does the location of the “misplaced” envelope affect one’s decision? Some might argue that an envelope dropped closer to the mailbox is more likely to be returned as it involves less hassle. However, is it the altruistic act itself or the extent of charitable giving that matters? Normally it is the act of giving (rather than extent) that generates utility. Therefore, after returning an envelope, an individual’s utility would increase regardless of the location it was misplaced.

I will also investigate the effect of “identifying” the recipient. In my experiment, I manipulate identifiability by including the title “Dr.” on half of my sample (i.e. 80 envelopes). This reveals a certain characteristic and consequently, I expect the response rate to be higher for envelopes including a title.

Another empirical idea is whether the ethnicity of the recipient matters. I will examine Bouckaert and Dhaene’s conclusion that average reciprocity is independent of ethnic origin. I expect a relatively higher response for English and Chinese recipients, as there is a higher number of English and Chinese students studying in the University of Warwick.

Many argue that islamophobia has spread widely in recent years. To assess this claim, I include 2 Arabic names: Fadi Al-Zain and Ahmed Al-Abdullah. While both names are clearly from the same ethnic origin, only the latter clearly reveals the religion of the recipient (Islam). In this case, some might expect that revealing this characteristic will lead to a negative reaction.

The final factor I will be testing is the time of day. As more people are on campus in the morning, when the majority of lectures take place, the probability of envelopes being returned is higher.
IV. Experimental Design and Procedures

First, it is crucial to note that my experimental treatment guarantees experimenter anonymity by being similar to the standard Hoffman et al. (1994, 1996) procedure. The whole experiment was conducted in the University of Warwick’s campus.

I will “misplace” 160 stamped, addressed envelopes. I use second-class postage on all envelopes, which all have my on campus address (see circle in the map included in the appendix). After misplacing the envelopes, I wait for 10 days and record whether they get delivered back or not.

The experiment is divided into 10 different weekdays, with 16 envelopes being misplaced every day. This is in order to avoid “weekend effects”, if any. These days are spread apart over a month, with an average of 2-3 “experiment days” per week.

All the addresses and “victim” names will be handwritten instead of typed to signal that the contents of the envelope are “personal”. I am assuming that “personal” letters are more likely to be returned when lost than “generic” mail. As I am writing all addresses myself, the problem of inconsistent handwriting is rendered superfluous.

Consistency is a vital part of any experimental procedure. Therefore, controlling all possible variables is essential. To do so, all envelopes will contain 2 index cards to appear more realistic and personal. From an administrative perspective, these cards are used to record when and where the envelope was misplaced. Additionally, identical envelopes will be used.

Various representations of the breakdown of envelopes are included in the appendix. As discussed previously, I will misplace envelopes at 2 different locations, which are selected based on their distance from the nearest mailbox. All the “Near” envelopes will be misplaced in the proximity of Costcutter (circle on the map included), which is less than 100 metres away from the mailbox. On the contrary, all the “Far” envelopes will be misplaced in the proximity of the library (circle). Individuals are expected to return the envelope to the mailbox (green circle) rather than to Heronbank (circle).

To investigate the effect of “identifiability”, a title (“Dr.”) is included on 80 envelopes. The remaining 80 envelopes (or 20 envelopes per name) will not have a title included.

I will divide the 160 equally amongst four “victims”, corresponding to 40 envelopes per name (see appendix for more details). Avoiding the possibility of sexual discrimination, all the recipients are male. Additionally, all names are fictitious and represent certain ethnicities:

- Ahmed Al-Abdullah (Arabic; Muslim)
- Alex Robinson (English; unidentifiable religion)
- Fadi Al-Zain (Arabic; unidentifiable religion)
- Zhi Chen (Chinese; unidentifiable religion)

As can be seen above, Ahmed and Fadi are clearly Arabic, but only the former is clearly Muslim. Comparing the response rates for both names would provide an insight into whether revealing the religion has a statistically significant effect in this case or not.

The time of day might affect the probability of the envelopes being picked up and,
eventually, returned. Therefore, I will perform the experiment at different times of day, which I divide into 2 categories for simplicity: morning and evening. “Morning” is defined as 8 – 11 am while “evening” is defined as 7 – 10 pm.

V. Data

Out of the 160 envelopes I misplaced, a total of 100 envelopes were returned. Therefore the overall probability that a misplaced envelope is returned is 62.5%, which is quite impressive. All envelopes returned either come back through Royal Mail (RM) or through the University’s internal mail. 79 of the 100 envelopes returned were delivered by RM.

Figure 1 shows the average response rates for the 10 experiment days. As can be seen, there is a wide range in rates, as the maximum is 93.75% and the minimum is 37.5%. In this case, the mean is equal to the median (62.5%), suggesting that the observations are not skewed in any particular direction:

![Figure 1: Response Rates on Different Days](image)

The obvious downward trend suggests the worrying possibility that individuals might have started getting suspicious of the envelopes. As a result, some might decide not to participate. Figure 1 is reproduced with a trend line below (Figure 2). The estimated slope of the trend line, -0.58, is the change in the number of envelopes returned per experiment day. That is, 1 less envelope is returned every 2 experiment days. However, the estimated trend line merely reflects the average change across the whole time period. Namely, 7 observations are above the trend line, while only 3 are below.
Figure 3 shows the average time it took for different sub samples of envelopes to be returned. Interestingly, as RM stamps all envelopes upon receipt, I can tell how many days individuals take to return the misplaced envelopes to the mailbox (“Time to Return”), assuming the mailbox is emptied frequently (which it is). Time to return is the most important in this case as it signals the willingness of individuals:

“Total Time” is the Time to Return plus the time it takes the post office to deliver
(“Post Office Delivery” in the Figure 5). Surprisingly, given the time of day, it takes longer for the “near” envelopes to be returned than “far” envelopes.

Figure 4 shows the trend of average Time to Return and Total Time across the experiment days. Again, this fuels any fears that individuals started getting suspicious. Time to Return is obviously growing from 31 January until 12 February, implying that individuals are taking longer to return the envelopes:

![Figure 4: Average Times for Different Dates](image)

Figure 5 shows the Post Office Delivery time across the 10 experiment days. It is imperative to mention that RM sent a note enquiring about the envelopes on 5 February. The average delivery time falls sharply to 2.10 days the next day and stays at low levels for a few days.

In order to further analyze the common trends in individual behaviour, it is helpful to look at Figures 6 and 7. While both reflect frequencies, the former shows what number of envelopes (observations) falls under each discrete value of Time to Return. Similarly, the latter plots the number of envelopes returned within a certain Time to Return.

The mean Time to Return is less than a day, specifically 0.97 days. This can be seen clearly as the majority of envelopes (49 + 27 = 76) are returned within one day.

However, excluding the 11 observations that took at least 3 days to be returned gives a mean of approximately 0.6, which is significantly different from 0.97, even at the 1% significance level. Additionally, 8 of these 11 envelopes were misplaced in the last 5 experiment days. On the other hand, 28 of the 49 envelopes (57.14%) returned within the same day were misplaced in the first 5 experiment days! The question here remains: did individuals become aware of the experiment?
The average Total Time taken for the envelopes to be returned is 2.76 days and can also be seen in Figure 7. This is a fair result as second-class postage has a delivery estimate of 2-3 days:
VI. Testing Statistical Significance of Data

Statistical significance in this experiment would be a significant difference in proportions (response rates). In this section, I briefly outline the main steps before highlighting my results.

Table 1 shows the total number of envelopes returned for each “victim” and the results of testing for statistically significant differences between corresponding response rates (envelopes returned out of 40) and the overall average (62.5%):

<table>
<thead>
<tr>
<th>Victim</th>
<th>Envelopes returned</th>
<th>Victim vs. Average</th>
<th>Significance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{p}_0$</td>
<td>z-value</td>
<td>1%</td>
</tr>
<tr>
<td>Ahmed (AA)</td>
<td>17</td>
<td>0.59</td>
<td>-2.30</td>
</tr>
<tr>
<td>Alex (AR)</td>
<td>29</td>
<td>0.65</td>
<td>1.18</td>
</tr>
<tr>
<td>Fadi (FZ)</td>
<td>31</td>
<td>0.66</td>
<td>1.78</td>
</tr>
<tr>
<td>Zhi (ZC)</td>
<td>23</td>
<td>0.62</td>
<td>-0.58</td>
</tr>
</tbody>
</table>
Hence, we have two extremes: Ahmed’s response rate is too low (rejected at 5% significance level) and Fadi’s is too high (rejected at 10% significance level).

Consider the difference between the response rates for Ahmed Al-Abdullah (AA) and Alex Robinson (AR). Underlying distributions are Bernoulli distributions as they take value 1 with “success” (probability = \( p \)) and value 0 with “failure” (probability = \( 1-p \)). However, as the number of observations is large, the distribution of sample means will be approximately normally distributed.

In the example provided, the test statistic lies to the left of the critical value. Therefore, the null hypothesis is rejected even at the 1% significance level, indicating a statistically significant difference between AA and AR. Table 2 summarizes the results of the remaining hypothesis tests:

<table>
<thead>
<tr>
<th></th>
<th>( \hat{p} )</th>
<th>z-value</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA vs. AR</td>
<td>0.58</td>
<td>-2.71</td>
<td>Reject ( H_o )</td>
<td>Reject ( H_o )</td>
<td>Reject ( H_o )</td>
</tr>
<tr>
<td>AA vs. FZ</td>
<td>0.60</td>
<td>-3.20</td>
<td>Reject ( H_o )</td>
<td>Reject ( H_o )</td>
<td>Reject ( H_o )</td>
</tr>
<tr>
<td>AA vs. ZC</td>
<td>0.50</td>
<td>-1.34</td>
<td>( H_o )</td>
<td>( H_o )</td>
<td>( H_o )</td>
</tr>
<tr>
<td>AR vs. FZ</td>
<td>0.75</td>
<td>-0.52</td>
<td>( H_o )</td>
<td>( H_o )</td>
<td>( H_o )</td>
</tr>
<tr>
<td>AR vs. ZC</td>
<td>0.65</td>
<td>1.41</td>
<td>( H_o )</td>
<td>( H_o )</td>
<td>( H_o )</td>
</tr>
<tr>
<td>FZ vs. ZC</td>
<td>0.68</td>
<td>1.91</td>
<td>( H_o )</td>
<td>( H_o )</td>
<td>( H_o )</td>
</tr>
<tr>
<td>With Title vs. Without</td>
<td>0.63</td>
<td>-0.65</td>
<td>( H_o )</td>
<td>( H_o )</td>
<td>( H_o )</td>
</tr>
<tr>
<td>Morning vs. Evening</td>
<td>0.63</td>
<td>1.31</td>
<td>( H_o )</td>
<td>( H_o )</td>
<td>( H_o )</td>
</tr>
<tr>
<td>Near vs. Far</td>
<td>0.63</td>
<td>0.98</td>
<td>( H_o )</td>
<td>( H_o )</td>
<td>( H_o )</td>
</tr>
</tbody>
</table>

**VII. Results**

Comparing the results with the expectations can help me creatively compare my findings with the existing literature. Through this comparison, I also highlight the additions of my results to the literature. While slightly more “near” envelopes (65.63%) were returned than “far” envelopes (59.38%), there is no significant difference between the two response rates. This coincides with Videras and Owen’s finding that the act of giving (rather than the extent) is what generates utility.

The results from examining the effect of identifiability sometimes contradict what the literature suggests. Loewenstein and Small observed that people might feel especially sympathetic to individuals with certain characteristics. Through my experiment, I revealed 3 different characteristics about the recipients: title, ethnicity, and religion. There is no significant difference between the response rates after adding “Dr.” before the recipient’s name (65%) or not (60%).

Testing whether the ethnicity of the victim affects the response rates yielded mixed results. First, while not always significant, average response rates for all 4 names were not independent of ethnic origin. This contradicts with Bouckaert and Dhaene’s conclusion. Second, Fadi Al-Zain has the highest response rate. Given that the number of Arabic stu-
dents in the University of Warwick is relatively small, the underlying motive for returning the envelopes has to be other than inter-ethnic reciprocity.

A reasonable justification for such result could be reverse psychology, which is also known in psychology as reactance. That is, individuals might reconsider their first decision if it simply means they are complying with society’s “views.”

Revealing the religion of the recipient appears to have a significant effect. The difference between the average response rates for Fadi Al-Zain (77.5%) and Ahmed Al-Abdullah (42.5%) is statistically significant even at the 1% level. This suggests that revealing the religion led people to react to the characteristic, instead of identifiability per se. However, it is crucial to note that the literature focuses on different applications of identifiability. This could justify why my findings are not always consistent with what the literature describes.

The final variable I tested was the time of day. As expected, more “morning” envelopes (67.5%) were returned than “evening” (57.5%). However, this difference is not statistically significant. Interestingly, this suggests that the probability of a misplaced envelope being returned does not necessarily depend on the number of people on campus.

VIII. Discussion

Due to budgetary and time constraints, my experiment was quite limited. In this section, I will highlight some shortcomings and discuss what improvements could be done in the future. I will also include possible extensions of my work.

As I dropped all 160 envelopes in the University of Warwick’s campus, the sample was limited to people on campus. The University’s records show that there are 15,969 full-time students and 4,921 staff members. Although highly unlikely, let us make the simplistic assumption that all students will be on campus at the same time. Additionally, we will unrealistically assume that each student only encounters one envelope. That is, no individual participates more than once. In this case, misplacing 160 envelopes corresponds to approximately 1% of students participating in the experiment. To address this issue, I suggest misplacing more envelopes in the future. Additionally, the experiment should not be limited to the campus only. For instance, performing the experiment in ethnic quarters (e.g. Chinatown or Little Italy) in nearby cities (e.g. London) might give alternative results.

The experimental design could be improved to enhance the quality of results. For example, with a bigger sample, the results would be more concrete. Other ways of improvement include using more names, locations, etc. For instance, examining the differences in response rates for English, Irish, and Scottish recipient names could potentially yield interesting results. Another possibility is to use a different address than Heronbank, which is the student accommodation I currently reside in. Furthermore, it is interesting to study the effect of using first-class or second-class postage.

The contents of envelopes could also be changed. For example, including coins in the
envelopes would probably lead to significantly lower response rates than otherwise, as individuals might overestimate the amount of money inside the envelope. Another factor is the envelope’s weight. For instance, including more coins would make the envelope slightly heavier. It is useful to note that including paper in the envelope could signal the presence of either a check or a letter, leaving some uncertainty.

As mentioned previously, all addresses and victim names were handwritten. I simply assumed that when the name and address are both handwritten, that signals it is “personal” mail. I also assumed that misplaced “personal” mail is more likely to be returned than misplaced “generic” mail. A possible extension is to scrutinize this assumption by comparing the response rates for handwritten and typed addresses and recipient names. Also, when choosing the four names, I attempted conveying a certain ethnicity and/or religion. Justifiably, I assumed that all individuals would easily recognize the victim’s ethnicity and/or religion. However, to make sure this is the case, a questionnaire could be distributed.

With the 10 experiment days spread over less than a month, the experiment was completed over a relatively short period of time. Spreading the experiment over a longer time period means there is a higher probability of international events changing the results, leading to more interesting, robust, and credible results. Furthermore, the experiment days were randomly selected. The lack of attention for weather or different events taking place on campus provides a possible extension of my work. Additionally, the dates can be chosen in order to examine whether the One World Week, which annually takes place on campus, leads to statistically significant differences.

While analyzing the results, several questions remain unanswered. For instance, who exactly is returning the envelopes? What is their ethnicity, age, sex, etc.? Another unknown is the number of people ignoring envelopes; I only knew how many envelopes were returned. By observing what happens after dropping an envelope, one can record different details about individuals, regardless of whether they return it or not. This way, one can also know whether an envelope is always returned after it is picked up or not.

Observing what happens after dropping an envelope permits investigation of the difference between the responses of the two categories of students: “home” and overseas. Given a total of 15,969 students and 4,086 overseas students, there are 11,883 home students. Without witnessing what happens, it would be challenging to divide the two categories. However, this method requires a lot of time and effort. Moreover, recorded details might not always be accurate.

Another possibility is to randomly select names from UK phone directories, post envelopes to their addresses, and ask them to return the envelopes. Consequently, it would be easier to get accurate details. If they do return the envelope, a follow-up questionnaire could be used to find out whether they report higher levels of life satisfaction and happiness or not. If so, what generates this utility exactly?

Therefore, as evident from the discussion above, there are several shortcomings in my experimental design. As a result, there is a lot of potential for further work. However, my robust findings still provide a valuable insight into the influential factors affecting one’s
decision whether to return a misplaced envelope or not.

IX. Conclusion

In a nutshell, through an application to returning misplaced mail, this paper linked the-ories of reciprocal altruism and inter-ethnic reciprocity together. The experiment in this paper involved misplacing envelopes and recording the response rates. While the results coincide with some findings in previous literature, they deviate from others. For instance, I conclude that it is the altruistic act itself that matters. Evidence for effects of identifiability was mixed. Unlike the literature, I cannot conclude that response rates were entirely independent of recipient ethnicities, the effects of which were not always significant. Additionally, revealing a characteristic by adding a title did not appear to significantly increase response rates. However, revealing the religion did make a significant difference, although negative. A plausible theory is that it caused individuals to feel less sympathetic with the “victim.” Thus, with 62.5% of misplaced mail eventually returned, regardless of various factors, it appears that humans are actually comparable with vampire bats in terms of reciprocal altruism!

References


Bouckaert, J. and G. Dhaene (2002). “Inter-Ethnic Trust and Reciprocity: Results of an Experiment with Small Business Entrepreneurs.”


THE IMPACT OF STOCK MARKET LIQUIDITY ON ECONOMIC GROWTH: AN EMPIRICAL ANALYSIS

Angelica da Rocha
University of Warwick

ABSTRACT

This project studies the effect of stock market liquidity on economic growth using panel data analysis to 64 countries between 1988 and 2005. The paper finds that stock markets positively affect growth and in particular this relationship changes in magnitude depending on whether the countries are developed or under development. The findings suggest diminishing returns to liquidity as a country becomes more developed.

I. Introduction

Does stock market liquidity exert a significant impact on long-run economic growth? The liquidity of a market refers to the extent to which it is continuous (without large price changes between trades), and deep (with many buyers and sellers willing to trade at the prevailing price) (Bhide 1993). In a liquid equities market it is relatively cheap to trade stocks, and there is relatively little uncertainty about the price, timing and settlement of the trades (Levine 2003). The theory is ambiguous on the effect of stock market liquidity on the growth of the economy. Some models suggest that a liquid stock market is favourable to resource allocation and economic activity since it lowers the cost of investing in long-run projects. An alternative view is that higher returns achieved from better resource allocation may reduce savings due to income effects and have an adverse impact on economic growth (Bencivenga and Smith 1991). Some authors also argue that the direction of causality is reversed, that is economic growth brings more liquidity to the markets (Robinson 1952). Although the theories of the adverse impact of stock market liquidity on economic growth seem reasonable, I could not find any empirical evidence to support such
Empirical evidence of the importance of stock market liquidity on economic growth will provide guidance to policymakers in the design of suitable changes to the financial system. Recent empirical literature has approached the matter from either a macro perspective, with the use of cross-country growth analysis or sophisticated panel data techniques; or a micro standpoint, where industry- and firm-level data have been used to deal with causality issues between economic growth and financial development.

Beck and Levine (2004) empirically examine the relationship between economic growth and both stock market liquidity and bank development. They use a panel of 46 countries over the period between 1976 and 1998 with data averaged over five years. The study controls for other determinants of growth and treats simultaneity and omitted variable bias. It uses the system of Generalised Method of Moments (GMM) estimator, which is a sophisticated econometric technique that combines first-differenced GMM with an additional set of equations in levels with suitably lagged first-differences as instruments. The paper finds that both stock markets and banks are jointly important for growth. Their results, however, are not entirely consistent to changes in specifications and econometric techniques.

Rousseau and Wachtel (2000) also empirically assess the importance of stock market and banks for economic growth. They employ panel techniques to annual data over 1980-1995 for 47 countries. As a measure of stock market liquidity they use two different proxies: value of the traded shares divided by GDP and market capitalization of all shares traded on the main stock exchange of a given country divided by GDP. To measure bank development they use M3 divided by GDP. First-differenced GMM estimator is used to analyse the relationship between the variables. Using slightly different proxies to stock market liquidity and bank development from Beck and Levine (2004) and different econometric methods, they find strong evidence that liquid stock markets have a significant positive effect on growth.

This paper extends the cross-country and time dimension of previous work using a panel data set with annual data over the period between 1988 and 2005 across 64 countries. It adds to the literature by exploiting the heterogeneity of the overall panel and investigating differences between developed and developing economies. Rousseau and Wachtel (2000) state (but do not investigate) that stock market liquidity is particularly important in emerging economies since it raises the confidence of investors in the value associated with information and risk diversification when trading on an organised exchange. I find evidence for such a statement.

I use the first-differenced Generalised Method of Moments (GMM) estimator suggested by Arellano and Bond (1991) and compare it with the Fixed Effects (FE) estimator. By and large both methods yield the same conclusion that stock market liquidity does affect economic growth. I show that both econometric techniques have advantages and disadvantages yielding slightly different results. The use of instrumental variables (IV) in the GMM allows for consistent estimation even in the presence of measurement error. However when the time series are persistent and the number of time periods (T) is small the first-
differenced GMM is biased. This is because lagged levels of the variables are weak instruments in those cases. The FE estimator, on the other hand, is much simpler and straightforward, however dynamic panel estimation using the FE estimator with $T$ small leads to bias. Fixed Effects estimator also does not deal with potential endogeneity problems. I shall assume that $T=18$ is sufficiently large to reduce the bias in the FE estimates.

I split the analysis in: (i) overall, where I examine the effect of stock market liquidity on economic growth including all the 64 countries, controlling for other growth determinants and considering the potential effect of bank development on growth; (ii) developed, in which I carry out similar analysis but only including the 21 developed countries in the sample; and (iii) developing, where comparable investigation is carried out with the 43 remaining developing countries.

In the overall analysis I find that stock market liquidity is important to economic performance, the results are robust to changes in econometric methods but not to changes in specifications. Similar results are found for developing countries. In the analysis of the developed countries I find liquidity to be statistically insignificant at the 5% significance level in all economic growth regressions. One may argue that the overall results may be driven by developing countries.

II. Data

I investigate the empirical relationship between stock market liquidity and economic growth in an unbalanced panel of 64 countries from 1988 to 2005. The choice of countries and time periods were essentially based on availability of data for the proxy to stock market liquidity.

The variables used in the analysis are largely divided into two groups:

- Control variables
  - Informational set
  - Policy set
- Study variables

In the ‘control variables’ group are the variables that potentially affect economic growth. The ‘informational set’ contains variables that control for convergence and human capital accumulation, i.e. lagged real GDP per capita and secondary school enrolment. The ‘policy set’ controls for economic factors that could potentially affect growth, such as, inflation rate, government expenditure and trade openness. As the name suggests the ‘study variables’ group includes the variables that I intend to investigate, i.e. stock market liquidity. Rousseau and Wachtel (2000) suggest that any study of the relationship between stock market liquidity and economic growth should simultaneously consider the effect of bank development. I also add a proxy to bank development to the study group to prevent stock market liquidity to capture the potential effect of banks on economic growth.

I measure stock market liquidity as the turnover ratio that is the ratio of the total value
of domestic shares to market capitalization of listed companies. This ratio captures the value of traded domestic stock relative to the size of the country’s stock market and has been used with success in previous research such as Beck and Levine (2004). High turnover is often used as an indicator of low transaction costs (Levine and Zervos 1998). Liquid stock markets reduce disincentives to long-run investments, since liquid markets provide a ready exit-option for investors. This can foster more efficient resource allocation and faster growth (Beck and Levine 2004).

According to Levine and Zervos (1998), researchers would ideally have cross-country measures of how well banks identify profitable activities, exert corporate governance, mobilise resources, manage risk, and facilitate transactions. Economists, however, have not been able to accurately measure bank development for a broad cross-section of countries. For this reason I use similar measure to Levine and Zervos (1998) that is domestic credit to private sector as a ratio to GDP. It refers to financial resources provided to the private sector, such as through loans, purchases of non-equity securities, and trade credits and other account receivables that establish a claim for repayment.

Table 1 presents the overall summary statistics of GDP per capita growth (DLGDPPC), stock market liquidity (TURNOVER) and bank development (BANK) and the overall correlations between those variables. The heterogeneity of the sample is evident by the extremely high standard deviation on all the three variables.

<table>
<thead>
<tr>
<th>TABLE 1: OVERALL SUMMARY STATISTICS AND CORRELATIONS 1988-2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLGDPPC</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
<tr>
<td>No observations</td>
</tr>
<tr>
<td>Correlation</td>
</tr>
<tr>
<td>DLGDPPC</td>
</tr>
<tr>
<td>TURNOVER</td>
</tr>
<tr>
<td>BANK</td>
</tr>
</tbody>
</table>

p-values are in brackets

I find that while Korea in 2000 had a turnover of approximately 622% of the total market capitalization of domestic listed companies, Barbados in the same year had a turnover ratio of only 0.4%. Such a high turnover ratio for Korea in 2000 raises three questions: (i) is Korea’s stock market extremely liquid? (ii) is the market capitalization of the domestic listed companies very small? (iii) are there measurement errors in the data? A brief analysis of Korea’s stock market history shows that the average turnover ratio in Korea over the period in analysis is approximately 200%. In 1999 turnover was around average and in
2001 it was half of what I found for 2000. Other years seem to be consistent with the history of turnover ratios, suggesting the existence of inconsistency in the data. The inconsistency of some of the data collected is also evident in the descriptive statistics for economic growth. Australia in 1997 shows the highest per capita GDP growth rate in the sample with almost 21%. The previous year Australia showed growth of 2.4% and in 1998 it shows growth of -13.54%, clearly a discrepancy. Banks in the US lent 260% of GDP to the private sector in 2005, while Polish banks lent only 1.68% of GDP in 1989.

Economic growth is significantly correlated with stock market liquidity and insignificantly correlated with bank development. Stock market liquidity is significantly correlated with bank development.

The sample with only developed countries, expectedly, shows relatively lower variance. The average turnover ratio is approximately 50% greater than the overall turnover. The same is found for banks. The proxies for stock market liquidity and bank development are not significantly correlated with economic growth however they are significantly correlated with each other. The developing countries sample shows a much higher standard deviation. The most disperse series is the turnover ratio with standard deviation twice the mean. The average BANK is a third lower compared to the overall sample. Both turnover and bank development are significantly correlated with growth.

I now pool the observations in OLS regressions. Despite its inefficiency, since the OLS estimator does not consider the structure of the error term when you have a panel data set, the estimates are consistent and give a rough idea of the relationship between economic growth and stock market liquidity. Table 2 presents the results of the overall pooled OLS regressions with standard errors asymptotically robust to heteroskedasticity and serial correlation. Each of the five regressions includes the lag of the log of GDP per capita to control for convergence and the log of secondary school enrolment to control for human capital accumulation. They also include the log of the turnover ratio (LTURNOVER) and the log of bank (LBANK). Each variable in the policy set of variables enter separately regressions (b), (c) and (d) and then together in regression (e). The pooled OLS regressions show a significant positive relationship between stock market liquidity and economic growth. The relationship is robust to change in specifications, remaining significant at the 10% significance level in all regressions and at the 5% significance level in the majority of the regressions. Bank development, on the other hand, is statistically highly insignificant in all regressions. Atje and Jovanovic (1993) also found that bank credit does not significantly affect economic growth.

The lag of GDP per capita enters the regressions significantly negative in most regressions and the size of the coefficient is consistent with previous research (Levine and Zervos, 1998). Secondary-school enrolment enters the regressions with a positive coefficient but insignificantly in most of the regressions, while government expenditure as a ratio to GDP and inflation rate enters the regressions with negative and significant coefficients. The $R^2$ of the regressions seem low when compared to the $R^2$ of other studies (Levine and Zervos, 1998 and Beck and Levine, 2004). However, one should not expect a high $R^2$ when dealing with pooled OLS in a panel data set.
The lag of GDP per capita enters the regressions significantly negative in most regressions and the size of the coefficient is consistent with previous research (Levine and Zervos, 1998). Secondary-school enrolment enters the regressions with a positive coefficient but insignificantly in most of the regressions, while government expenditure as a ratio to GDP and inflation rate enters the regressions with negative and significant coefficients. The $R^2$ of the regressions seem low when compared to the $R^2$ of other studies (Levine and Zervos, 1998 and Beck and Levine, 2004). However, one should not expect a high $R^2$ when dealing with pooled OLS in a panel data set.

Tables 3 and 4 report the results of similar OLS regressions for developed and developing countries, respectively. To avoid repetition I only present coefficient estimates on TURNOVER and BANK.

I find that stock market liquidity is not statistically significant in any of the economic growth regressions for the developed countries. Bank development also appears not to affect growth in the majority of the regressions. The $R^2$ of the regressions are consider-

### Table 2: Overall Pooled OLS

<table>
<thead>
<tr>
<th>Variable</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>0.0590124</td>
<td>0.0623695</td>
<td>0.0285028</td>
<td>0.0927368</td>
<td>0.0380786</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.020)</td>
<td>(0.405)</td>
<td>(0.000)</td>
<td>(0.264)</td>
</tr>
<tr>
<td>L.LGDPPC</td>
<td>-0.0072621</td>
<td>-0.0069333</td>
<td>-0.0056667</td>
<td>-0.0083404</td>
<td>-0.0062179</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.040)</td>
<td>(0.109)</td>
<td>(0.015)</td>
<td>(0.082)</td>
</tr>
<tr>
<td>LSCHOOLa</td>
<td>0.0240984</td>
<td>0.0151667</td>
<td>0.0352996</td>
<td>0.0287037</td>
<td>0.0296548</td>
</tr>
<tr>
<td></td>
<td>(0.174)</td>
<td>(0.404)</td>
<td>(0.048)</td>
<td>(0.103)</td>
<td>(0.101)</td>
</tr>
<tr>
<td>LCHCPFa</td>
<td>-0.0182657</td>
<td></td>
<td></td>
<td></td>
<td>-0.016613</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td>(0.035)</td>
</tr>
<tr>
<td>LGOVEXP</td>
<td></td>
<td>-0.0175249</td>
<td></td>
<td>-0.0172227</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.011)</td>
<td></td>
<td>(0.012)</td>
<td></td>
</tr>
<tr>
<td>LEXPIMP</td>
<td></td>
<td></td>
<td></td>
<td>0.104648</td>
<td>0.0092763</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>LBANK</td>
<td>0.0054477</td>
<td>0.00374</td>
<td>0.0059815</td>
<td>0.0030798</td>
<td>0.0021834</td>
</tr>
<tr>
<td></td>
<td>(0.145)</td>
<td>(0.319)</td>
<td>(0.157)</td>
<td>(0.421)</td>
<td>(0.600)</td>
</tr>
<tr>
<td>LTURNOVER</td>
<td>0.0047029</td>
<td>0.0053186</td>
<td>0.003844</td>
<td>0.0065029</td>
<td>0.0060604</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.028)</td>
<td>(0.078)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.1096</td>
<td>0.1383</td>
<td>0.1377</td>
<td>0.1357</td>
<td>0.1839</td>
</tr>
<tr>
<td>Joint signi-</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>ficancetest (p-value)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

p-values are in brackets

Dummy variables for each year are included in the regression and are not reported.

a This variable is included as log(1+VARIABLE).
ably higher than the overall pooled OLS. Note that, while in some of the regressions none of the explanatory variables are individually significant, they are jointly significant.

**Table 3: Developed Pooled OLS**

<table>
<thead>
<tr>
<th>Variable</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBANK</td>
<td>-0.0047869 (0.218)</td>
<td>-0.005434 (0.153)</td>
<td>-0.0086482 (0.057)</td>
<td>-0.0041365 (0.247)</td>
<td>-0.0082344 (0.013)</td>
</tr>
<tr>
<td>LTURNOVER</td>
<td>-0.0009542 (0.585)</td>
<td>-0.0014375 (0.405)</td>
<td>-0.0012434 (0.528)</td>
<td>-0.0001997 (0.917)</td>
<td>-0.0007945 (0.662)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.1995</td>
<td>0.2172</td>
<td>0.2319</td>
<td>0.2320</td>
<td>0.2660</td>
</tr>
</tbody>
</table>

*p-values are in brackets*

Clearly stock market liquidity is important for the economic performance of developing countries. Turnover ratio is significant at the 1% significance level in all the regressions, while bank development oscillates between significance and non-significance depending on the specifications of the model. While the $R^2$s are greater than the overall regressions they are consistently lower than the regressions of the developed countries.

**Table 4: Developing Pooled OLS**

<table>
<thead>
<tr>
<th>Variable</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBANK</td>
<td>0.008166 (0.063)</td>
<td>0.006086 (0.165)</td>
<td>0.0098644 (0.062)</td>
<td>0.0054964 (0.281)</td>
<td>0.0061812 (0.296)</td>
</tr>
<tr>
<td>LTURNOVER</td>
<td>0.0058672 (0.005)</td>
<td>0.0066119 (0.002)</td>
<td>0.0047665 (0.012)</td>
<td>0.0070445 (0.000)</td>
<td>0.0061779 (0.001)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.1599</td>
<td>0.2000</td>
<td>0.1786</td>
<td>0.1704</td>
<td>0.2241</td>
</tr>
</tbody>
</table>

*p-values are in brackets*

If I take a country like Bangladesh and use regression (a) from Table 2, I find that had Bangladesh had the sample mean turnover ratio (0.4480186) rather than its own average over the time periods (0.2818615), other things being equal, the country would have gained 0.22% on its average growth. A number that is economically significant and consistent with findings by previous research (Levine and Zervos 1998). If instead I use regression (a) from Table 4 and the sample mean turnover ration for developing countries (0.3873718), the gain in average growth is 0.19%.

Note that all the statistical inferences are made on the assumption that the variables are stationary. For that I carry out the Fisher unit root test for unbalanced panels. Fisher tests for panel unit root using an Augmented Dickey-Fuller test. The null hypothesis of a unit root is rejected for all series only including lag differences of the series. A trend needs to be included to reject unit root for LGDP. It is clearly not reasonable to assume that per capita GDP has a constant mean over time, and cointegration of panel data is not trivial. Hence I choose to include dummy variables for each year to allow for common long-run growth in GDP per capita. The time dummies transform the variables into deviation from
time means. Therefore, any arbitrary pattern in the time means is consistent with a constant mean of the transformed series for each country (Bond, S., Hoeflfler, A. and Temple, J. 2001). Although this transformation may not be ideal, and results may be distorted if the assumption is not true, such transformation has been used in recent literature, especially in empirical growth models using system GMM.

III. Methodology

To examine the relationship between stock market liquidity and economic growth in a panel, I use the fixed effects (FE) estimator and the first-differenced GMM estimator. I estimate the relationship using each estimator separately and compare the results.

I chose FE as an estimator as opposed to the random effects (RE) estimator because the country specific effects are likely to be correlated with the explanatory variables, particularly if those individual effects represent omitted variables. Choosing RE estimator in this case would yield inconsistent results. FE estimates are generally biased in a dynamic panel when T is small, however the bias tends to zero as T approaches infinity. The crucial question is ‘How large should T be so that the bias can be ignored?’ There are 18 time periods in the sample. I shall assume that is a sufficiently large T to reduce the bias. The FE estimator by itself does not control for potential simultaneity. To treat such a problem we need to include instrumental variables to the regression. It is also worth noting that measurement error and time-varying omitted variables generally cause contemporaneous correlation between the error term and the explanatory variables which yields inconsistent FE estimates.

Ideally, the econometric methods for estimating dynamic panel data should allow for endogeneity, measurement error and omitted variables. One approach to tackle these problems is the first-differenced GMM but such econometric technique is not on the scope of this paper.

IV. Results

The results in Table 5 show FE coefficients estimates using the whole sample. The standard errors are robust for heteroskedasticity and serial correlation.

The statistical significance of LTURNOVER varies according to which variables are included in the regression. The size of the coefficient is considerably larger than the pooled OLS results (Table 2) for all regressions except regression (e) where it is smaller. The coefficient estimates for LBANK are insignificant for all regressions. L.LGDPPC shows significant negative coefficients while LSCHOOL is statistically insignificant in all regressions. All the variables in the ‘Policy set’ enter the regressions individually significant both when taken one at a time and when they are all together in the same regression. Note that while some of the variables are individually insignificant they are jointly significant in all regressions.
TABLE 5: OVERALL FE ESTIMATES

<table>
<thead>
<tr>
<th>Variable</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Coefficient</td>
<td>Coefficient</td>
<td>Coefficient</td>
<td>Coefficient</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>0.7417408</td>
<td>0.817481</td>
<td>0.6880445</td>
<td>0.8062875</td>
<td>0.8356943</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.000)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>L.LGDPPC</td>
<td>-0.0821673</td>
<td>-0.0911075</td>
<td>-0.0867703</td>
<td>-0.0884989</td>
<td>-0.1026988</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>LSCHOOLa</td>
<td>0.0023089</td>
<td>-0.0303951</td>
<td>0.0150923</td>
<td>-0.0009252</td>
<td>-0.021184</td>
</tr>
<tr>
<td></td>
<td>(0.940)</td>
<td>(0.287)</td>
<td>(0.632)</td>
<td>(0.975)</td>
<td>(0.495)</td>
</tr>
<tr>
<td>LCHCPIa</td>
<td>-0.0184062</td>
<td>-0.0328154</td>
<td></td>
<td></td>
<td>-0.0230274</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.002)</td>
<td></td>
<td></td>
<td>(0.013)</td>
</tr>
<tr>
<td>LGOVEXP</td>
<td></td>
<td>-0.0263033</td>
<td></td>
<td></td>
<td>-0.0448845</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.036)</td>
<td></td>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>LEXPIMP</td>
<td></td>
<td></td>
<td>0.0263033</td>
<td></td>
<td>0.0268686</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.036)</td>
<td></td>
<td>(0.026)</td>
</tr>
<tr>
<td>LBANK</td>
<td>0.0000987</td>
<td>0.0017723</td>
<td>0.0033247</td>
<td>0.0004163</td>
<td>0.0068311</td>
</tr>
<tr>
<td></td>
<td>(0.988)</td>
<td>(0.794)</td>
<td>(0.643)</td>
<td>(0.952)</td>
<td>(0.343)</td>
</tr>
<tr>
<td>LTURNOVER</td>
<td>0.0071005</td>
<td>0.0079275</td>
<td>0.0050813</td>
<td>0.006843</td>
<td>0.004962</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.035)</td>
<td>(0.141)</td>
<td>(0.078)</td>
<td>(0.150)</td>
</tr>
<tr>
<td>Joint signifi-</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
canccetest (p- value) |                      |                      |                      |                      |                      |

p-values are in brackets

Dummy variables for each year are included in the regression and are not reported.

This variable is included as log(1+VARIABLE).

The statistical significance of LTURNOVER varies according to which variables are included in the regression. The size of the coefficient is considerably larger than the pooled OLS results (Table 2) for all regressions except regression (e) where it is smaller. The coefficient estimates for LBANK are insignificant for all regressions. L.LGDPPC shows significant negative coefficients while LSCHOOL is statistically insignificant in all regressions. All the variables in the ‘Policy set’ enter the regressions individually significant both when taken one at a time and when they are all together in the same regression. Note that while some of the variables are individually insignificant they are jointly significant in all regressions.

Tables 6 and 7 report the results of FE estimates for developed and developing countries, respectively. The standard errors are robust for heteroskedasticity and serial correlation.

I find that neither stock market liquidity nor bank development is statistically significant at the 5% significance level in determining economic growth in developed countries. Such results seem reasonable since developed countries generally have much more com-
plex financial markets and the proxy used in this study may not actually reflect the liquidity in those markets.

It is reasonable to say that between 1980 and 1991 the financial markets in developed countries were much less sophisticated than they are today. On the other hand, during the same time period stock markets in developing countries may have contributed very little for economic growth. Nowadays the picture has changed which is reflected on the results found in this paper.

### Table 6: Developed FE Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBANK</td>
<td>-0.0084546 (0.227)</td>
<td>-0.0082627 (0.222)</td>
<td>-0.0066944 (0.279)</td>
<td>-0.0068857 (0.233)</td>
<td>-0.0048441 (0.296)</td>
</tr>
<tr>
<td>LTURNOVER</td>
<td>0.0055198 (0.075)</td>
<td>0.0048397 (0.112)</td>
<td>0.0047984 (0.099)</td>
<td>0.0040184 (0.224)</td>
<td>0.0023371 (0.459)</td>
</tr>
</tbody>
</table>

p-values are in brackets

The results in Table 6 are significantly different, particular for LTURNOVER, from the OLS estimates in Table 3. However both econometric techniques yield the same conclusion that stock market liquidity is not statistically significant for economic growth in developed countries.

In the developing countries LTURNOVER is significant at the 10% significance level in most regressions. These results seem realistic as financial markets in developing countries tend to be much less sophisticated when compared to developed countries. And the proxy used here reflects the stock market liquidity in those countries more accurately than in developed countries. LBANK does not contribute to the economic growth in those countries.

The FE estimates here yield considerably different results from the previous OLS estimates (Table 4), particularly with regards to the p-values. It is worth noting that the OLS estimates treat equally both sources of variation (time and cross-country), it does not take into account the structure of the error term.

### Table 7: Developing FE Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBANK</td>
<td>0.0071735 (0.454)</td>
<td>0.0093917 (0.344)</td>
<td>0.0128691 (0.253)</td>
<td>0.0080317 (0.467)</td>
<td>0.0172973 (0.125)</td>
</tr>
<tr>
<td>LTURNOVER</td>
<td>0.0053107 (0.061)</td>
<td>0.0054055 (0.075)</td>
<td>0.003491 (0.187)</td>
<td>0.0051518 (0.080)</td>
<td>0.0029109 (0.251)</td>
</tr>
</tbody>
</table>

p-values are in brackets

One may question the direction of the causality of the relationship. The FE estimator does not address the potential problem of endogeneity. I test for reverse causality and find
that there is simultaneity in the relationship between LTURNOVER and DLGDPPC. To address such problem suitable instrumental variables would need to be included in the FE regression.

FE estimates using LTURNOVER lagged two periods as instrument (not shown) are very similar to the results presented above. The p-values on LTURNOVER are lower in the estimates using instrumental variables (IV) than the ones shown above for the developing countries and higher for the developed countries, presenting further support to our findings. LBANK also seems to contribute to the economic growth of developing countries, showing high levels of significance in the majority of the regressions using IV.

The first-differenced GMM estimator addresses the potential endogeneity, measurement error and omitted variables problems. The results in Table 8 show GMM coefficients estimates using the whole sample. The standard errors are robust for heteroskedasticity and serial correlation. Again with this more sophisticated econometric technique I find a significant positive relationship between stock market liquidity and economic growth. The size of the coefficient on LTURNOVER is significantly larger than the results found using the FE estimator. LBANK is again found to be statistically insignificant in all regressions.

<table>
<thead>
<tr>
<th>Variable</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>0.0025274 (0.488)</td>
<td>0.0049781 (0.015)</td>
<td>0.0052354 (0.017)</td>
<td>0.005702 (0.020)</td>
<td>0.0050459 (0.011)</td>
</tr>
<tr>
<td>L.LGDPPC</td>
<td>-0.3175145 (0.000)</td>
<td>-0.32517640 (0.000)</td>
<td>-0.3150804 (0.000)</td>
<td>-0.3218584 (0.000)</td>
<td>-0.32059 (0.000)</td>
</tr>
<tr>
<td>LSCHOOLa</td>
<td>0.0439783 (0.367)</td>
<td>0.0484926 (0.265)</td>
<td>0.0374538 (0.497)</td>
<td>0.0376962 (0.437)</td>
<td>0.0418292 (0.394)</td>
</tr>
<tr>
<td>LCHCPIa</td>
<td>-0.0255552 (0.035)</td>
<td></td>
<td>-0.0284138 (0.032)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGOVEXP</td>
<td></td>
<td>-0.0576569 (0.072)</td>
<td></td>
<td>-0.0601183 (0.036)</td>
<td></td>
</tr>
<tr>
<td>LEXPIMP</td>
<td></td>
<td></td>
<td>-0.0165373 (0.332)</td>
<td></td>
<td>-0.0072174 (0.692)</td>
</tr>
<tr>
<td>LBANK</td>
<td>-0.007997 (0.366)</td>
<td>-0.0018002 (0.821)</td>
<td>0.0096009 (0.684)</td>
<td>-0.0091265 (0.407)</td>
<td>0.0027522 (0.784)</td>
</tr>
<tr>
<td>LTURNOVER</td>
<td>0.0115553 (0.035)</td>
<td>0.0121917 (0.026)</td>
<td>0.0052354 (0.082)</td>
<td>0.0112642 (0.047)</td>
<td>0.0096851 (0.072)</td>
</tr>
<tr>
<td>AR(1) p-values</td>
<td>0.0094</td>
<td>0.0091</td>
<td>0.0094</td>
<td>0.0106</td>
<td>0.095</td>
</tr>
<tr>
<td>AR(2) p-values</td>
<td>0.9865</td>
<td>0.8941</td>
<td>0.9631</td>
<td>0.9359</td>
<td>0.8730</td>
</tr>
</tbody>
</table>

p-values are in brackets
Dummy variables for each year are included in the regression and are not reported.
The consistency of the GMM estimator depends on the validity of the instruments and the assumption that the error terms are not serially correlated. The Sargan test of over-identifying restrictions tests the validity of the instruments, under the null hypothesis that the instruments used are valid instruments. The test fails to reject the null hypothesis giving support to the model. To address the issue of serial correlation I test the null hypothesis that the error term is not serially correlated. The test rejects the hypothesis of no serial correlation of the AR(1) type and fails to reject the null hypothesis of no serial correlation of the AR(2) type. This is the case because by construction the differenced error term is generally first order serially correlated even if the original error term is not. Both tests (the Sargan and the serial correlation tests) give support to the GMM model. I shall express concern about the power of the Sargan test.

I also run a GMM regression separately for the developed and developing countries (Tables 9 and 10). The results from Table 9 are somewhat different from the results obtained using the FE estimator (Table 6), particularly for LTURNOVER. The p-values on LTURNOVER are closer to the ones obtained in the pooled OLS regressions (Table 3). However all three econometric techniques come to the same conclusion: LTURNOVER does not explain economic growth in developed countries.

### Table 9: Developed GMM Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBANK</td>
<td>-0.0008357</td>
<td>-0.0012302</td>
<td>0.0010853</td>
<td>-0.0048468</td>
<td>-0.0012592</td>
</tr>
<tr>
<td>LTURNOVER</td>
<td>0.0006282</td>
<td>0.0003807</td>
<td>0.0000572</td>
<td>-0.0002038</td>
<td>-0.0018666</td>
</tr>
</tbody>
</table>

p-values are in brackets

In the analysis of the developing countries, I find the coefficient on LBANK to be insignificant in all regressions. LTURNOVER is significant in most regressions and the coefficients are significantly larger than the FE estimates. The conclusion again is that stock market liquidity does affect economic growth positively.

### Table 10: Developing GMM Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBANK</td>
<td>-0.0039391</td>
<td>0.0087073</td>
<td>0.0028322</td>
<td>-0.0048729</td>
<td>0.0205851</td>
</tr>
<tr>
<td>LTURNOVER</td>
<td>0.0076072</td>
<td>0.007318</td>
<td>0.0057878</td>
<td>0.0071942</td>
<td>0.004932</td>
</tr>
</tbody>
</table>

p-values are in brackets
V. Conclusion

All in all, the results point to the direction of a positive relationship between stock market liquidity and economic growth. The level of significance of this relationship varies depending on the econometric technique used and control variables included in the regression.

The sub-sample analysis suggests no effect of stock market liquidity on economic growth in developed countries and a positive relationship in developing countries. A question remains whether the overall results has been driven by developing countries.

The issue of reverse causality needs to be addressed in a more formal manner, perhaps using system GMM to sub-samples of developing and developed countries. Regional differences within developed and developing countries could possibly be addressed in the future with the introduction of dummies. Future research should focus on more accurate measures of stock market liquidity. Liquidity in other securities market, such as bonds should also be addresses in future analysis. A study that uses annual data, such as this one, is particularly influenced by the business cycle, especially because data for TURNOVER was only available from 1988, right after 1987 stock crash, when economies were probably close to the bottom of the cycle. Future analysis may want to take that into consideration.

References


INTRAHOUSEHOLD BARGAINING AND DIVORCE NORMS IN INDONESIA

Nicholas Hartman
Georgetown University

ABSTRACT

This paper develops a model that tests whether customary laws related to divorce are a useful proxy for relative power within the household. I hypothesize that a more favorable distribution of assets in the event of household dissolution improves an individual’s bargaining power, thereby increasing that individual’s influence over the allocation of household resources during marriage. I use household and community level survey data from Indonesia to test empirically the effect of divorce norms on a series of household welfare indicators which I posit are sensitive to bargaining between husbands and wives. Results suggest that divorce norms play a greater role in determining a household’s food expenditures compared to expenditures on clothing and education.

I. Introduction

Traditional models treat the household as a homogenous unit with a single set of preferences that determine a myriad production and consumption decisions. These models assume either that all individuals within the household have identical preferences or that one individual has complete control over the household decision-making process. This unitary model of the household has long been used across the social sciences. However, in recent years alternative models, coupled with empirical evidence, have called into question the validity of the unitary model and its main assumptions. Individuals within the households display preferences different from other members, and often the decision-making process is not dominated by a single individual but rather is the result of a process of bargaining or negotiation among members. If this is true, then an individual’s power rel-
ative to other members will play an important role in determining household decisions with respect to items such as food, clothing, education, and leisure goods. As a result, relative power may have direct and far-reaching implications for household welfare.

The main difficulty with measuring power is that it is not clear what to use as its proxy in empirical tests. A measure of bargaining power should ideally be unrelated and exogenous to any outcome that might result from an intrahousehold bargaining process, specifically one between husbands and wives. In this paper, I use as a proxy an indicator that is less likely to suffer from these endogeneity issues: the rights of an individual to retain household assets in the case of divorce. I hypothesize that during marriage a spouse derives power from the value of the assets he or she would receive were divorce to occur. Additionally, I use a more standard measure of power - human capital brought to marriage by each spouse - along with other control variables to predict household welfare indicators, namely expenditures on food, clothing, and education. To test this hypothesis I use data from the Indonesian Family Life Survey (IFLS), a national survey that included over 7,500 households. IFLS, in addition to standard household-level data, provides community-level data on adat, or the customary laws and traditions that define “how individuals relate to each other with respect to matters of marriage, divorce, inheritance, land, and property rights” (Thomas, Contreras, & Frankenberg 2002).

Results suggest that women tend to prefer nutritious foods, while men seem to have distaste for staples. Overall, divorce norms seem to do better at explaining preference differences related to food than those for clothing or education. Educational spending is more influenced by the individual characteristics of men and women than by their rights to household assets in the case of divorce.

II. Literature Review

The household has traditionally been modeled as if it were an individual whose behavior could be analyzed within the framework of classical consumer choice theory. In the unitary model, it is assumed that even if initially (at the point of household formation, for example) members of the household have different preferences, those differences disappear and a set of common preferences emerge to reflect the apparent uniformity of the household unit. Either preferences are consolidated through the formation of consensus (Samuelson, 1956) or a “dictator” who singularly controls the decision-making process rises within the household (Becker 1991). Under either of these assumptions, a representative individual can be used to study the entire household (Haddad et al., 1997; Rosenzweig and Stark 1997). This model over time has proved a useful tool for understanding the behavior of households; however, it overlooks intrahousehold interactions that may result in decisions different than those expected within the unitary framework (Rangel 2006).

In recent years new models have emerged attempting to better understand these interactions and how they might impact household welfare. One class of these models focuses on household decisions as outcomes of a bargaining process between individuals.
INTRAHOUSEHOLD BARGAINING AND DIVORCE NORMS

(Manser and Brown 1980; McElroy and Horney, 1981; Ulph, 1988; Lundberg and Pollak, 1993). Another set of models considers these outcomes the result of a series of repeated interactions that achieve efficiency in the distribution of household resources (Chiappori 1988, 1992; Browning and Chiappori 1998). Although they disagree about the mechanism by which outcomes are reached, these models share the common premise that individuals within the household have distinct utility functions and seek to influence the allocation of household resources according to their own preferences. Presumably, the more “power” an individual possesses relative to other members of the household, the more success that person will have in realizing his or her preferences and imposing them on the entire group.

McElroy (1990) suggests that indicators of welfare outside the household are the most effective proxies for power because they are presumably exogenous to any negotiation or bargaining process within the household. More specifically, indicators such as family support, marriage market conditions, or anything that reflects one’s economic prospects in the case of household dissolution, which McElroy calls “extrahousehold environmental parameters” (EEP), could prove useful measures of household power. More favorable EEP grant credibility to threats made by a spouse to leave the household (through divorce), thereby improving his or her bargaining position and, by extension, influence over household decision-making.

In this paper I focus on another EEP – the resources an individual would take away from the household in the case of divorce. I hypothesize that a more favorable distribution of assets in the event of divorce improves an individual’s bargaining power, thereby increasing his or her influence over allocation of resources while the household is intact. I use a set of indicators designed to reflect differences in preferences related to the household’s diet or nutritional status. Although overall nutrition is difficult to quantify using a single metric, some papers use measures of household calorie intake while others focus on food’s share of total household expenditures. In this paper, I use the share of a household’s food budget spent on food groups of varying nutritional value – staples (including rice, corn, cassava, potatoes, etc.), meat and dairy, and leisure items like tobacco and alcohol – to reflect differences in the preferences of husbands and wives. In Indonesia, women are often responsible for the purchase and preparation of food, and also bear most of the child-rearing responsibilities (Soemardjan 1985). Frankenburp et al. (1996) found a positive correlation between children’s height and weight – strong indicators of nutrition – and maternal education. They note that better-educated women may improve nutrition by “preparing healthier food in adequate quantities.” In Indonesia, however, such preferences likely vary across regions (given the heterogeneity of the country culturally), making generalizations based on qualitative evidence speculative at best. In the end, the question is an empirical one.

III. Methodology

This paper uses household-level data in conjunction with community-level data. This combined data set includes a sample of over 2,300 households with a married couple pres-
ent residing in 270 villages where the module containing divorce norm data was administered.

I structure my regression models in a way similar to those of Carroll (2006) and Quisumbing & Maluccio (2003). In order to test impact of relative power on the composition of a household’s food budget, I run the following model using Ordinary Least Squares (OLS):

\[
foodproportion_{ij} = \beta_0 + \beta_1 \ln hhincome_{ij} + \beta_2 \ln hhsize_{ij} + \beta_3 maleeduc_{ij} + \beta_4 maleage_{ij} + \beta_5 femeduc_{ij} + \beta_6 femage_{ij} + \beta_7 divorcerights_j + \beta_8 urban_{ij} + \beta_9 Bali_{ij} + \beta_{10} Sumatra_{ij} + \beta_{11} Java_{ij} + \varepsilon_{ij}
\]

where \( i \) indexes households; \( j \) indexes villages; \( foodproportion \) is the proportion of household food budget spent on a given food type; \( \ln hhincome \) is the natural log of the household’s combined farm and non-farm income (in Rupiahs); \( \ln hhsize \) is the natural log of the size of the household in number of members; \( maleeduc \) and \( femeduc \) indicate the level of education of each spouse in years; \( maleage \) and \( femage \) are the age of each spouse; \( divorcerights \) represents probable asset division at divorce according to local customary norms; \( urban \) is a dummy variable that takes the value of one if the household resides in an urban area and zero otherwise; \( Bali, Sumatra, \) and \( Java \) are dummies that take the value one if the household is located in that region. I control for household income in order to account for the fact that households may spend more on luxury goods (like tobacco and alcohol) once spending on necessities has reached saturation but household income continues to rise.

**IV. Empirical Results**

Results for the first model using expenditures on meat and dairy as the dependent variable are given in Table 1. These results suggest that households located in

<table>
<thead>
<tr>
<th>Proportion of Meat and Dairy</th>
<th>Coefficient</th>
<th>T-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(hhincome)</td>
<td>0.01200</td>
<td>6.85***</td>
</tr>
<tr>
<td>ln(hhsize)</td>
<td>0.00510</td>
<td>0.85</td>
</tr>
<tr>
<td>Maleeduc</td>
<td>0.00752</td>
<td>4.04***</td>
</tr>
<tr>
<td>Maleage</td>
<td>-0.00028</td>
<td>-0.71</td>
</tr>
<tr>
<td>Femeduc</td>
<td>0.00715</td>
<td>3.03***</td>
</tr>
<tr>
<td>Femage</td>
<td>0.00060</td>
<td>1.38</td>
</tr>
<tr>
<td>Urban</td>
<td>0.00722</td>
<td>1.27</td>
</tr>
<tr>
<td>Bali</td>
<td>-0.00761</td>
<td>-0.68</td>
</tr>
<tr>
<td>Java</td>
<td>-0.01276</td>
<td>-2.36**</td>
</tr>
<tr>
<td>Sumatra</td>
<td>-0.02273</td>
<td>-3.06***</td>
</tr>
</tbody>
</table>
communities where women have an absolute right to assets at divorce allocate a greater proportion of their food budget to protein-rich, nutritious foods – meat and dairy products. The response representing when women receive all household assets is positive and significant at the 10 percent level, relative to cases where individuals retain the assets they brought to marriage or were responsible for acquiring during marriage. I note that all other divorce norms have either a negative or insignificant impact on meat and dairy expenditures. Even in communities where assets are divided evenly, spending on nutritious foods is less. Why might women prefer nutritious foods more than men? Women in Indonesia are often fully employed in the farming of land inherited from their families. On the other hand, men are often expected to seek non-farm employment outside of the household in addition to helping with farming. If farming is more labor intensive, then women may require better nutrition. Also, if a woman relies on her children to help her with farming, she would have another incentive to allocate more of the household’s food budget to highly nutritious foods because the children would be more capable workers.

The results for staple foods are given below in Table 2. Given that 90 percent of households in the sample own farms and many of these produce at least some of their own food, households were asked about the value of self-produced crops that they had consumed in the prior week. These values are combined with market purchases made by the household in order to obtain a single value for expenditures on each type of food, even though these values reflect market and non-market activity.

In these results, strong rights for women at divorce fail to have a significant impact on proportion of household food budget spent on staples. However, the variable representing cases where children receive all household assets at divorce is positive and highly significant. Because women get custody of the children in nearly 70 percent of sample communities, women in most cases retain control over these assets. One might consider this an
indirect method by which wives utilize divorce as a threat point in household bargaining – the assets will remain with the children but in reality the woman gets them. However, I have less confidence in the usefulness of this variable as a proxy for relative power than for the more clear-cut cases where one spouse receives all of the assets.

Table 2: Share of Food Budget Spent on Staples

<table>
<thead>
<tr>
<th>Proportion of Staples</th>
<th>Coefficient</th>
<th>T-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(hhincome)</td>
<td>-0.02102</td>
<td>-9.08***</td>
</tr>
<tr>
<td>Ln(hhsize)</td>
<td>0.03461</td>
<td>4.39***</td>
</tr>
<tr>
<td>Maleeduc</td>
<td>-0.01402</td>
<td>-5.7***</td>
</tr>
<tr>
<td>Maleage</td>
<td>0.00086</td>
<td>1.64*</td>
</tr>
<tr>
<td>Femeduc</td>
<td>-0.00530</td>
<td>-1.7*</td>
</tr>
<tr>
<td>Femage</td>
<td>-0.00076</td>
<td>-1.31</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.03592</td>
<td>-4.78***</td>
</tr>
<tr>
<td>Bali</td>
<td>0.04363</td>
<td>2.94***</td>
</tr>
<tr>
<td>Java</td>
<td>0.01172</td>
<td>1.64*</td>
</tr>
<tr>
<td>Sumatra</td>
<td>0.05659</td>
<td>5.78***</td>
</tr>
</tbody>
</table>

Divorce (case of own-asset retention excluded)

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>T-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even division of assets</td>
<td>0.01226</td>
<td>1.37</td>
</tr>
<tr>
<td>Husband takes all</td>
<td>-0.04236</td>
<td>-2.25**</td>
</tr>
<tr>
<td>Wife takes all</td>
<td>-0.00846</td>
<td>-0.56</td>
</tr>
<tr>
<td>Children get everything</td>
<td>0.03192</td>
<td>2.84***</td>
</tr>
<tr>
<td>Other</td>
<td>0.00317</td>
<td>0.22</td>
</tr>
<tr>
<td>Religious rules</td>
<td>-0.02779</td>
<td>-1.21</td>
</tr>
<tr>
<td>Husband 2/3 wife1/3</td>
<td>0.05142</td>
<td>2.81***</td>
</tr>
<tr>
<td>Constant</td>
<td>0.53012</td>
<td>14.42***</td>
</tr>
</tbody>
</table>

Observations: 2343
R-squared = 0.1367
F(17, 2325) = 21.66

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

For example, if strong rights for children are an advantage for women and this variable is positive and significant, then why are strong rights for women not positive and significant as well? Where husbands retain all the assets, expenditures on staples are strongly and negatively affected. Staples, especially rice, are the most common type of crop cultivated by farm households in Indonesia. Men may have weaker preferences for consuming staples given their more minor role in farming relative to women. Expenditures on staples seem to be better explained by other factors. Staples, although an important source of calories, are not as rich in micronutrients as meat and dairy products. As a result, staples may not reflect a household’s nutritional preferences as well as meat and dairy. Here, staple expenditures are more strongly influenced by household income and location effects. Households in urban areas spend less on staples than rural households, perhaps
because they have access to a wider variety of goods than most rural households. Household income is strongly negative, which implies that staples are an inferior good whose budget share shrinks as household income rises. This is consistent with the predictions of Engel’s law that spending on necessities shrinks with increasing income.

V. Conclusion

Overall, the impact of divorce norms on food expenditure shares is more significant than on levels of expenditure on clothing and education (results not shown). In communities where divorce norms strongly favor women, households tend to spend more on very nutritious foods – meat and dairy products. Women may choose to spend more on these items than their husbands for a variety of reasons. Since women retain custody of their children in over two-thirds of communities, they may view themselves as having a greater stake in the future health of their children. In regards to staple foods, households spend relatively more when children are likely to receive all of the household’s assets at divorce. While we can speculate that such a division of assets favors women, extending this result to the household bargaining framework and drawing conclusions about the preferences of women is risky. Wealthier families tend to spend less on staples, suggesting that staples are treated as inferior goods. Urban families also spend less on staples, likely because they produce less of their own food and have a greater variety of foods available for consumption.

Based on these results, however, I refrain from making firm policy recommendations. The results for food suggest that more resources in the hands of women could improve household nutrition. However, implementing policies that intervene directly and seek to shift relative power within the household demand careful consideration and require supporting evidence beyond what is offered here.

References


THE EFFECTS OF THE SARBANES-OXLEY ACT ON U.S. MERGER ACTIVITY

Freddy Tsai
University of British Columbia

I. Introduction

The “Public Company Accounting Reform and Investor Protection Act of 2002” (more commonly referred to as Sarbanes-Oxley, or SOX) was signed into U.S. federal law on July 30 of 2002 in reaction to major corporate governance and accounting scandals (most notably WorldCom, Enron, and Tyco International). Its far-reaching effects prompted George W. Bush to hail the act as “the most far-reaching reforms of American business practices since the time of Franklin D. Roosevelt” (Bumiller, 2002, New York Times).

Despite the proposed numerous benefits of the act, there has been widespread dissent in the business industry against the substantial costs of complying with and implementing SOX. Critics of the act have ranged from academics and CEO’s to U.S. Treasury Secretary Henry Paulson. (Wright, 2006). The costs of SOX are very real – one often cited statistic is the large drop in U.S. IPO’s after the passage of Sarbanes-Oxley. For instance, only 5% of the value of global initial public offerings was raised in the U.S. in 2005 compared with 50% in 2000 (Bawden, 2006). In response, the securities and accounting regulators have recently conceded that interpretations of SOX may be too stringent and have stated that it will issue guidance on interpreting SOX in a way to save firms time and money (Scanell & Solomon, 2006).

In this paper I analyze whether the passage of the Sarbanes-Oxley Act resulted in a decline in U.S. M&A activity through multivariate analysis. There is evidence to suggest that the substantial costs of compliance and implementing SOX are a possible impediment to M&A deals. In a survey by RHR International (2004), corporations reported an average increase of 77 percent ($16 million) in compliance costs due to Sarbanes-Oxley. However, the costs of SOX go beyond the direct costs of compliance and include many indirect opportunity costs such as diversion of attention from conducting business and discouraging value-increasing risky investment (Zhang, 2005). In fact, in what has been called “the most extensive and persuasive study of SOX’s costs” (Ribstein, 2005, p. 15),
Zhang (2005) estimated the total (abnormal) loss in total market value from SOX to be $1.4 trillion. Empirically, the hypothesis that SOX is an impediment to M&A activity is corroborated by the fact that the total number of U.S. mergers/acquisitions declined 11.8% from 10294 deals in 2001 to 9077 deals in 2002 (Thompson SDC Platinum Special Mergers Sectors Database).

Despite these substantial costs, it is still unclear as to whether SOX has a negative effect on M&A activity. As Greenberg & Walton note (2003, p. 13), “if [Sarbanes-Oxley] succeeds in improving the quality of public disclosure, it will facilitate M&A transactions due to greater transparency and confidence in public filings.” Moreover, a cross-country analysis of mergers and acquisitions by Rossi and Volpin (2003, p. 1) found that “the volume of M&A activity is significantly larger in countries with better accounting standards and stronger shareholder protection.” However, it could possibly be that once a threshold level of accounting standards is met, additional regulations are, in fact, a hindrance to M&A activity. Thus, a side implication of this study is to examine via a case study, whether this is possibly true in the country with arguably, the most stringent accounting regulations in the world, the United States.

There has been relatively little work investigating the effects of macroeconomic determinants on aggregate merger and acquisitions activity. In addition, to my knowledge, there has been only one other event study on M&A activity (See Benzing 1991). Lastly, there are serious shortcomings with previous research in terms of both methodology and data. This is a result of the fact that until the launch of Thomson Financial’s SDC Platinum database, mergers and acquisitions data was highly limited and resulted in researchers continually utilizing the same data over and over again. Almost all prior research on aggregate M&A activity and macroeconomic determinants has been confined to data on the number of annual completed mergers in the U.S. mining and manufacturing sector from 1919 to 1979 (e.g. Weston, 1961; Beckenstein, 1979; Polonchek & Shushka, 1987; Benzing 1991; Crook, 1996; Clark, Chakrabati and Chiang 1988). The results of this fact are that: 1) samples sizes are relatively small which may lead to poor and inaccurate estimation, 2) time-period bias. The consequences of time-period bias are significant as we expect significant structural changes to the U.S. economy since then, (most notably the rise of the technology and service sectors and the decline of mining and manufacturing) which prohibits generalization of previous results to the current economy. This paper benefits by utilizing the Special Mergers Sectors Database provided by Thomson SDC Platinum which allows for generalization of results to the entire economy.

Aside from improving on previous research methodology and data, this paper’s significance comes from examining whether heightened accounting regulations or accounting standards are conducive towards mergers and acquisitions (and hence investment) or, in fact, inhibitive. In addition, it adds to the ongoing policy debate of whether the benefits of SOX outweigh its costs (see Zhang, 2005; The Economist, 2005). Specifically it investigates the effects of SOX on mergers and acquisitions, which have yet to be quantified. All this comes at a time when concern for corporate governance is at an all-time high, and with countries considering increasing their accounting standards.
In this study, I find that the Sarbanes-Oxley Act is negatively related to U.S. M&A activity at the 10% significance level and has decreased the number of mergers per month by roughly 8.6%. One surprising result is that contrary to previous research, stock prices are found to be negatively related to merger activity. This suggests that further research needs to be conducted to determine the cause of this, possibly through seasonality-adjustment, and accounting for the number of firms and Tobin’s Q ratio.

II. Effects of Sarbanes-Oxley in Mergers & Acquisitions

The Sarbanes-Oxley Act consists of various heightened corporate governance and accounting regulations designed to improve financial reporting and restore investor confidence. SOX consists of 60 sections and the most heavily criticized and difficult to implement sections are Sections 302 (“Corporate Responsibility for Financial Reports”) and Section 404 (“Management Assessment of Internal Controls.”) (The Economist, 2005). The requirements of Sections 302 and Section 404 are that: 1) management must establish and maintain a system of adequate internal controls, 2) management must report annually on “the effectiveness of the internal control structure and procedures of the issuer for financial reporting” and 3) The external auditor of the company must testify as to management’s assessment of the company’s internal controls.

The problem with Sections 302 and Section 404 lies in the fact that according to John Thain, Chief Executive of the New York Stock Exchange, is that they lack a “materiality standard” (Scannell & Solomon, WSJ). The same Wall Street Journal article reports that interpretation of Sections 404 (and to a certain extent 302) is overly broad and requires companies to spend countless hours and millions of dollars “documenting things that have nothing to do with the integrity of their financial statements.” In fact, the article also reports that auditors interpret the rules to such a degree that management must document and account for things such as who has access to an office key.

The burden of SOX is most evident for small-cap companies, especially considering the fact that SOX compliance is largely a non-value adding activity (i.e. it does not promote sales). A report by Mark Cecil (2003) notes that small firms face substantial compliance, insurance, regulatory, and other costs because of SOX. The same report observes that independent audit costs may have increased 15% to 100% due to SOX and that the cost of director and officer insurance has potentially increased by 400% due to SOX. To illustrate this fact, the report highlights a company with a market-capitalization of $50-million estimated the expense of being public and subject to SOX was $1 million per year - $1 million that could otherwise be invested in the company.

However, it is important to note that while SOX materially affects small companies more, their costs certainly extend to companies and organizations of all sizes. Companies are now hiring officers to deal with compliance (Accounting Web.com, 2004). The Sarbanes-Oxley act has also driven up the costs of director fees as it increases director qualifications (thus limiting the pool of qualified directors) and increases the work of directors. Linck et al. (2006) found that the workloads and risks of directors increased after
SOX – for example, audit committees meet more than twice as often after SOX than before SOX. Also, they found that Director and Officer insurance premiums have more than doubled. Link et al. (2006) found that the director fees paid by small firms increased from $2.34 per $1,000 in net sales in 2001, to $3.19 per $1,000 in net sales in 2004. Similarly, the costs for large firms increased from $0.25 per $1,000 in net sales in 2001, to $0.32 per $1,000 in net sales in 2004.

By now it should be evident that SOX involves substantial costs to companies. But, the question that remains to be asked is ‘how exactly do these costs work themselves into M&A transactions and the deal making process?’ There are two roots to the costs of SOX: 1) Direct costs of compliance, and 2) Indirect Costs.

First off, Sarbanes could stand to inhibit M&A Activity in that it pertains only to SEC registrants. As Karan & Sharifi (2006) note, in terms of SOX, there are four possible types of deals depending on the target and acquirer’s SEC registration status: 1) Both target and acquirer are SEC Registrants, 2) Acquirer is an SEC Registrant, Target is a non-SEC Registrant, 3) Target is an SEC-Registrant, Acquirer is a non-SEC Registrant, 4) Target and Acquirer are both non-SEC Registrants.

In addition to the direct costs of compliance, there are indirect and opportunity costs to be considered with SOX. First, it diverts management and employee attention away from the company’s operations to documentation and compliance with SOX. As stated by the chief accounting officer of General Motors Corp, “The real cost isn’t the incremental dollars, it is having people that should be focused on the business focused instead on complying with the details of the rules.” (Solomon, 2004, WSJ). Secondly, SOX increases the risks and penalties of litigation for company executives. As a result, CEO’s are likely to take less risky but profitable actions which could stifle growth and potentially reduce the value of their firms. (Wallison, September 2003, WSJ). The opportunity costs of SOX are not to be underestimated, in what has been called “the most extensive and persuasive study of SOX’s costs” (Ribstein, 2005, p. 15), Zhang (2005) estimated the total (abnormal) loss in total market value from SOX to be $1.4 trillion.

Despite these substantial costs, it is still unclear as to whether SOX has a negative effect on M&A activity. As Greenberg & Walton note (2003), “if [Sarbanes-Oxley] succeeds in improving the quality of public disclosure, it will facilitate M&A transactions due to greater transparency and confidence in public filings.” Moreover, a cross-country analysis of mergers and acquisitions by Rossi and Volpin (2003) found that “the volume of M&A activity is significantly larger in countries with better accounting standards and stronger shareholder protection.”

III. Literature Review

As with many other studies, this paper draws mainly on papers from two separate fields in order to produce an original work of research. In this case, the study draws from literature on: 1) the costs and effects of Sarbanes-Oxley, and 2) the determinants of M&A activity.
EFFECTS OF THE SARBANES-OXLEY ACT

The vast majority of quantitative research conducted on the effects of Sarbanes-Oxley has been on its effect on stock prices and the stock market (e.g. Zhang, 2005; Li, et al. 2003; Rezaee & Jain 2004). In fact, there has been no quantitative research done on the effects of SOX on mergers and acquisitions. While there are two works exploring the effects of SOX specifically on M&A, the treatment provided by Greenberg & Walton (2003) and Karan & Sharifi (2006) is limited to a qualitative discussion of the topic. Greenberg & Walton note that SOX has had a substantial impact on the due diligence process for M&A deals as well as the negotiation and documentation of the transactions. Karan & Sharifi details the process by which SOX affects M&A deals, most notably through the costs of bringing non-compliant companies into compliance, the costs of privatizing companies, and possible costs of future compliance.

There is a considerable volume of literature on the determinants of mergers and acquisitions. The relative few empirical studies that examine the macroeconomic determinants of M&A activity on a national level (and thus most similar to this study) have been restricted to data before 1980 and is restricted to data on the U.S. mining and manufacturing sector rather than the aggregate economy (Crook 1996). Together, Benzing (1991) and Crook (1996) provide an excellent overview of various studies examining the macroeconomic determinants of M&A activity. They summarize that all previous studies have found stock prices to be significantly positively related to merger activity (e.g. Nelson, 1959; Weston, 1961; Beckenstein, 1979; Melicher et al., 1983; Polonchek & Shushka 1987). In addition, Benzing (1991) and Polonchek & Shushka (1987) found merger activity to be negatively related to merger activity.

However, for all other variables, the evidence is mixed as to whether they are significant. For interest rates/corporate bond yields, Melicher et al., Golbe & White, and Polonchek and Sushka found bond yields to be negatively related to merger activity. On the other hand, Beckenstein found a positive relationship while Benzing found that bond yields and merger activity were positively related before 1950, but negatively related after 1950, possibly due to the Celler-Kefauver Act. With GNP and production levels, Nelson and Golbe & White found that they had a significant effect on merger activity while Beckenstein, and Clark et al. did not. Lastly, with lagged dependent variables of merger activity, Clark et al. (1988) found that merger activity is autoregressive to the second order through ADF tests and ARIMA modeling. Contrarily, Benzing did not find the lagged dependent variables to be significant.

IV. Data

This paper utilizes a monthly time-series dataset for the sample period January 1st, 1996 – February 28, 2007 and is comprised of variables for: monthly number of mergers/acquisitions, corporate AAA bond yields, monthly unemployment rate, and the level of the S&P 500 Index. The dependent variable of the number of mergers per month was obtained from Thomson Financial SDC’s Special Mergers Sectors Database. Its extensive coverage includes “All corporate transactions involving at least 5% of the ownership
of a company where the transaction was valued at $1 million or more (after 1992, deals of any value are covered) or where the value of the transaction was undisclosed. Public and private transactions are covered.” Deals included were those in which either the target or acquirer was a U.S. company. We exclude deals in which neither target or acquirer is an SEC Registrant, as we expect the effects of SOX on the deals to be negligible. In addition, it is important to note that while technically foreign companies can be SEC Registrants, SOX is generally not applicable to foreign companies as most foreign issuers in the U.S. file under rule 144A which allows less strict reporting requirements for foreign issuers (The Economist, 2006).

Corporate AAA bond yields are expressed in percents and provided by Moody’s Investors Services. It is derived by averaging daily data. The monthly unemployment rate is also expressed in percents and is provided by the U.S. Bureau of Labor Statistics. In addition, the data is seasonally adjusted. Lastly, as a method of proxying stock prices, we use the level on the S&P 500 index provided by the University of Chicago’s Center for Research in Security Prices. The original data indicates the level of the Standard & Poor’s 500 Composite Index at the end of last trading day for the month, but I have lagged the data such that it becomes the level of the S&P 500 Index for the first trading day of the month. The data is available only until December 2006, so it is supplanted by data from Economagic for the months of January 2007 and February 2007.

**Regression Model**

Our objective is to examine whether the Sarbanes-Oxley act has had either a positive or negative increase in the number of mergers and acquisitions in the United States as theoretically, the overall effect is ambiguous. Thus in working with our time-series data, the general model for our regression equation is:

$$ M_t = \beta_0 + \beta_1 M_{t-1} + \beta_2 M_{t-2} + \beta_3 SP500_t + \beta_4 Yield_t + \beta_5 Unemp_{t-1} + \delta_{1} Sarbanes_t + u_t $$

Note that the regression equation is identical to the one in Benzing (1991), aside from the lagged unemployment rate variable and the Sarbanes-Oxley dummy variable. The dependent variable $M_t$ is the number of announced M&A transactions per month where either the target or acquirer is a U.S. company and covers the period from January 1st, 1996 until February 28, 2007.

The variables $M_{t-1}$ and $M_{t-2}$ are lagged dependent variables and are included as Clark (1988) found merger activity to be second-order autoregressive. Note that the lagged dependent variables did not induce multicollinearity for Benzing. The variable ‘$SP500_t$’ represents the opening market value for the month for the S&P 500 Index and controls for the effect that M&A activity increases during periods of higher stock prices as it is easier to acquire companies through stock-for-stock deals when the acquiring company’s stock is highly valued. In addition, higher stock prices indicate an optimistic economic outlook - suggesting higher rates of return on investment, and thus encouraging merger activity (Benzing 1991). The variable ‘$Yield_t$’ represents the monthly average yield on corporate Aaa bonds. As acquisitions are often financed with debt, the variable controls for the fact
that M&A deals are easier to conduct during periods of cheap financing.

The variable ‘Unemp_{t-1}’ represents the previous month’s unemployment rate. The variable is lagged by one month in recognition of the fact that the Bureau of Labour Statistics does not release unemployment rate data until a month later (Bureau of Labour Statistics, 2006). Like Benzing (1991), I use it to proxy the business cycle and control for the fact that mergers activity is likely to increase during economic growth periods. Lastly, the dummy variable ‘Sarbanes_{t}’ takes the value ‘1’ if the period in consideration is after July 2002, and ‘0’ otherwise and is meant to capture the effect of Sarbanes-Oxley on M&A activity. Also, to control for possible seasonality, we include monthly dummies. Intuitively, one might expect merger activity to decline at the end of the year as companies slow down activities for the holidays and avoid complex integration and accounting issues right before the fiscal year end. Similarly, we might expect that this might create a backlog of M&A deals, leading to increased M&A activity at the start of the year.

In summary, the hypothesized relationship between the level of merger activity and the macroeconomic determinants is as follows:

Mergers = f(mergers_{t-1}, mergers_{t-2}, cost of funds, unemployment rate, stock prices, SOX)

As a cautionary note, I would like to note that similar to all other event studies, this event study suffers from the drawback that the event dummy can also capture effects from other news and events. While one cannot ignore this fact, it should also be realized that other events are likely to be controlled for through their effects on the unemployment rate, corporate bond yields, and especially the S&P 500 Index. In addition, as part of her study Zhang (2005) ran a search of events likely to affect abnormal market returns from SOX such thing as legislative activities, economic statistics, and accounting scandals. Her conclusion was that, while there were legislative activities, accounting scandals, and economic news during the period, they were unlikely to be the key determinants of the abnormal stock returns for the period” (Zhang, 2005). Thus it also seems reasonable to assume that there were unlikely to be other key drivers of M&A activity during the period (since M&A activity and stock returns have similar determinants) not already accounted for through our other variables.

**Regression Analysis**

Our dependent variable is the number of merger announcements in a month. In employing OLS we follow the work of Weston (1961), Beekenstein (1979), Polonchek & Shushka (1987), Benzing (1991) and Crook (1996). While the simple linear model has may have drawbacks for limited dependent variables, we recognize that “if a strictly positive variable takes on many different values, a special econometric model is rarely necessary” (Wooldridge, 2006, p. 582). Note also that the variable is not dominated by zeroes or even ones. In fact, from Table 1.1 we see that the minimum value is 536. In referring to Figure 1.1, we see that the dependent variable takes on a wide range of values (and is also somewhat normally distributed). However, as there is no reason to suspect that homoskedasticity is present, we report heteroskedasticity-robust standard errors.
V. Regression Results

The regression results are reported in Tables 2.1 and 2.2. First we note that including the lagged dependent variables induces multicollinearity in the model, causing all other explanatory variables to be statistically insignificant. As the basis of this paper, we find that in OLS(2) the Sarbanes coefficient has a value of -73.34, but is statistically insignificant with a two-sided p-value of 0.146. However, after controlling for seasonality through monthly dummies in OLS(3), we find that the Sarbanes coefficient becomes significant at a 10% level with a coefficient of -88.687. Thus, we expect that after controlling for other factors, the Sarbanes-Oxley Act reduces the number of monthly mergers by roughly 89 deals on average (or 86/1002 where 1002 is the mean monthly mergers, represents an 8.6% decrease). Given the 10% significance level and the large coefficient on the Sarbanes variable, from a policy standpoint it may be worthwhile to conclude that the Sarbanes-Oxley Act has a negative effect on M&A activity. Note that while we should theoretically use a two-sided test, given the persuasive argument of the costs of Sarbanes-Oxley, it may be more practical to utilize a one-sided test and conclude that the Sarbanes coefficient is statistically significant at a 5% level.

A Wald test on the monthly dummies shows an F-statistic of 2.05 and a p-value of 0.0296, suggesting that there is some seasonality present in our data. This suggests that the study would benefit from seasonally adjusting the data through ARIMA modeling as in the case of Melicher et al. (1985) and Clark et al. (1999).

A surprising result that conflicts with all prior research is that we find that stock prices are significantly negatively related to merger activity under all model specifications (excluding those with lagged dependent variables). However, as Weston (1961), Crook (1996), and Beckenstein (1979) did not include a proxy for the business cycle, it is possible that the coefficients for their stock prices were inflated. Theoretically however, the result may not be as surprising as higher stock prices also mean that the cost of acquiring firms is higher as target companies have higher stock prices. In addition, from the work of Tremblay & Tremblay (1988), it is possible that when the stock market is rising few firms see the need to merge or be acquired as a means of avoiding bankruptcy. In addition, it is possible that in periods where the market is doing well, firms have high enough internal growth and less justification to expand through mergers and acquisitions.

The rest of the regression results are as expected and are consistent with previous findings. We find that like Clark (1988), merger activity is second-order autoregressive. However, it also induces multicollinearity. Lastly, we find that under all specifications (aside from those with lagged dependent variables) unemployment and corporate bond yields are negatively related to M&A activity, and statistically significant.

VI. Conclusion and Future Research

This paper investigates the effects of the Sarbanes-Oxley Act on aggregate U.S. mergers and acquisition activity. Theoretically, the effect is ambiguous as increased accounting
and financial reporting standards should increase confidence in deals and encourage M&A deals. On the other hand, the Sarbanes-Oxley Act imposes substantial compliance costs and indirect/opportunity costs on firms. This study finds that there is slight statistical evidence to suggest that the Act has reduced M&A activity in the U.S. The finding that Sarbanes-Oxley reduces M&A activity offers evidence against Rossi and Volpin’s (2005) finding that heightened accounting standards encourage mergers activity. It is possible that once accounting standards reach a given threshold, higher accounting standards serve as an obstacle to merger activity.

The initial findings corroborate with previous research that unemployment and the cost of financing are negatively related to merger activity. A surprising result however, is that this study finds that stock prices are negatively related to merger activity. In terms of future research, much of it can be guided by the surprising result that stock prices are negatively related to merger activity. It is important to determine whether this result is due to model mis-specification or is indeed the correct result. Theoretically, the effect is ambiguous.

For future research I propose that ARIMA modeling be undertaken as there are initial indications that there is some seasonality at play. In addition, the number of firms should be worked into the model (though preliminary searching indicates that such data is difficult to find, and there is also the question of what firms should be included). If data on the number of firms is available, it should be possible to create a dependent variable of (monthly mergers/number of firms) and employ a binary response model such as the Tobit or logit model. Both models are possibly more theoretically valid than OLS. Lastly, rather than (or in addition) to stock prices, it may be useful to include Tobin’s Q ratio as an explanatory variable.

Inclusion would effectively control for growth prospects of firms as a higher Q ratio is an indicator of an optimistic outlook. However, one would expect substantial correlation between Tobin’s Q and stock prices which could induce multicollinearity.

References


Linck, J., Netter, Jeffry., & Yang, Tina. “The Effects and Unintended Consequences of the Sarbanes-
EFFECTS OF THE SARBANES-OXLEY ACT


**TABLE 1: DESCRIPTIVE STATISTICS OF MONTHLY MERGERS**

<table>
<thead>
<tr>
<th>Summary Statistics - Monthly Mergers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Mode</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Sample Variance</td>
</tr>
<tr>
<td>Kurtosis</td>
</tr>
<tr>
<td>Skewness</td>
</tr>
<tr>
<td>Range</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Sum</td>
</tr>
<tr>
<td>Count</td>
</tr>
<tr>
<td>Largest(1)</td>
</tr>
<tr>
<td>Smallest(1)</td>
</tr>
<tr>
<td>Confidence Level(95.0%)</td>
</tr>
</tbody>
</table>

**Figure 1.1**

[Histogram of Monthly Mergers]
### Table 2.1: OLS and Poisson Regression Results

<table>
<thead>
<tr>
<th></th>
<th>OLS (1)</th>
<th>OLS (2)</th>
<th>OLS (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co</td>
<td>0.541</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.44</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Mergers$_{t-2}$</td>
<td>0.339</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.64</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Unemp$_{t-1}$</td>
<td>-14.784</td>
<td>-313.74</td>
<td>-326.666</td>
</tr>
<tr>
<td></td>
<td>-0.39</td>
<td>9.94</td>
<td>-10.25</td>
</tr>
<tr>
<td></td>
<td>(0.698)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>ln of S&amp;P 500</td>
<td>-16.287</td>
<td>-477.52</td>
<td>-508.014</td>
</tr>
<tr>
<td></td>
<td>-0.18</td>
<td>-5.48</td>
<td>-6.07</td>
</tr>
<tr>
<td></td>
<td>(0.854)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>bondyield</td>
<td>3.333</td>
<td>-74.55</td>
<td>-88.687</td>
</tr>
<tr>
<td></td>
<td>0.13</td>
<td>-2.58</td>
<td>-2.94</td>
</tr>
<tr>
<td></td>
<td>(0.898)</td>
<td>(0.011)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Sarbanes</td>
<td>3.594</td>
<td>-73.34</td>
<td>-82.0515</td>
</tr>
<tr>
<td></td>
<td>0.12</td>
<td>-1.46</td>
<td>-1.66</td>
</tr>
<tr>
<td></td>
<td>(0.903)</td>
<td>(0.146)</td>
<td>(0.099)</td>
</tr>
<tr>
<td>Constant</td>
<td>284.72</td>
<td>6419.783</td>
<td>6811.588</td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td>7.61</td>
<td>8.10</td>
</tr>
<tr>
<td></td>
<td>(0.768)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Monthly Dummies</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Robust Option</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.78</td>
<td>0.53</td>
<td>0.59</td>
</tr>
<tr>
<td>Observations</td>
<td>132</td>
<td>133</td>
<td>133</td>
</tr>
</tbody>
</table>

Where the top number is the coefficient, the second number is the t-statistic, and the number in parentheses is the p-value.
### Table 2.2: Results for OLS (3)

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>bondyield</td>
<td>-88.68705</td>
<td>(0.004)</td>
</tr>
<tr>
<td>unemp1</td>
<td>-326.6557</td>
<td>(0.000)</td>
</tr>
<tr>
<td>lsp500</td>
<td>-508.0138</td>
<td>(0.000)</td>
</tr>
<tr>
<td>sarbanes</td>
<td>-82.0515</td>
<td>(0.099)</td>
</tr>
<tr>
<td>february</td>
<td>-107.2271</td>
<td>(0.021)</td>
</tr>
<tr>
<td>march</td>
<td>33.62784</td>
<td>(0.610)</td>
</tr>
<tr>
<td>april</td>
<td>-11.0081</td>
<td>(0.828)</td>
</tr>
<tr>
<td>may</td>
<td>7.745926</td>
<td>(0.860)</td>
</tr>
<tr>
<td>june</td>
<td>22.27791</td>
<td>(0.680)</td>
</tr>
<tr>
<td>july</td>
<td>53.51091</td>
<td>(0.377)</td>
</tr>
<tr>
<td>august</td>
<td>-11.75161</td>
<td>(0.825)</td>
</tr>
<tr>
<td>september</td>
<td>-57.30114</td>
<td>(0.387)</td>
</tr>
<tr>
<td>october</td>
<td>-16.54553</td>
<td>(0.756)</td>
</tr>
<tr>
<td>november</td>
<td>-115.8308</td>
<td>(0.007)</td>
</tr>
<tr>
<td>december</td>
<td>-15.74069</td>
<td>(0.757)</td>
</tr>
</tbody>
</table>

P-values are in parentheses

### Figure 1.2: Monthly Merger Trend

![Monthly Merger Trend Graph](chart.png)

- **Monthly Mergers**
- **4 per. Mov. Avg. (Monthly Mergers)**
OIL PRICE SHOCKS AND MONETARY POLICY REVISITED:
AN ENERGY-AUGMENTED NEW KEYNESIAN DSGE MODEL FOR THE
ANALYSIS OF THE ROLE OF SYSTEMATIC MONETARY POLICY IN THE
OIL PRICE-MACROECONOMY RELATIONSHIP

Sören Radde
University of Bayreuth

ABSTRACT

It is a widely held belief that the 1970s’ oil price crises were responsible for the severe stagflations witnessed during this period in many industrialized economies. A number of authors have provided evidence in favour of an alternative interpretation of the 1970s’ stagflation and the subsequent economic stability attributing the majority of the adverse effects of oil price surges to the weak response of systematic monetary policy. This perception has been confirmed by simulations of empirically estimated monetary policy rules in an ad hoc New Keynesian Dynamic Stochastic General Equilibrium model by Clarida, Gali and Gertler (2000). These authors remain silent, however, on the precise transmission channel of oil price shocks as their model lacks an appropriate micro-foundation. Therefore, I develop a model which explicitly accounts for the role of energy and find that the insights of the monetary explanation of the 1970s’ stagflation can be reproduced qualitatively, though not quantitatively. This result suggests that oil price shocks are either over-estimated as a direct and indirect – source of business cycle fluctuations or that the economic environment proposed by Clarida, Gali and Gertler and replicated by my model is not rich enough to account for the interplay of monetary policy and oil shocks.

I. Introduction

The impact of oil price shocks on the industrialized economies seems to have changed radically over the past 35 years. Whereas the oil crises of the 1970’s were associated with highly persistent inflation and significant contractions in output the oil price movements
we witness since the mid-1990s and especially since 2004 do not seem to have such disturbing macro-economic impacts. This fact has cast doubt on the purely oil price-focussed explanation of U.S. business cycle dynamics and motivated the search for alternative answers. Thus, a number of authors have argued that the adverse effects of the oil price surges in the 1970’s are largely attributable to the inadequate response of monetary policy which failed to combat inflation by raising the real interest rate. According to this interpretation a switch to a more aggressive monetary regime which occurred sometime at the end of the 1970s or at the beginning of the 1980s is at the source of the seemingly changed oil price-macroeconomy relationship.

This observation has led Clarida, Galì and Gertler (2000; CGG henceforth) to test the performance of monetary policy rules à la Taylor estimated for the period prior and posterior to 1979 within an ad hoc New Keynesian Dynamic Stochastic General Equilibrium (DGSE) model. Their major insight is that aggressive monetary policy is very successful in stabilizing the economy while only a weak monetary performance allows for the empirically observed volatility in output as well as high and persistent inflation in response to exogenous fundamental shocks. Taking the 1970s’ oil price shocks as an example of such fundamental disturbances, CGG’s result seems to support the notion that economic instability during this period is largely attributable to monetary policy. However, CGG do not take a stance on how exactly such shocks work their way through the economy as their model lacks an appropriate micro-foundation. Therefore, this approach might be inappropriate to capture the oil price-macroeconomy relationship.

In order to investigate whether CGG’s conclusions can nonetheless be reinforced with respect to oil price shocks I develop a micro-founded DSGE model sharing most features with the model used by CGG, but explicitly accounting for the role of energy in the economy. Apart from the baseline specification of my model exhibiting only nominal rigidity, I also introduce a modified version with both nominal and real rigidity which allows me to flesh out the most severe consequences an oil price shock can have in this framework. As a second step, I evaluate the performance of interest rate rules estimated for the pre- and post-1979 period in response to an oil price shock against both environments. I find that my model is able to replicate qualitatively CGG’s conclusions with respect to the interplay of oil price shocks and monetary policy– as predicted, in both versions of the model output and inflation react much stronger in the monetary regime which accommodates inflationary pressure rather than taking an aggressive stance on it. Still, even in the presence of a substantial degree of real rigidity the fluctuations in economic aggregates are so small that they cannot even remotely approximate the corresponding volatility observed in real world data. Therefore, I conclude that either oil price shocks are far more innocuous than commonly believed – at least in economies which are reasonably well described by the model structure - or the energy-augmented model is not rich enough to capture the oil price-macroeconomy relationship due to a number of reasons discussed in the concluding section.

The rest of the paper is organized as follows. Section 2 offers a brief overview of the empirical literature on the standard interpretation of the link between oil and the economy.
as well as challenges to this purely oil price-focussed view putting forward the significance of systematic monetary policy. In section 3 the baseline and the modified version of the energy augmented-model are developed including a qualitative discussion of the impact of some parameters on equilibrium dynamics. The model is calibrated in section 4 and the simulation results are presented in section 5, where both, accommodating and aggressive monetary policy is evaluated against the economic environments specified by the baseline and the modified model. This section stresses the qualitative discussion of the dynamics at hand. Section 6 finally highlights the central insights of the paper and offers concluding remarks on some aspects left to future research.

II. Oil, Monetary Policy and the Macroeconomy

Stylized Facts on Oil Price Shocks and Business Cycle Fluctuations

During the 1970s most industrialized economies faced a phenomenon often referred to as the ‘Great Stagflation’, describing the coincidence of severe recessions and soaring inflation. The fact that two previously unparalleled oil price shocks preceded these economic downturns established the widely held belief that oil price movements were the major cause of economic fluctuations during this period. This conventional wisdom has been lent scientific support by a whole strand of literature focussing on the oil price–macroeconomy relationship.

In his much noticed 1983 article Hamilton demonstrates, for instance, that a structural relationship between oil prices and output contractions indeed existed ever since World War II in the U.S.. Following common narrative evidence about the determinants of oil price shocks he first makes a case for the exogeneity of these disturbances. In this vein, oil price surges since 1945 are generally attributed to political factors. In particular, the 1973 OAPEC embargo on oil exports in the course of the Yom Kippur War and the resulting quadruplication of crude prices are hold responsible for the 1973-75 recession. Supply disruptions in the wake of the Iranian revolution in 1979 and the following Iran-Iraq war lead to the post-war all-time peak in real crude oil prices and are believed to have caused the 1980 and 1981–82 depressions (cf. Figure 1). Based on these arguments in favour of the exogenous nature of oil price shocks with respect to the U.S. economy and their apparent adverse effects, Hamilton presents forceful econometric evidence for the correlation between petroleum prices and output. Thus, he rejects the hypothesis of an arbitrary historical coincidence of oil price surges and economic downturns and refutes the proposition that some third endogenous factor was responsible for volatility in oil prices and output at the same time. Moreover, using a modified indicator of oil price increases Hamilton reinforces the evidence of a systematic relation between oil prices and economic contractions in a later article (cf. Hamilton 1996) by extending it to the post 1986 period: Until the mid-1990s the only significant disruption in oil prices, which occurred subsequent to the Iraqi invasion of Kuwait in 1990, was accompanied by a recession in the US. In light of these findings Hamilton concludes with a prophecy.

“To summarize, the evidence since 1983 has strengthened, not weakened, my earlier convictions. My 1985 article concluded with the statement: ‘The political his
tory of the Middle East makes it almost inevitable that sometime within the next decade economists will be granted some more data with which to assess the economic effects of oil supply disruptions. ‘This is exactly what happened in 1990 when Iraq invaded Kuwait, and surely this oil shock was a key factor in the recession that followed. But for those who have yet to be convinced, I hereby renew the forecast – sometime again within the next ten years, turmoil in the Middle East will produce another major disruption to world petroleum supplies. The crisis will produce a recession in the United States.’ (Hamilton 1996, p.220)

And indeed, about five years later a new recession occurred – although with some delay – after a substantial increase in oil prices due to OPEC production cutbacks. The historical evolution of the real world oil price together with some decisive political events, which seem to have driven its dynamics, is shown below.

**Figure 1: Major Political Events and Real World Oil Price, 1970-2006**

(Prices adjusted by quarterly GDP deflator, $2005)

Author’s calculations. Source: EIA

In early 2004 oil prices began to soar again until they peaked in August 2006 at $65.83 per barrel, exceeding by far the price level of the 1973-75 energy crisis. This development was initially fuelled by the countermovement of demand and supply due to increasing oil thirst in the U.S. and Asian economies and fears of disruptions in oil production in the wake of the second Gulf War. In the face of hurricanes hitting the Gulf of Mexico, disturbances in Nigeria, escalating tensions between Israel and Lebanon, additional struggle in
the Middle East due to the Iranian nuclear program as well as ongoing unrest in Iraq. Oil supply chains were further deteriorated in 2005 and 2006 so that even substantial OPEC and non-OPEC production increases could not halt the boost in petroleum prices. Moreover, the loss of production capacity in several OPEC countries combined with increased production to meet growing international demand and to counteract the extreme price surges led to the erosion of OPEC excess oil production capacity, adding a significant risk premium to crude prices.

As these price movements persisted over an unusually long period of time and took threatening dimensions even by historical standards, they renewed fears of a severe recession hitting the world economy. However, U.S. GDP seemed to respond far less vigorously than previously experienced. There are a number of reasons why the developed economies might have responded differently to the most recent oil price shock. First of all, energy intensity - as measured by the “energy consumption or use to GDP-ratio” - and therewith oil-dependency has declined constantly over the last decades in the industrialized countries. Second, the relative importance of oil has decreased due to substitution processes set off by the 1970s oil price crises. Finally, the very nature of the oil price shocks was different, as contrary to the shocks of the 1970s demand pressure played a major role. Hence, one could assume that the direct effects of oil shocks on economic activity have changed to explain their reduced impact.

Paradoxically, however, the share of oil in production and consumption has never been very high, even throughout the 1970s. Therefore, one could come to doubt more fundamentally, that oil shocks alone could ever have accounted for the significant volatility in macroeconomic aggregates. Alternatively, a number of economists emphasize the impact of systematic monetary policy on U.S. business cycle dynamics – a factor which Hamilton himself alludes to in concluding that the correlation between oil prices and output should not “[…] be viewed as an immutable structural relation. Changes in expected inflation, the response of monetary policy to oil shocks, or the regime in which oil prices are determined could be expected to give rise to a different dynamic pattern.” (Hamilton 1983, p.247)

**New evidence and current debate**

Proponents of a monetary explanation of the 1970s recessions point to two shortcomings of Hamilton’s econometric approach. His bivariate analysis of oil price and output movements cannot account for further endogenous variables through which oil prices might exert an indirect effect on economic activity. More specifically, Hamilton does not elaborate on the extent to which volatility in output can be attributed to the direct effects of oil price shocks or the monetary response to these shocks. Likewise, the link between oil price shocks and the high and persistent inflation, which was also characteristic of the time, remains blurred. Therefore, empirical studies by Bernanke et al. (1997) and Barsky and Kilian (2001, 2004), among others, have confronted these two central aspects in the debate over the oil price-macroeconomy relationship.

As a starting point, Bernanke et al. (1997) note that vector autoregression (VAR) models reveal the irrelevance of “erratic and unpredictable fluctuations” in monetary policy
for substantial variations in output. However, they proceed to observe that endogenous, systematic monetary policy changes may have significant impacts on economic activity. Indeed, “essentially all U.S. recessions of the past 30 years have been preceded by both oil price increases and a tightening of monetary policy, which raises the question to what extent the ensuing economic declines can be attributed to each factor” (Bernanke et al. 1997, p.93). This identification problem is illustrated by Figure 2, which plots the nominal interest rate, measured by the effective Federal Funds Rate, against the real oil price.

**Figure 2: Nominal Interest Rate vs. Real World Oil Price, 1970-2006**  
*Prices Adjusted by GDP Deflator $2005*

Author’s calculations. Source: EIA, Federal Reserve.
To assess the relative importance of both factors in explaining business cycle dynamics, Bernanke et al. flesh out the impacts of oil price shocks in the presence and absence of systematic monetary policy in a multivariate VAR system over the sample period 1965-95. They find that the 1970s recessions are not well explained by oil shocks if monetary policy is shut off, so that “the majority of the impact of an oil price shock on the real economy is attributable to the central bank’s response to the inflationary pressures engendered by the shock” (Bernanke et al. 1997, p.122). This result helps “to resolve the long-standing puzzle of the apparently disproportionate effect of oil price increases on the economy” (Bernanke et al. 1997, p.94). Besides, the authors find a substantial degree of subsample instability in the link between oil and the macroeconomy. However, they ascribe this observed instability to variations in the Federal Reserve’s reaction function over time,
which only supports their previous conclusion. Barsky and Kilian, in turn, extend their analysis to inflation and challenge the assumption that the 1970s’ oil shocks can account for the pattern of price-level increases observed during this period. They present evidence, that the petroleum price surges were preceded by “sharp and across-board increase in industrial commodity prices” (Barsky and Kilian 2001, p.14) which were not driven by cost push, but by demand shocks induced through expansionary monetary policy.

Building on these observations, CGG have addressed the role of monetary policy on macro-economic stability more systematically in their 2000 seminal paper. In line with the literature on rule-based monetary policy initiated by Taylor (1993) they assume that the behaviour of the Federal Reserve can be captured by an interest rate rule treating the Federal Funds rate as its instrument and targeting the output gap as well as the deviation of inflation from its target level (cf. section 3.4.3 for a brief discussion of monetary policy rules or Walsh 2003, Woodford 2003 for extensive analyses). Following the evidence on structural breaks in the Taylor rule over the sample period 1960-96 (cf. section 4.2), CGG estimate reaction functions for the pre- and post-1979 period, as the Federal Reserve’s policy is believed to have changed after Paul Volcker took over its chair in 1979. While monetary policy prior to 1979 is said to have been “accommodative” of inflation, policy under Volcker and Greenspan is found to have satisfied the Taylor principle, i.e. to have reacted sufficiently to inflationary pressures, raising not only the nominal, but also the real interest rate. This general pattern is reflected in Figure 3, which displays the real interest rate, given by the effective Federal Funds rate less inflation, versus inflation.

**Figure 3: Real Interest Rate vs. Inflation, 1960-2006**

Shaded areas indicate U.S. recessions as determined by the NBER.

Author’s calculations. Source: BEA, Federal Reserve.
Obviously, only the monetary regime after 1979 successfully raised the real interest rate to combat inflation (CGG note that the significant increase in the real rate in the beginning of the Volcker tenure leads the subsequent decline in inflation). To investigate whether this change in monetary policy is at the source of the switch from the 1970s’ stagflation to subsequent macroeconomic stability as suggested by the empirical evidence discussed above, CGG implement the estimated Taylor rules into a standard New Keynesian business cycle model. They then compare the responses of the output gap and inflation to shocks under these alternative rules and demonstrate that “holding constant the volatility of exogenous fundamental shocks, the economy exhibits greater stability under the post-1979 rule than under a rule that closely approximates monetary policy pre-1979.” (CGG 2000, p.149). Thus, the propensity of oil price shocks, identified as supply shocks in the CGG framework, to produce stagflation hinges on the aggressiveness of the monetary reaction.

This conclusion provides an alternative explanation for the seemingly reduced impact of petroleum price increases on economic aggregates after the 1970s. As mentioned in the preceding section, the latest surges in crude prices did neither lead to considerable contractions in output nor to runaway inflation. On the contrary, the above figure suggests that monetary policy might have reacted aggressively enough to prevent inflation from rising, thereby stabilizing output as well.

Despite the plausibility of the above argument, CGG’s modelling approach is confined in an important way with respect to the treatment of energy shocks. Lacking an appropriate microfoundation, it remains silent on the exact transmission channel of oil price shocks, the latter simply being introduced as any other supply disruption. The very transmission mechanism is, however, crucial to determine both the nature and amplitude of the dynamics precipitated by oil shocks (cf. Jones et al. 2004 and Hamilton 2003, who argues for non-linear effects of these shocks). Hence, CGG’s adhoc specification as such is inadequate to unravel the oil price-macroeconomy relationship. Therefore, it is the major objective of the present analysis to test the ability of a micro-founded and explicitly energy-augmented but otherwise similar New Keynesian model for the business cycle to reproduce both qualitatively and quantitatively CGGs results with respect to the relative importance of monetary policy in the face of oil price shocks.

The specific features of my model including the energy transmission channel are presented in the succeeding section.

III. The Energy-Augmented Model

General Features

The analysis of monetary policy has long been dominated by two different approaches to explaining the non-neutrality of money with respect to real economic activity. Macroeconomic models in the spirit of the Keynesian doctrine have relied on the existence of nominal rigidities to derive a Phillips Curve relation, which was expected to hold in the medium run. These models suggest that monetary policy should exploit the trade-off between output and inflation for output and employment objectives. However, the Lucas
Critique proved these models inadequate for policy analysis as they neglect the effects of expectations on behaviour. Owing to this critique, flexible price models were developed in which systematic and predictable monetary policy could have no real effects at all. Exploiting imperfect information on the part of economic agents rather than the existence of nominal rigidities, monetary policy could only affect output by unanticipated changes in money supply. As expectations were assumed to adapt quickly, these effects had to be essentially short-lived. This observation led to the policy irrelevance hypothesis.

Inspired by the Lucas Critique, another class of New Classical models, known as real business cycle (RBC) models, emerged. They attributed fluctuations in economic aggregates to supply or demand shocks in competitive markets instead of relying on nominal rigidities as the Keynesian models had done. In the absence of rigidities, i.e. with perfectly flexible prices and wages, these models left no room for policy interventions at all and money was believed to be neutral even in the short run. Technically, an important feature of these models is their rigorous micro-foundation, as the response of economic variables to shocks is entirely explained by the optimizing calculus of consumers and firms. The fact that the RBC literature has stressed the importance of exogenous disturbances as the primary cause of business cycle fluctuations explains the popularity of these models in the analysis of the oil price-macroeconomy relationship. Oil price surges are viewed as an emblematic example of exogenous supply-side disturbances which, much like adverse technology shocks, directly affect output, employment and inflation in line with the evidence presented by Hamilton and others. Hence, these models explicitly account for the role of energy by introducing it either directly as an input factor to production (cf., among others, de Miguel et al. 2003, Kim and Loungani 2002, Schmidt and Zimmermann 2005) or indirectly via capital utilization rates (Finn 1996, 2000).

Both, the New Classical models allowing for transitory real effects of unanticipated monetary policy changes as well as the RBC models have grown less popular with many economists, however, because of their failure to account for the significant and persistent non-neutrality of systematic monetary policy found in empirical studies (cf. the above discussion of the paper by Bernanke et al. 1997). Therefore, most recent macroeconomic models extend the RBC framework to include New Keynesian elements, i.e. wage and/or price rigidities to allow for a substantial influence of systematic monetary policy on real aggregates. The remarkable difference with respect to earlier Keynesian macro-models is, however, that their New Keynesian successors are rigorously micro-founded, so that they are consistent with the intertemporal optimizing behaviour of consumers and firms despite the incorporation of rigidities. The latter are attributed to institutional factors. In particular, many models drop the assumption of perfect competition in favour of a monopolistically competitive economy à la Dixit-Stiglitz with differentiated goods so as to give some price setting power to firms. This does not yet lead to monetary non-neutrality, however. Therefore, stickiness is introduced via staggered price (and/or wage) setting decisions, where prices are assumed to adjust infrequently following Calvo (1983), for instance. Such an economic setting is able to account both for sluggish adjustment in output and in inflation in response to disturbances.
Linear approximations to dynamic versions of these models, called Dynamic Stochastic General Equilibrium models, where monetary policy is captured by interest rate rules, have become the standard approach for policy analysis in recent years. They have also been used to address the oil price-macroeconomy relationship as in Rotemberg and Woodford (1996), de Fiore et al. (2006) or Kamps and Pierdzioch (2002), for example. Unlike CGG, who use a standard linearized DSGE model to begin with, these studies derive their linearized models from an extensive microfoundation. This approach allows them to explicitly develop the energy transmission channel alike the RBC models mentioned above and to identify the structural features which are crucial to generate any significant direct and/or indirect effects of oil price shocks. In the Rotemberg-Woodford model, for instance, energy works its way through the production function into the economy. However, as the energy share in production is small, these authors emphasize the importance of imperfect competition allowing firms to increase their mark-ups in response to a negative oil price shock to generate a substantive response in output. De Fiore et al. (2006) introduce energy as an input factor to an intermediate sector producing energy-loaded capital which, in turn, serves as an input in the final goods sector. At the same time, they assume that capital utilization is costly in terms of energy, so on the supply side they use the same transmission channel as Finn (1996 or 2000). Moreover, de Fiore et al. also consider the demand-side effects of oil shocks using an open economy model with an energy-producing and two energy-dependent countries. Their structure is very rich including many supply and demand disruptions besides the oil shock.

The energy-augmented model developed here draws on these examples. However, its primary aim is to replicate the features of CGG’s adhoc macro-structure in order to clarify whether this very structure is rich enough to allow for substantial economic volatility in response to oil price shocks under accommodating as opposed to aggressive monetary policy. Therefore, I develop a New Keynesian DSGE model with the following features:

Basically, the economy consists of households which consume, hold bonds and supply labour, firms hiring labour and producing differentiated final goods and a central bank which sets the nominal interest rate. Besides, there are a number of other important structural choices. First of all, capital is abstracted away as it is found to contribute little to the dynamics (cf. Walsh, p.231). Second, the economy is assumed to be closed as is the case in CGG’s version. This assumption has important consequences for the oil price transmission mechanism as it precludes the possibility of treating energy as a consumer good. Rather than being produced and sold by a specific sector, energy enters the economy as a non-produced input in the final goods sector whose relative price is exogenously determined. This is in line with the perception that the relative price of energy is exogenous to the economy and subject to political events as pointed out in the first section (a fact which goes unchallenged even by most advocates of the monetary explanation of business cycle fluctuations). With energy entering through the production function only, the analysis of oil shocks is confined to supply-side effects. These are aggravated by nominal rigidities in the goods market introduced via monopolistic competition and Calvo pricing, as is assumed by CGG. Different monetary policy regimes are finally represented by estimated
interest rate rules. These are adopted both from CGG and more recent estimation exercises.

In addition to this baseline specification of the energy-augmented model I develop a modified version including not only nominal rigidity in the goods market, but also real rigidity in the labour market. This modification suspends the ability of monetary policy to simultaneously stabilize inflation and output. With monetary policy being less effective in combating the adverse effects of oil shocks, this specification will flesh out the most severe consequences of oil price increases which can be represented with this model. If the order of magnitude of the increase in inflation and the contraction in output generated by oil price shocks in this setting is not close to real-world data, it must be concluded that the present model structure is not rich enough to capture the oil price-macroeconomy relationship quantitatively.

Still, in the concluding paragraph of their paper CGG assert that

“[…] the results we obtain in the context of our simple New Keynesian model are largely robust to using a much broader set of macroeconomic models. Two model features are critical, but they are features that are commonplace in most conventional macroeconomic frameworks. First, there must be an inverse relationship between output and the ex ante real rate (i.e., an IS-type relationship). Second, there must be a positive short-run link between output and inflation (i.e., a Phillips curve). Given these features, a monetary policy rule that accommodates inflationary pressure is, in general, more likely to bring about higher unconditional volatility of inflation and the output gap than otherwise.” (CGG 2000, p.177)

Whether this holds true with the micro-foundation proposed here will be revealed by the following analysis.

**Households**

The utility function

The demand side of the model is captured by an infinitely lived representative household having preferences over composite consumption \( C_t \) and leisure \( 1-N_t \). Total disposable time, which can be divided between leisure and labour \( N_t \), is normalized to 1. Hence, choosing leisure amounts to choosing the fraction of work time and vice versa. Hence, its optimization problem involves choosing composite consumption and leisure optimally so as to maximize the expected present discounted value of utility. Utility is assumed to be represented by a Constant Relative Risk Aversion (CRRA) function, where \( \sigma \) can be interpreted as the household’s preference for consumption smoothing or the intertemporal elasticity of substitution in consumption and \( \eta \) as the elasticity of labour supply.

\[
\max E_t \sum_{i=0}^{\infty} \beta^i \left[ \frac{C_{t+i}^{1-\sigma} - N_{t+i}^{1+\eta}}{1 - \sigma \cdot 1 + \eta} \right] \tag{1}
\]

The bundle of differentiated final consumption goods is produced by a continuum of monopo-listically competitive firms of measure 1, where individual firm \( j \) produces good \( c_{jt} \).
Where \((-\theta)\) is the price elasticity of individual demand (or the elasticity of substitution between different varieties). The smaller \(\theta\), the less individual demand responds to changes in individual prices. Hence, a small \(\theta\) corresponds to a high degree of market power. For, \(\theta \rightarrow \infty\), however, individual goods \(c_{jt}\) become near perfect substitutes, so that firms lose market power and perfect competition is ultimately induced.

The Optimization Problem

In the absence of explicitly modelled labour market imperfections the optimization problem is two-fold. First, the minimal cost combinations for a given level of \(C_t\) are derived. Then, given the costs of achieving a certain level of composite consumption, the household chooses \(C_t\) and \(N_t\) optimally so as to maximize the expected present discounted value of utility.

The minimization problem for any level of the final goods basket \(C_t\) can be written as:

\[
\begin{align*}
\min_{C_t} \int_0^1 p_{jt} c_{jt} \, dj & \quad \text{s.t.} \quad C_t = \left[ \int_0^1 c_{jt}^{\bar{\theta}} \, dj \right]^{\frac{\bar{\theta}}{\bar{\theta} - 1}} \\
\end{align*}
\]

This yields the usual equations for the aggregate price level of the final goods basket, i.e. the producer price index (PPI) of final goods, and individual demand for good \(j\):

\[
\lambda_t^C = P_t = \left[ \int_0^1 p_{jt}^{1-\bar{\theta}} \, dj \right]^{1-\bar{\theta}}
\]

Where \(\lambda_t^C\) is the Langrangian multiplier of the minimization problem. As households consume only one composite good, the price index for this good corresponds to the consumer price index, i.e. PPI and CPI coincide. Individual demand is:

\[
c_{jt} = \left( \frac{p_{jt}}{P_t} \right)^{-\bar{\theta}} \, C_t
\]

Once households have determined the minimal cost combinations for any given level of total composite consumption, they can proceed to the second stage of their decision problem and choose optimal levels of \(C_t\) and \(N_t\) subject to their budget constraint. In real terms, the latter is:
Total composite consumption and real holdings of bonds in period $t$ have to be financed by real income (real wage times hours worked), real money balances as well as bond holdings from the preceding period and real profits. The value of bonds from the previous period is increased by the nominal interest rate $i$.

Given this maximization problem, the following conditions for optimal household behaviour have to be satisfied in equilibrium:

\[
C_t = \beta (1 + i_t) E_t \left( \frac{P_t}{P_{t+1}} \right) C_{t+1}^{\sigma} \quad \text{or} \quad \hat{c}_t = E_t \hat{c}_{t+1} - \frac{1}{\sigma} (i_t - \hat{E}_t \hat{\pi}_{t+1})
\]

\[
\frac{N_t^\eta}{C_t^{\sigma}} = \frac{W_t}{P_t} \quad \text{or} \quad \hat{w}_t - \hat{p}_t = \eta \hat{n}_t + \sigma \hat{c}_t
\]

The Euler equation in (6) describes the optimal inter-temporal choice of consumption. Its linear approximation around the zero-inflation steady state will be used later to yield an IS-relation. It establishes the “inverse relationship between output and the ex ante real rate” required by CGG. Relation (7a) equates the marginal rate of substitution between labour supply and consumption (MRS) with the real wage.

**Firms**

In an environment of monopolistic competition firms produce differentiated goods with a common production technology. Additionally, prices are assumed to be sticky due to staggered pricing opportunities for firms à la Calvo.

Firms use two inputs for production, labour and energy, with capital being ignored. Output is furthermore subject to a random technology shock $\epsilon_t$.

\[
y_{jt} = \epsilon_t N_{jt}^{1-\alpha} e_{jt}^{\alpha}
\]

**Marginal Costs**

To facilitate subsequent analysis an expression for the firms’ marginal cost is derived before turning to the actual profit maximization problem. Much like households, firms will always minimize factor costs for a given level of production $y_{jt}$. Hence,

\[
\min_{N, \epsilon} \left( \frac{W_t}{P_t} \right) N_{jt} + \left( \frac{P_t^e}{P_t} \right) \epsilon_{jt} \simeq \epsilon \quad y_{jt} = \epsilon N_{jt}^{1-\alpha} e_{jt}^{\alpha}
\]

Where the relative price of energy is given by $\left( \frac{P_t^e}{P_t} \right)$.
From the first order conditions for this problem marginal costs can be derived as

\[ mc_t = 1 - \frac{\left( \frac{P_t^e}{P_t} \right)^{\alpha} \left( \frac{W_t}{P_t} \right)^{1-\alpha}}{\epsilon_t (1-\alpha)^{1-\alpha} \epsilon^\alpha} \quad \text{or} \quad \hat{c}_t = (1-\alpha)\left( \hat{y}_t - \hat{t}_t \right) + \alpha \left( \hat{p}_t^e - \hat{t}_t \right) - \hat{t}_t \]

Hence, marginal costs are driven by factor prices \( \frac{P_t^e}{P_t} \) and \( \frac{W_t}{P_t} \) as well as the level of technology \( \epsilon_t \).

Total costs are equal to

\[ TC_t = mc_t y_{jt} \]

**Optimal Pricing**

In the monopolistically competitive final goods market price setting is assumed to be staggered à la Calvo. An exogenously determined, random fraction \( (1-\psi) \) of all firms is allowed to adjust prices each period. Attempting to maximize the expected present discounted value of their current and future profits the adjusting firms will uniformly choose the optimal relative price according to:

\[ Q_t = \frac{P_t^*}{P_t} = \frac{E_t \sum_{i=0}^\infty \psi^i \beta^i C_t^{1-\eta} \left[ \left( \frac{P_{t+i}}{P_t} \right)^{\eta} mc_{t+i} \right]}{E_t \sum_{i=0}^\infty \psi^i \beta^i C_t^{1-\eta} \left[ \left( \frac{P_{t+i}}{P_t} \right)^{\eta-1} \right]} \]

Thus the optimal relative price is a weighted average of current and expected future marginal costs, where the probability of non-adjustment \( \psi \) amounts to a measure of the degree of nominal rigidity.

Assuming away such price stickiness and letting firms adjust prices every period \( (\psi = 0) \) the standard result of models of monopolistic competition is easily obtained. (10) then reduces to:

\[ \frac{P_t^*}{P_t} = \frac{\theta}{\theta-1} mc_t^f = \chi mc_t^f = 1 \quad ; \quad \chi > 1 \]

This shows that the relative price is set as a mark-up \( \chi \) over marginal costs (where the super-script \( f \) denotes the value of \( mc_t \) under flexible prices).

Finally, the aggregate price level of the final goods basket from (3) can now be expressed as an average of the optimal price charged by adjusting firms in period \( t \) and the aggregate price level of the previous period.

\[ P_t = \left[ (1-\psi) P_t^{*1-\theta} + \psi P_{t-1}^{1-\theta} \right]^{\frac{1}{1-\theta}} \]
Adjusting firms are randomly selected and will set the same optimal price \( p^* \) in period \( t \). As the remaining firms are by virtue of the random selection of adjusting firms also a random subset of the firms in \( t-1 \), their average price will equal the overall average price of the previous period, i.e. the aggregate price level of the final goods basket in \( t-1 \).

**The Linearized Model with Nominal Rigidities**

Before discussing the linearization of the energy-augmented model some assumptions are introduced to facilitate the subsequent derivations.

First, as energy price shocks are the focus of attention, technology shocks are assumed away by setting the technology parameter \( \varepsilon_t \) in the production function equal to one, i.e. equal to zero in terms of percentage deviations around the zero-inflation steady state. Thus, the technology term disappears in the linearized versions of both the production function (8) and the marginal costs equation (9). To be sure, this does not bias the results obtained in the presence or absence of sluggish price adjustment as the technology disturbance term is exogenous and therefore \( \hat{\varepsilon}_t^f = \hat{\varepsilon}_t \).

Second, energy is assumed to be a non-produced input factor to production in a closed economy setting, and its relative price is exogenously determined in order to investigate the impact of price shocks on the economic equilibrium. This exogeneity of the relative price of energy implies that it is the same both in the absence of any rigidities (i.e. with flexible prices) and in the presence of rigidities. Therefore, \( (\hat{p}_t^f - \hat{p}_t) = (\hat{p}_t^* - \hat{p}_t) \).

**The Output Gap**

The households’ Euler condition (6) for the optimal inter-temporal choice of consumption implicitly defines an IS-relation describing the demand side of the model.

In equilibrium aggregate demand is given by \( Y_t = C_t \) or \( \hat{y}_t = \hat{c}_t \). Therefore, replacing consumption in (6) by aggregate production, subtracting the flexible-price (alternatively, ‘natural’ or ‘potential’) output \( \hat{y}_t^f \) from both sides and extending the right side by \( E_t \hat{y}_t^f - E_t \hat{y}_t^f \) yields an expression for the output gap.

\[
x_t = E_t x_{t+1} - \frac{1}{\sigma} (i_t - E_t \pi_{t+1}) + E_t \hat{y}_t^f - \hat{y}_t^f \quad \text{where} \quad x_t = \hat{y}_t - \hat{y}_t^f \quad (13)
\]

In order to specify this relation, the flexible price output has to be determined. The latter can be extracted from the expression for percentage deviations of marginal costs from their steady-state value under flexible prices.

As a first step to expressing marginal costs in terms of output and the relative price of energy, the real wage in (7a) has to be reformulated. Substituting labour supply by the production function and using once more the AD-relation \( \hat{y}_t = \hat{c}_t \) it can be expressed as

\[
(\hat{w}_t - \hat{p}_t) = \frac{\eta}{1 - \alpha} (\hat{y}_t - \alpha \hat{c}_t) + \sigma \hat{y}_t
\]
However, energy has to be eliminated from the above equation, as the relative price of energy is exogenously given and the quantity of energy used in the economy will be determined by its factor input in production. This factor demand can be deduced from the optimality conditions of the firms’ cost minimization problem, i.e.:

$$e_{jt} = \alpha \mu_t y_{jt} \left( \frac{\hat{p}_t^e}{\hat{P}_t} \right)^{-1}$$

or $$\hat{e}_t = m \hat{c}_t + \hat{y}_{jt} - (\hat{p}_t^e - \hat{p}_t)$$ with $$m \hat{c}_t = \hat{\mu}_t$$

Note that the Lagrangian multiplier associated with cost minimization is equal to marginal costs. Substituting this relation for energy and replacing marginal costs with (9) gives real wages as

$$\left( \hat{w}_t - \hat{p}_t \right) = \frac{1}{1+\eta \alpha} \left[ (\sigma + \eta) \hat{y}_t + \eta \alpha (\hat{p}_t^e - \hat{p}_t) \right] \quad (7a*)$$

Reinserting this expression into the specification of marginal costs in (9) and neglecting the technology disturbance, real marginal costs finally become

$$m \hat{c}_t = \left[ \frac{(1-\alpha)(\sigma+\eta)}{1+\eta \alpha} \right] \hat{y}_t + \left[ \frac{(1+\eta \alpha)}{1+\eta \alpha} \right] (\hat{p}_t^e - \hat{p}_t) \quad (14)$$

Under flexible prices all firms would face the same conditions and would therefore charge the same price (cf. relation (11)), so that real marginal costs would equal their steady-state-value. Hence, the deviation from this value would be zero so that

$$m \hat{c}_t^f = 0 = \left[ \frac{(1-\alpha)(\sigma+\eta)}{1+\eta \alpha} \right] \hat{y}_t^f + \left[ \frac{(1+\eta \alpha)}{1+\eta \alpha} \right] (\hat{p}_t^e - \hat{p}_t)$$

Solving for the flexible price output gives

$$\hat{y}_t^f = -\alpha \left( \frac{1+\eta}{(1-\alpha)(\sigma+\eta)} \right) (\hat{p}_t^e - \hat{p}_t) \quad (15a)$$

So the natural output – or, more precisely, the deviation of output from its zero-inflation steady-state value in the absence of any rigidity - is entirely determined by the exogenous relative price of energy and would be equal to zero without shocks (if the technology disturbance parameter had been included, it would also show in the above equation and consequently reappear in the output gap relation).

Finally, combining (15a) with (13) the equation for the output gap can be written as
So apart from the expected output gap and the ex ante real interest rate, the current deviation of output from its natural rate depends on expected inflation in the exogenous relative price of energy.

The New Keynesian Phillips Curve

If relations (10) and (12) for the optimal price and the aggregate price level are approximated around a zero-inflation steady-state, a standard New Keynesian Phillips Curve can be derived. This log-linear curve describes how inflation in the PPI, and in this case also CPI, deviates from its steady-state value of zero with movements in marginal costs.

\[ \pi_t = \beta E_t \pi_{t+1} + \lambda \hat{m}_{ct}, \text{ where } \lambda = \frac{(1-\psi)(1-\psi\beta)}{\psi} \]

(17)

Thus, inflation in the goods market depends on expectations of future inflation and the percentage deviations of marginal costs around the steady state. Future inflation has to be taken into account by price setters in their optimization calculus as they can only adjust prices infrequently. Accordingly, the forward-looking nature of the Phillips Curve is induced by the nominal rigidities in the goods market which were introduced via Calvo pricing.

Yet, to be consistent with the demand side of the model the New Keynesian Phillips Curve is usually formulated in relation to the output gap. Therefore, marginal costs have to be rewritten in terms of the output gap, which can be achieved by subtracting the expression for marginal costs under flexible prices from the corresponding expression under sticky prices (14). This is done without loss of generality as marginal costs in the absence of rigidity are equal to zero.

\[ \hat{m}_{ct} = \frac{(1-\alpha)(\sigma+\eta)}{1+\eta\alpha} \]

(18)

(18) describes marginal costs solely as a function of the output gap. Hence, the New Keynesian Phillips Curve can ultimately be written in terms of expected inflation and the output gap.

\[ \pi_t = \beta E_t \pi_{t+1} + \kappa \pi_t, \text{ where } \kappa = \frac{(1-\psi)(1-\psi\beta)}{\psi} \frac{(1-\alpha)(\sigma+\eta)}{1+\eta\alpha} \]

(19)

Equation (19) is obtained from a specification of the optimizing behaviour of price setters. This explicit micro-foundation specifically allows for an understanding of the role of the structural parameters characterizing the economy like the discount rate \(\beta\), the degree of nominal rigidity \(\phi\), the energy share \(\alpha\), the households’ preference for consumption smoothing \(\sigma\) and the corresponding parameter for leisure \(\eta\).

Monetary Policy

In his 1993 seminal article “Discretion versus policy rules in practice” John B. Taylor
shows that simple interest rate rules are able to describe the empirical behaviour of monetary authorities reasonably well. More specifically, they provide a simple means to capture the conduct of central banks which are assumed to respond to fluctuations in the output gap and inflation via the short-term interest rate.

Building on Taylor’s findings but diverting from the initial specification of the monetary policy rule, numerous authors have stressed some features which improve its empirical fit. For instance, most estimation exercises find a significant degree of interest rate smoothing reflecting the central bank’s intention to gradually adjust its policy (so that agents may form reliable expectations about the bank’s behaviour) (cf., for instance, the estimation exercises presented in Table 4, p.55).

Moreover, Taylor’s original formulation “[...] does not take into due account the necessity of modelling the forward-looking behaviour of a central bank. In fact, due to the existence of a lag in the transmission mechanism of monetary policy, the monetary policy decision-makers should be guided by the outlook for prices rather than by the current developments in inflation.” (Gerdesmeier and Roffia 2003, p.12)

Other studies extend this reasoning to the output gap (given that the latter is difficult to measure so that estimations of the current output gap are hardly available) and yet other studies specify backward-looking rules taking into account lagged inflation and output (cf. Smets and Wouters 2003, 2005). Hence, the general formulation of the class of Taylor rules relevant to the present investigation allows for different target horizons.

Finally, another issue raised by recent work in the field of empirically estimated monetary policy rules points to the importance of the data set which is available to central banks at the time of their policy decisions. If, for example, their prediction of the current or expected output gap significantly diverts from ex post evaluations, estimations of policy parameters on the basis of revised data are prone to substantial bias. In light of these limitations many studies turn to the estimation of Taylor-like rules from real-time data in order to realistically approximate the actual informational basis of policy makers (cf. Gerdesmeier and Roffia 2004 or Bovin 2006, for some recent examples).

Given these considerations, the general structure of the estimated rules discussed in section 4.2 is captured by equation (20).

\[ i_t = i_{lag}^t + (1-i_{lag}^t) \left( \rho_x E_{t+h}^x x_{t+h} + \rho_y E_{t+h}^y r_{t+h} \right) \] (20)

The parameter \( i_{lag}^t \) describes the degree of inertia inherent to monetary policy; \( \rho_x \) and \( \rho_y \) quantify the central bank’s response to (expected) inflation and the (expected) output gap. Different quarterly target horizons are introduced by different choices of \( k \) and \( h \).

The Relative Price of Energy

In line with the general assumptions about the model economy, the relative price of energy is exogenously given as energy is assumed to be a non-produced factor input to production. Specifically, energy prices are taken to follow an AR(1) process, a standard assumption about the stochastic variables in models of this kind.
While $\varphi$ describes the persistence of energy price movements, shocks are captured by the parameter $\varepsilon_t^e$ with standard deviation $\sigma_e$.

General Equilibrium
With the interest rate being the monetary policy instrument, the equilibrium conditions of the energy-augmented model are given by (6), (10), (12) and the AD-relation $Y_t = C_t$. The fully specified linearized DSGE model derived from these conditions is summarized below.

**Table 1: The Linearized Model with Nominal Rigidities**

\[
\begin{align*}
x_t &= (E_t x_{t+1}) \left\{ \frac{1}{\sigma} \left[ (\varepsilon_t - E_t \pi_{t+1}) - \alpha \left( \frac{1+\eta}{1-\alpha(\sigma+\eta)} \right) \left[ (E_t \hat{\pi}_t^e - \hat{\pi}_{t+1}^e) - (\hat{\pi}_t^e - \hat{\pi}_t) \right] \right] \right\} \\
\pi_t &= \beta E_t \pi_{t+1} + \kappa x_t \quad \text{where} \quad \kappa = \left[ \frac{(1-\psi)(1-\psi\beta)}{\psi} \right] \left[ \frac{(1-\alpha)(\sigma+\eta)}{1+\eta\alpha} \right] \\
i_t &= i_{lag}^t i_{lag} - (1-i_{lag}) \left[ \rho_x E_t \pi_{t+1} \pi_{t+k} + \rho_y E_t \pi_{t+k} x_{t+k} \right] \\
\hat{\pi}_t^e - \hat{\pi}_t &= \varphi(\hat{\pi}_t^e - \hat{\pi}_{t+1}^e) + \varepsilon_t^e \quad \text{with} \quad \varepsilon_t^e \sim N(0, \sigma_e)
\end{align*}
\]

The demand side of the model is entirely characterized by the output gap and the supply side by the New Keynesian Phillips Curve. The monetary policy rule closes the model with respect to the nominal interest rate and the only exogenous parameter is the relative price of energy.

Before extending the basic model, I would like to draw attention to an important feature of the above formalization, namely the absence of a trade-off between the monetary authority’s objectives of price and output stability. The reason for this “divine coincidence” (Blanchard and Gali (forthcoming)) is that in the absence of real rigidities marginal costs can be entirely characterized by the output gap. Hence, energy price shocks, which drive a wedge between actual and flexible-price output in the face of nominal rigidities, do not have a separate, independent effect on marginal costs and therewith inflation. Offsetting deviations in the output gap at all times therefore amounts to stabilizing inflation as well. In short,

“if the output objective is interpreted as meaning that output should be stabilized around its flexible-price equilibrium level, then […] the central bank can always achieve a zero output gap (i.e., keep output at its flexible-price equilibrium level) and keep inflation equal to zero." (Walsh 2003, p.252)

To facilitate subsequent analysis, it is helpful to furthermore discuss how the choice of
some non-standard or empirically debatable parameters – ceteris paribus - affects equilibrium dynamics. First of all, the degree of nominal rigidity captured by $\psi$ is crucial for the extent to which inflation responds to fluctuations in the current output gap. The reason for this is that fewer firms are able to adjust their current prices to the prevailing economic conditions, i.e. to changes in marginal costs, the higher the probability of non-adjustment, $\psi$, becomes (cf. equation (12)). In other words, if firms can adjust their prices less often, current conditions become less important. This relation can be directly observed from the initial formulation of the Phillips Curve in (17), where the factor associated with marginal costs, $\lambda$, declines with $\psi$. Therefore, the impact of the current output gap on inflation is moderated by increasing stickiness.

Second, as outlined below in section 4, the value of the elasticity of labour supply, $\eta$, varies substantially between studies, so I briefly comment on its impact in the model. The higher its value, the stronger workers will adjust their labour supply to changes in real wages (cf. equation (7a)). Hence, more volatility is induced in production which, through (14) or (18), respectively, translates into the variability of marginal costs. Consequently, the sensitivity of inflation with respect to changes in real conditions rises.

However, with the inter-temporal elasticity of consumption being equal to the standard choice of one, the output gap is completely insensitive to the elasticity of labour supply. Furthermore, its impact on the output gap is negligible even with values of $\sigma$ substantially above one.

Finally, the energy share in production, $\alpha$, naturally is the most crucial factor in the analysis of energy price shocks. The more important this input in production, the larger the elasticity of the output gap with respect to energy price shocks becomes. The straight-forward reason for this is, that energy price shocks drive a wedge between actual and potential output. More explicitly, marginal costs, expressed in terms of factor prices, are directly affected by energy price shocks. Increasing marginal costs, for instance, cause firms to adjust their prices upward whenever given the opportunity to do so. Higher inflation devalues real wages thereby reducing labour supply and ultimately resulting in decreasing production. However, with nominal rigidities prices adjust more slowly, so that for a given energy shock actual output declines less than potential output, leaving a positive gap.

A more thorough analysis of the impulse responses to an energy price shock, which builds on the above considerations, is deferred to section 5.

The Linearized Model with Real and Nominal Rigidities

As an alternative to the above specification a second model exhibiting not only nominal, but also real rigidities in form of sticky wages is introduced next. The reason for establishing yet another source of distortions is that nominal rigidities might not be sufficient for energy price shocks to generate substantial contractions even in the presence of an accommodating monetary policy. As noted, extending the analysis of the basic model to the modified one will allow me to judge whether the economic environment discussed here is rich enough to account for the adverse effects of energy shocks or whether substantial structural changes as regarding the production technology, the consumption basket or the openness of the economy would have to be incorporated.
Following Blanchard and Gali (forthcoming), I assume that real wages adjust sluggishly because of some labour market imperfections. To be specific, the current real wage is introduced as a weighted average of the lagged and the “optimal” real wage determined by the households’ marginal rate of substitution between consumption and labour. The weight assigned to the past real wage, $\rho$, can be viewed as a measure of real rigidity, analogous to the parameter of price staggering $\psi$.

$$\hat{w}_t - \hat{p}_t = \rho \hat{w}_t - \hat{p}_t + (1 - \rho) \text{mrs}$$

(22)

Blanchard and Gali

"[...] view equation [22] as an admittedly ad-hoc but parsimonious way of modelling the slow adjustment of wages to labor market conditions, as found in a variety of models of real wage rigidities, without taking a stand on what the "right" model is. [...] The important assumption underlying equation [22] is that the slow adjustment be the result of distortions rather than preferences [...]" (Blanchard and Gali (forthcoming), p.9)

Two features of the above approach are worth noting. First, Blanchard and Gali (like many other authors including Gali 2002, Smets and Wouters 2003, de Walque et al. 2005 and de Fiore et al. 2006) also derive real wage rigidity from staggered wage setting decisions similar in spirit to Calvo pricing. However, they conclude that the results of the more explicitly micro-founded approach are basically identical to the ad-hoc-version of sluggish real wage adjustment. Therefore, I rely on the simpler introduction of real rigidities (as no precision would be added to the oil shock transmission channel by an explicit micro-foundation of this feature).

Second, Blanchard and Gali emphasize that wage rigidity is a consequence of distortions and not preference shocks. This is crucial, since a taste shock which changes the MRS, for instance,

"[...] also affects the flexible-price equilibrium level of output. [...] Thus, if the output gap variable in the inflation adjustment equation is correctly measured as the deviation of output from the flexible-price equilibrium level, [the shock] no longer has a separate, independent impact on [inflation]." (Walsh 2003, p.253-254)

However, frictions in the labour market resulting in sluggish adjustment of the above kind do not affect potential output and thus drive a further wedge between output in the presence and in the absence of rigidities which is going to be reflected in the Phillips Curve specification.

As real wages are sticky and not set optimally anymore, the consumers’ MRS is no longer equal to the actual real wage. Instead, it represents the optimal wage obtained from the households’ optimization problem. Hence, condition (7a) has to be rewritten as

$$\frac{N_i^\eta}{C_i^{\sigma-\eta}} = \left(\frac{W_t}{P_t}\right)^* \text{ or mrs} = (\hat{w}_t - \hat{p}_t)^* = \eta \hat{w}_t + \sigma \hat{c}_t$$

(7b)
The modifications this new specification of real wages inflicts on the model are presented below after discussing potential output in the modified framework.

Potential Output and the Output Gap

In the absence of both nominal and real rigidities wages still adjust optimally so that (7a) describes not the optimal, but the actual current wage. Hence, marginal costs can be derived as in the flexible prices case and potential output in the modified model will accordingly be equal to the flexible price output of the basic model, i.e.

\[
\hat{y}_t^{\text{pot}} = \hat{y}_t^f = -\alpha \left( \frac{1+\eta}{1-\alpha(\sigma+\eta)} \right) \left( \hat{p}_t^s - \hat{p}_t \right) \tag{15b}
\]

Furthermore, the demand side of the basic model is not subject to any changes albeit the introduction of labour market imperfections. As the sluggish adjustment of wages is ascribed to (unmodelled) distortions rather than the structure of tastes, it is not taken into account in the optimization rationale of the households. So the Euler equation (6) derived above and the IS-relation implicitly defined by it are still valid. Both, the IS-relation and potential output being unaffected by real wage stickiness, the specification of the output gap as laid out in (16) is unaltered as well. The structural description of the demand side, therefore, remains the same in both versions of the energy-augmented model. Nevertheless, the dynamics of the output gap under sticky wages will, of course, respond to the new economic conditions as the analysis of the impulse response functions to energy price shocks will reveal.

The New Keynesian Phillips Curve

Real marginal costs are the channel through which sluggishly adjusting wages affect the economic equilibrium. Let me recall the expression for marginal costs obtained from the firms’ cost minimization problem:

\[
m \hat{c}_t = (1-\alpha)(\hat{w}_t - \hat{p}_t) + \alpha(\hat{p}_t^s - \hat{p}_t) \tag{9}
\]

In the absence of real rigidity the real wage in this expression was replaced by a formulation of the MRS in terms of output and real energy prices and marginal costs were finally expressed by means of the gap between the sticky prices output and the flexible prices output. However, real wages behave differently now so that this procedure is not feasible anymore.

Still, the real wage in (22) can be specified with regard to output and the relative price of energy. This is achieved by first inserting the optimal wage from (7b). If, then, labour is replaced with the corresponding expression from the production function, energy by the equation for energy demand, consumption by output making use of the AD-relation and marginal costs with equation (9), the evolution of real wages is given by:

\[
(\hat{w}_t - \hat{p}_t) = \frac{\rho_w}{1+(1-\rho_w)\eta \alpha} (\hat{w}_{t-1} - \hat{p}_{t-1}) + \frac{1-\rho_w}{1+(1-\rho_w)\eta \alpha} \left[ (\sigma+\eta) \hat{y}_t + \eta \alpha(\hat{p}_t^s - \hat{p}_t) \right] \tag{7b*}
\]
Equation (9) together with (7b*) defines the initial Phillips Curve (17) in terms of the lagged real wage, the deviation of output from its steady-state value and the relative price of energy. To compare this Phillips Curve relation with the one obtained in the basic model, one could aim at introducing the output gap in (7b*). However, even if the flexible price output was subtracted from this expression, the relative price of energy would not disappear. Therefore, as a result of staggered wage setting the New Keynesian Phillips Curve includes a stochastic term, the energy price shock, which embeds a trade-off between inflation and output stabilization alike the cost shock in the ad-hoc specifications of the Phillips Curve in Gali (2002, p.29f) or Walsh (2003, p.252ff). To the extent that the exogenous disturbance is independent of fluctuations in potential output and is thus not captured by the gap between actual and potential output, stabilizing this gap would no longer be sufficient to keep inflation equal to zero. Hence, the implications for monetary policy in the modified framework are in sharp contrast to the model exhibiting only nominal rigidities.

General Equilibrium

The following table summarizes the linearized version of the modified model:

\[
\begin{align*}
\hat{y}_t^\text{pot} &= \hat{y}_t^f = -\alpha \left( \frac{1+\eta}{(1-\alpha)(\sigma+\eta)} \right) \left( \hat{p}_t^e - \hat{p}_t \right) \\
x_t &= (E_t x_{t+1}) - \frac{1}{\sigma} (i_t - E_t \pi_{t+1}) - \alpha \left( \frac{1+\eta}{(1-\alpha)(\sigma+\eta)} \right) \left[ (E_t \hat{p}_t^{e+1} - \hat{p}_{t+1}) - (\hat{p}_t^e - \hat{p}_t) \right] \\
\pi_t &= \beta E_t \pi_{t+1} + \lambda m \hat{c}_t \quad \text{where} \quad \lambda = \frac{(1-\psi)(1-\psi\beta)}{\psi} \\
m \hat{c}_t &= (1-\alpha)(\hat{w}_t - \hat{p}_t) + \alpha (\hat{p}_t^e - \hat{p}_t) \\
(\hat{w}_t - \hat{p}_t) &= \frac{\rho_w}{1 + (1-\rho_w)\eta\alpha} (\hat{w}_{t-1} - \hat{p}_{t-1}) + \frac{1 - \rho_w}{1 + (1-\rho_w)\eta\alpha} \left[ (\sigma + \eta) \hat{y}_t + \eta \alpha (\hat{p}_t^e - \hat{p}_t) \right] \\
i_t &= i_{t-1} + (1-i_{t-1}) \left[ \rho_y E_{t+h} \pi_{t+h} + \rho_y E_{t+h} x_{t+h} \right] \\
\hat{p}_t^e - \hat{p}_t &= \varphi (\hat{p}_{t-1}^e - \hat{p}_{t-1}) + \epsilon_t^e \quad \text{with} \quad \epsilon_t^e \sim N(0,\sigma_e)
\end{align*}
\]

Potential output, the output gap, the interest rate rule and the exogenous relative price of energy are unaltered. However, as developed above, unlike the basic model this new specification comprises a more complex building block for inflation.

The description of inflation in this second formulation notably differs from the first in
that it incorporates the degree of newly introduced real wage rigidity, $\rho_w$. Increasing stickiness translates into a larger weight of the lagged real wage in the equation for the current real wage. That means that the effect of energy price shocks on real wages builds up more slowly and hence delays the effect on output. Eventually, this delayed response of output to the exogenous disturbance may result in a negative output gap and therefore significantly alter or even reverse equilibrium dynamics, inducing higher volatility in the aggregate variables. These effects increase in the degree of wage stickiness.

**IV. Model Calibration and Policy Rules**

**Structural Parameters**

In section 3 I have developed a baseline New Keynesian DSGE model augmented by energy as an input to production and discussed the qualitative impact of several parameters on equilibrium dynamics. As a second step, I calibrate the model in order to examine the quantitative behaviour of the economy and the performance of different monetary policy rules as specified below. To this end impulse responses to exogenous variations in the relative price of oil will be generated in the subsequent section.

The values pertaining to preference and technology parameters as well as the stochastic properties of the energy price shock are summarized below:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>US</th>
<th>EA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discount factor</td>
<td>$\beta$ (beta)</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Elasticity of labour supply</td>
<td>$\eta^{-1}$ (eta)</td>
<td>0.36</td>
<td>0.40</td>
</tr>
<tr>
<td>Inter-temporal elasticity of substitution</td>
<td>$\sigma^{-1}$ (sigma)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy share in production</td>
<td>$\alpha$ (alpha)</td>
<td>2.0%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Probability of non-adjustment</td>
<td>$\psi$ (psi)</td>
<td>0.70</td>
<td>0.75</td>
</tr>
<tr>
<td>Energy Price Shock</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autocorrelation</td>
<td>$\varphi$ (phi)</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Standard error</td>
<td>$\sigma_e$</td>
<td>8.5%</td>
<td>8.5%</td>
</tr>
</tbody>
</table>

All parameters are estimated on a quarterly basis and are assumed to be stable over the period considered. Where not otherwise specified, the structural parameters in the US and Euro area are assumed to coincide. Preference parameters correspond to standard choices in the business cycle literature. Thus, the quarterly discount factor is assumed to equal $\beta = 0.99$, implying an annual steady-state real interest rate of 4 percent. Similarly, a value of $\sigma^{-1} = 1.00$ for the elasticity of inter-temporal substitution in consumption can be found in
most RBC and New Keynesian models without habit formation (estimations of models with habit formation tend to generate a significantly smaller value as in Smets and Wouters (2003), de Walque et al. (2005), for example). There is less agreement about the value of the elasticity of labour supply. While Smets and Wouters (2003) and de Walque et al. (2005) find values of 0.40 - 0.53 for the euro area (EA) and 0.34 – 0.38 for the US, the RBC literature (cf. Cooley and Prescott 1995, for instance) typically assumes a much higher value. In order to replicate the empirical conditions as closely as possible I choose values close to the above estimates, i.e. 0.40 for the EA (which is also used by the ECB) and 0.36 for the US.

Taylor (1998a) evaluates econometric estimates and survey evidence on the stickiness of wages and prices. He concludes that even “for the range of inflation rates observed in the developed economies in the 1970s the average duration of wages and prices remained high” (p.26). Prices are found to remain constant for approximately one year on average which is replicated in the model by assuming a value of $\psi = 0.75$ for the probability of non-adjustment in the EA. Empirical studies find less nominal rigidity for the US, however (cf. de Walque et al. (2005) who find a value of 0.74 for the EA and 0.70 for the US). The energy share cannot be adopted as easily from other studies which have explicitly accounted for the role of energy in production, since most of these models introduce energy via a capital-energy bundle. However, de Fiore et al. (2006) report oil shares of between 1.00 and 1.5% for the EU and between 1.5 and 2.0% for the US. I adopt the upper bounds of these parameters to construct a worst case scenario with respect to energy price shocks.

Following de Walque et al. (2005), the exogenous process of the relative energy price is specified with a high autocorrelation coefficient of $\varphi = 0.96$ and a standard error of $\sigma_e = 8.5\%$ (cf. also Rotemberg and Woodford (1996)).

Finally, the degree of stickiness in real wages is varied to investigate the impact of increasing real frictions on equilibrium dynamics.

**Estimated Monetary Policy Rules**

To answer the central question of this paper, namely whether the interplay of monetary policy, oil price shocks and macroeconomic fluctuations examined by CGG (2000) can be captured in a model with an explicit micro foundation accounting for the role of energy as an input factor to production, I compare monetary policy rules of different episodes which have witnessed oil price shocks.

CGG (2000) distinguish the monetary policy adopted by the Federal Reserve in the 1960s and 70s from the regime of Volcker and Greenspan from 1979 to 1996. I extend this analysis in two respects. First, as laid out in the introductory section, I expand the period under consideration to contrast the monetary policy in the face of the 1970s oil price shocks with more recent estimated policy rules to be able to account more persuasively for the oil price surges in the aftermath of the second war in Iraq. Second, all monetary policy regimes under consideration are evaluated in light of their performance in economies exhibiting nominal (as in CGG) as well as both nominal and real rigidities.

A survey of recently estimated monetary policy rules for the US and the EA reveals that despite broad consensus on some general features, there is considerable disagreement
about the nature of the rule and some parameter values (cf. the output gap coefficient, for instance). These differences are due to the choice of the data set on the grounds of which the rule is estimated (i.e. revised vs. real-time data; different measures of inflation and the output gap), the estimation method itself as well as assumptions about the nature of the policy function (forward- vs. backward-looking, inclusion of additional parameters), target horizons (from 1 quarter to 1 year) and discrete structural breaks vs. gradual changes (revealed by the use of time-varying coefficients). Investigating the occurrence of structural changes in the policy rule of the Federal Reserve, Jean Bovin even concludes that

“[t]he rich pattern of changes in the response to inflation and unemployment is not well captured by a discrete shift, occurring in 1979. The changes are neither unidirectional nor synchronized across parameters. Moreover, the transition under Volcker appears to have taken place over a few years, with the most important changes occurring between 1980 and 1982. More than real-time data issues, the complex nature of the changes might explain why conflicting results have been reported in the literature.” (Bovin 2006, p.1171)

In light of the empirical and conceptual difficulties discussed above, all results come with a grain of salt and the explicatory power of the estimated rules should not be overestimated. However, a significant degree of interest rate smoothing, the realization of the Taylor principle in more recent policy rules, the general pattern of changes in monetary policy over time as well as broad periods of stable monetary policy should be emphasized as robust findings of all studies.

For the purpose of simulation, US monetary policy in the pre-Volcker period will be approximated with the rule found by CGG 2000 with a one-quarter forward-looking horizon for the inflation target \((k = 1; h = 0)\). Policy in the Volcker and Greenspan era is identified with the rule estimated by Chadha et al. 2003 using the GDP deflator \((k = 1; h = 1)\). For the ECB I adopt the Taylor rule estimated by Gerdesmeier and Roffia 2004 \((k = 4; h = 0)\).

All policy rules relevant to this analysis are forward-looking, diverting only in their respective target horizons for inflation and the output gap. Note, besides, that the estimate for the inflation coefficient in both US rules is based on the GDP deflator, whereas the corresponding coefficient in the EA rule relies on a CPI measure (which might introduce a potential bias in the comparison of the two currency areas).

V. Simulation Results

The impulse response functions of differently specified economies hit by energy price shocks are presented in this section. As a first step, the monetary policies adopted by the Fed during the pre-Volcker era and the Volcker-Greenspan tenure are contrasted in an economy exhibiting stickiness in prices (as in CGG 2000). Then, the performance of both policies is tested against an environment with sluggish adjustment in wages in addition to rigid prices. ECB policy is considered as well in both economic settings and is found to yield very similar results.

In both environments exogenous energy price shocks are the only source of deviations
of economic aggregates from their steady-state values. These shocks enter the economy via the supply side as they drive up the firms’ production costs. Increases in marginal costs will gradually translate into rising consumer prices given that monopolistically competitive firms set their sales prices as a mark-up over marginal costs. However, the speed of adjustment of prices as well as the behaviour of the other aggregates depends on the degree of nominal and real rigidity. Therefore, the exact transmission mechanism of energy shocks and the dynamics set off by them are discussed in detail in the following scenarios.

5.1 Oil Price Shocks and Nominal Rigidities

Accommodating Monetary Policy

Most estimations of the policy parameters for the pre-Volcker era reported in CGG 2000, including those of the robustness analysis, yield an inflation coefficient, $\rho_\pi$, significantly below one.

"Values of $[\rho_\pi]$ in this range lead to indeterminacy of the equilibrium, and raise the possibility of fluctuations in output and inflation around their steady state values that result from self-fulfilling revisions in expectations. The intuition is straightforward: with $[\rho_\pi]$ below unity, a rise in anticipated inflation leads to a decline in the real interest rate. The decline in the real rate then stimulates aggregate demand which, in turn, induces a rise in inflation. The initial rise in expected inflation thus becomes self-confirmed." (CGG 2000, p.171)

Still, some standard errors are too large to rule out the possibility of an inflation coefficient equal to or slightly above one (which would hold the real rate roughly constant). Additionally, Jondeau and Bihan (2000) actually estimate a value slightly above unity for $\rho_\pi$ in the pre-Volcker era, using a different sample than CGG. In this limit case, endogenous fluctuations due to accommodating monetary policy do not occur. However, the economy will still exhibit much more volatility than under an aggressive monetary regime (cf. CGG 2000, p.174ff). In the following analysis the dynamics and standard errors of economic aggregates in the case of near-indeterminacy are compared to the case of determinacy due to aggressive policy.

Taking into account what has just been said, $\rho_\pi$ is set equal to one to simulate accommodating monetary policy in the pre-Volcker period whereas the persistence and output coefficient are taken from the estimated rules in Table 4. The impulse response functions in Figure 4 reveal the reaction of a model US-economy characterized by monopolistic competition in the final goods market, sticky prices and an accommodating monetary policy as estimated by CGG 2000 for the period of 1960-79 to a shock in the relative price of energy of 8.5% above its steady-state value.

The basic pass-through of the energy price shock is the same, whether the economy exhibits nominal rigidities or not. Increasing factor prices drive up marginal costs which translates into a rise in consumer prices as described above. However, in the presence of staggered price setting the initial price adjustment to the shock is smaller than under flexible prices, as only a fraction $(1-\psi)$ of firms may adjust their prices. With nominal prices rising more slowly, real wages, (which initially respond positively to the shock presum-
ably due to a substitution of energy with labour in production), decline less. As real wages are equivalent to the opportunity cost of leisure this decelerated decrease implies that households reduce their labour supply more reluctantly than in the flexible prices case so that actual production declines more sluggishly as well.

**Figure 4: Pre-Volcker Monetary Policy and Nominal Rigidities (US)**

This is reflected in the impulse response functions by the fact that the full response of output builds up gradually. The wedge which is thus driven between actual and potential output in the face of an adverse supply shock is captured by the positive output gap.

Both, the current positive output gap and the expectation of future inflation, due to the persistence of the shock, induce the monetary authorities to raise, the nominal interest rate, though not enough for an increase in the ex ante real rate. Hence, the return of economic aggregates to their steady state is rather driven by the diminishing effect of the oil shock.

Quantitatively, the output gap is widened by about 0.1% on impact of the shock (and exhibits a standard deviation of 0.11%). This is well in line with de Walque et al. who estimate an impact on GDP of around 0.1% in the first and second year after the shock (cf. de Walque et al. 2005, p.24). Inflation initially deviates by 0.122% from its steady-state value of zero (with a std. dev. of 0.23%) and the interest rate response gradually builds up to a maximum value of 0.05% after about one year (the std. dev. being 0.2%).

**Aggressive Monetary Policy**

In order to evaluate the stabilizing potential of monetary policy under Volcker and Greenspan, who reacted more aggressively to fluctuations in inflation and the output gap, the impulse response functions of an economy where the policy rule is characterized by
the estimates of Chadha et al. (2003) are retrieved below.

**FIGURE 5: GREEFSNAP MONETARY POLICY AND NOMINAL RIGIDITIES (US)**

Technically, the dynamics are the same as presented above. However, in contrast to the case of accommodating monetary policy the more aggressive regime raises the ex ante real interest rate in response to expected inflation rather than holding it constant. This has countercyclical effects as according to equation (16) the output gap declines in the ex ante real rate. The reason is that households have an incentive to postpone current consumption to the future with increasing interest rates, as long as the gains through higher returns on savings are not offset by future inflation, i.e. as long as not only the current nominal, but the ex ante real rate increases (cf. Euler equation (6)). Due to the equilibrium identity of consumption and output, current actual output will decline along with consumption. This is not true, however, of potential output, which, according to equation (15a), only diverts from its steady-state value due to exogenous disturbances. Because of the direct impact of the shock potential output had initially declined faster than actual output leaving a positive gap. As actual output now decreases further with the ex ante real rate, this gap is reduced. Finally, a smaller output gap eventually exerts a dampening impact on inflation as given by (19).

Besides, as all agents in the economy are familiar with the model structure including the features of the monetary policy regime, they anticipate the aggressive reaction to inflationary pressure so that they correct their expectation of future inflation downwards. Through the Phillips Curve relation these expectations also translate into diminishing actu-
al inflation. Hence, both volatility and persistence in inflation as well as volatility in the output gap are reduced due to the stabilizing effect of monetary policy on expectations. Moreover, for a given nominal rate the smaller expected inflation in an aggressive monetary regime additionally increases the \textit{ex ante} real rate. Therefore, the nominal rate has to increase by less to achieve a given degree of stabilization than under a feeble regime.

These findings are reflected in the quantitative behaviour of the economic aggregates. The output gap increases by 0.02% on impact (with a std. dev. of 0.02%) and inflation by 0.02% (std. dev. of 0.03%). Hence, the volatility in both aggregates is reduced by more than factor five. Moreover, by comparison the “tail” of the impulse response of inflation in the first scenario is much longer than in the second. This suggests that monetary policy is crucial in accounting for the persistence of inflation. Finally, the interest rate builds up gradually again, with a maximum value of 0.009% (std. dev. of 0.03%) - far less than under the accommodating regime.

Thus, the most central findings of Bernanke et al. (1997) and CGG (2000) as regards the importance of systematic monetary policy in business cycles can be reproduced with the basic energy-augmented model. Generally, the above analysis reveals that it is unlikely that energy price shocks alone were responsible for the economic contractions in the 1970s. Rather,

“[e]ven assuming that the oil price shocks played a critical role, the ultimate impact of these disturbances on output and inflation depends very much on the feedback monetary policy rule that is in place.” (CGG 2000, p.168)

Specifically, holding constant the volatility in exogenous disturbances, output fluctuations significantly diminish and inflation becomes much less volatile and persistent with aggressive monetary policy, as it is conducted today by the Fed or the ECB, for instance. This result may help explain the relative stability of inflation and economic activity even in the face of sharp oil price surges after 2004.

However plausible these findings appear, the magnitude of the economic slowdown and the degree of inflationary pressure caused by the energy price shock are quantitatively altogether negligible - independently of the monetary regime (the inflation rate rose up to 11% and GDP decreased by over 2% in response to the shock compared to 0.122% and 0.1% in the model). Hence, the quantitative differences in the volatility of economic aggregates under the two monetary reaction functions are minor as well. Accordingly, the basic energy-augmented model is only qualitatively and not quantitatively capable of capturing the link between energy shocks and monetary policy. Therefore, I will extend the basic model to a richer environment embracing real rigidity to investigate whether the fundamental pass-through of energy price shocks proposed here is appropriate to capture the quantitative dimensions of energy price shocks.

\textbf{Oil Price Shocks and Real Rigidities}

\textbf{Accommodating Monetary Policy}

The inclusion of real rigidities has two consequences in the model. First, it leads to the already much discussed abolition of the divine coincidence of simultaneous output and inflation stabilization and second, the response of output is further delayed by the sluggish
reaction of real wages. The question now is, whether these modifications result in stronger volatility in the economic aggregates. The impulse response functions below describe the behaviour of an economy with a very high degree of real rigidity of $\rho_w = 0.9$, i.e. where 90\% of the current real wage are explained by the previous real wage. This extreme degree of real rigidity is assumed to flesh out its consequences as clearly as possible.

The dynamics in the flexible prices case being unaffected, potential output responds exactly as before. As far as the other variables are concerned, their initial response is the same as well. The energy shock drives up marginal costs and therewith inflation, leading to a devaluation of real wages, which in turn provokes a decline in labour supply so that output finally decreases. As before, the decline in real wages is slowed down by the presence of rigidities, so that actual output declines less than potential leaving a positive output gap. However, in contrast to the preceding example there are now two sources for the sluggish adjustment of real wages. This is why the build-up of the full effect of the shock on the real wage takes much more time (about 10 quarters as opposed to about 4 without real rigidity). While actual output continues to fall because of the slow transmission of the oil shock through the real wage, potential output already returns to its steady-state value of zero. Thus, a negative output gap emerges which calls for a smaller interest rate. In fact, the interest rate declines much faster than without real rigidity. Hence, the pattern of the adjustment has substantially changed.

Nonetheless, real rigidity does not seem to contribute to a much stronger quantitative
response in aggregate variables. Thus, the initial response of the output gap and inflation has been 0.09% and 0.128% and the standard deviation 0.1% and 0.23%, respectively. The interest rate peaks at 0.06% and exhibits a std. dev. of 0.19%. These values almost coincide with those encountered in the nominal rigidity case.

Aggressive Monetary Policy

Under the monetary regime estimated for the Volcker-Greenspan tenure the following impulse responses are retrieved in the presence of real rigidities.

As before, the monetary policy response to surges in inflation and the output gap is sufficient to increase the ex ante real interest rate in this second scenario. The dampening effect this policy reaction exerts on output contributes to the decline of the latter, adding to the effect of the delayed response to the energy price shock due to sticky wages. Thus, actual output eventually drops further below potential output than under an accommodating monetary policy, which is also reflected in the output gap. The sharper decrease in actual output due to the increase in the ex ante real rate translates into lower inflation, which even becomes negative, according to equations (17), (9) and (7b*). The combination of effects - the negative output gap and negative inflation at the same time - in turn call for a negative nominal rate. To summarize, monetary policy in the second scenario of the modified model exerts again a much stronger influence on the path of return to the steady state. But in comparison to the basic model the dynamics displayed by the modified version exhibit much more volatility so that monetary policy is less successful and the economy far more instable.
This qualitative instability is reflected in the amplitude of the responses to the exoge-
nous shock in comparison to the case of nominal rigidity. On impact both the output gap
and inflation respond much stronger (0.036% and 0.078% vs. 0.02% each in the corre-
sponding basic model scenario). The effect is even more pronounced in the standard devi-
ations as they are a measure of volatility (0.08% and 0.11% vs. 0.02% and 0.03%). The
interest rate peaks three times as high (at 0.027%) with a std. dev. which is twice as big
(0.06%). Still, for policy analyses the order of magnitude of these differences is quantita-
tively negligible.

VI. Conclusion
Bernanke et al. (1997) and Barsky and Kilian (2001) have presented evidence point-
ing at the impact of systematic monetary policy on U.S. business cycle fluctuations follow-
ing on the heels of massive oil price increases. Being motivated by these findings and
arguing for a structural break in the way monetary policy was conducted pre- and post-
1979, CGG estimate interest rate rules for these different sample periods and evaluate their
performance in the face of oil price shocks in a standard and ad hoc version of a New
Keynesian DSGE model. Giving their model an explicit micro-foundation this paper has
added precision to the transmission channel of oil price shocks to investigate whether this
structure is appropriate to reproduce CGG’s results both qualitatively and quantitatively.
Accordingly, I have modelled energy as an input factor in production to capture the effects
of oil price spikes and assessed the performance of empirically estimated policy rules in
response to an energy price shock, holding constant the volatility of the shock.

My main conclusion is that while I can confirm the findings of Bernanke et al. (1997),
Barsky and Kilian (2001) and CGG qualitatively by an intra-model comparison, my results
are quantitatively implausible when compared to actual data. Thus, I showed that a mon-
etary regime which reacts aggressively to exogenous oil price shocks by raising the ex ante
real interest rate exerts a dampening impact on the amplitude of the responses of both
inflation and the output gap to the shock. At the same time, inflation persistence seems to
be perpetuated by weak monetary policy while a stronger stance achieves a much faster
reduction. These results are robust to the inclusion of real rigidity in the labour market in
addition to nominal rigidity in the final goods market, although the second model exhibits
more volatility and monetary policy is slightly less effective in muting the consequences
of oil price shocks. Generally, my insights seem to support the view that the apparent
change in the oil price-macroeconomy relationship may in fact be attributable to a change
in monetary policy.

By way of contrast, the quantitative effects during actual recessions were at times more
than a hundred-fold stronger than predicted by the model even in the presence of an
extreme degree of real rigidity (thus, the strongest response of inflation, for instance, was
0.128% against values of over 10% measured during the 1970s). The differences in the
performance of alternative policy rules in economies characterized by stickiness in prices
or real wages are therefore insignificant and would even decline in the degree of nominal
and real flexibility. This result allows for all but one policy recommendation – if the trans-
mission channel of oil price shocks proposed here is actually dominant and if the oil share in production is low, oil price shocks do not seem to have very serious consequences on economic stability so that they should perhaps not be the primary concern of monetary authorities.

However, this conclusion hinges on a number of specific conceptual and structural assumptions about the economy and the work-through of energy shocks. Hence, my model only provides a baseline case for the analysis of the oil price-macroeconomy relationship and ought to be extended along several lines. Assuming a Cobb-Douglas functional form for my production technology is a somewhat arbitrary choice, for instance. Other specifications of the production function exhibiting a smaller elasticity of substitution between energy and labour might aggravate the consequences of energy price shocks. One way to modify the production technology is via capital utilization being costly in terms of energy as proposed by Finn. Second, to generate a stronger response to energy shocks demand-side effects should be included as an additional transmission mechanism. This is most conveniently done by modelling an open economy where energy is imported both for production and final consumption as in de Fiore et al. 2006 (thereby leaving CGG’s framework as they focus on a closed economy). Moreover, in an open economy setting where aggregate consumption includes energy, shocks to this commodity may drive a wedge between different measures of inflation so that the choice of the appropriate target inflation indicator becomes a highly important issue for monetary policy (cf. Kamps and Pierdzioch 2002). Finally, in more recent work Hamilton (2003) argues that the responses of economic aggregates to oil price increases and decreases are essentially asymmetric. These effects cannot be captured by models taking a linear approach to the relation between oil and production.

With these issues being unaddressed it is evident that the present analysis is not the final note to the debate over the oil price-macroeconomy relationship and the policy relevance of oil price shocks. With increasing crude oil scarcity and limited substitutability the impacts of oil price surges on economic stability are rather likely to be a much-debated and probably growing concern in the future as well.
References


FITTING MARTINGALE WITH HORSESHOES FOR A DETERMINISTIC WALK TOWARDS CHAOS:
A MODEL OF HETEROGENEOUS EXPECTATIONS AMONG TRADERS

Erik Eggum
University of Warwick

ABSTRACT

This paper argues the importance and investigates the implications of nonlinear approaches to economics and introduces mathematical concepts associated with chaos theory that may not be part of an economist’s standard toolbox. Furthermore, we apply these concepts to study the implications of endogenously formed heterogeneous expectations in a simple nonlinear asset pricing model using an extended version of the adaptive rational equilibrium dynamics (ARED) introduced by Brock and Hommes (1997, 1998). In this model agents make a bounded rational choice between a finite set of predictors in forming their beliefs about future prices. It is shown that chaotic price fluctuations emerge as the intensity of choice to switch prediction strategies becomes high, giving rise to strange attractors and possibly horseshoes.

I. Introduction

“You believe in the God who plays dice, and I in complete law and order.”
Albert Einstein, Letter to Max Born

The efficient markets hypothesis (EMH) based on rational expectations stipulates that because prices are determined by fundamentals and agents have homogeneous expectations, price volatility is a result of random exogenous factors only. Although this paradigm is not a bad approximation of reality, it leaves much up to chance. The aim of this paper is, therefore, to challenge the underlying rational expectations assumptions of the EMH. Since information is costly both to acquire and process, markets should exhibit informational asymmetry. Indeed, even granting symmetric information, agents would arguably still interpret information differently. This is what the current paradigm ignores. Allowing for heterogeneous expectations may therefore enhance our understanding of the conse-
quences of agents’ interactions in the market.

The observation that both the largest price and volatility changes are unrelated to real-world events, and that trading volume and volatility of returns are large and auto-correlated (Gaunersdorfer 2000) indicates an endogenous dynamics possibly due to market psychology and “animal spirits.” Such a mechanism can be captured by heterogeneous expectations which may amplify the impact of “news” and in turn effectively generate “news.” Simple nonlinear deterministic models (Day and Chen 1993) can generate complex (chaotic) behavior similar to a random walk, suggesting they may enhance our understanding of economic phenomena. Furthermore, price movements exhibit dynamics strikingly similar to those of fluid turbulence (Farmer et al 2005), suggesting that economic theory may have much to gain from physics.

To study the dynamics of heterogeneous expectations, I use a framework introduced by Brock and Hommes (1997), henceforth BH. They examine Adaptive Belief Systems where agents use a finite set of predictors that are functions of past information to form expectations. Each predictor has a publicly available performance measure based on the predictor’s past realized profits. The agents’ rational choice between predictors yields the Adaptive Rational Equilibrium Dynamics (ARED). BH (1997, 1998) and subsequently Gaunersdorfer (2000, 2001) apply ARED to a simple discounted present value asset-pricing model with one costly sophisticated predictor \( \theta_1 \) and one (or several) ‘naïve’ predictor(s) \( \theta_2 \). When the (absolute) difference between the observed price \( p_t \) and the fundamental price \( p^* \) is small, most traders rationally use the costless \( \theta_2 \), thus increasing \( |p_t - p^*| \). As the forecast error of \( \theta_2 \) increases, an increasing number of traders find it profitable to switch to \( \theta_1 \) eventually causing \( p_t \) to approach \( p^* \) again. As the forecast error of \( \theta_2 \) decreases, the story repeats. The interaction of this centrifugal force as \( |p_t - p^*| \) is small and the centripetal force as \( |p_t - p^*| \) becomes large may lead to market instability and the emergence of strange attractors and chaotic price volatility.

This paper extends BH’s model by using two predictors with 1) risk-adjusted performance measures and 2) differentiated risk aversion for different traders. Using a mixture of analytical and numerical methods, I study the dynamics of the model and offer an economic interpretation of the results. Some short remarks conclude the paper and provide a brief discussion of possible future work.

II. The Model

The EMH cannot always hold. The definition of competitive equilibrium as a situation where all arbitrage profits are eliminated is inconsistent with the assumption of rationality when arbitrage is costly (Grossman and Stiglitz 1980). Hence, given that agents are rational, profitable arbitrage must occasionally be possible. This in turn suggests that agents have heterogeneous beliefs. To study the possible implications of expectations heterogeneity, I follow BH (1998), and present a simple discounted present value asset-pric-
ing model with a constant number of $N$ traders facing one risky and one risk-free asset.

**Theoretical Framework**

We assume the risk-free asset is supplied perfectly elastically at gross return $R > 1$ and let

$$p_t = \text{Price (ex-dividend) per share of the risky asset at time } t;$$

$$\{y_t\} = \text{The stochastic process of dividends of the risky asset, assumed to be iid;}$$

$$z_t = \text{The number of shares of the risky asset purchased at time } t.$$

The traders’ wealth in $t+1$ is given by

$$W_{t+1} = RW_t + R_{t+1}z_t,$$

$$=RW_t + (p_{t+1} + y_{t+1} - Rp_t)z_t \quad \text{[Dynamics of wealth]} \quad (2.1)$$

where bold font indicates random variables. Furthermore, I denote the conditional expectation and variance operators respectively:

$$E_{ht} = \text{Investor } h's 'beliefs' about the conditional expectation operator } E_t.$$  

$$V_{ht} = \text{Investor } h's 'beliefs' about the conditional variance operator } V_t.$$  

$E_{ht}$ and $V_{ht}$ are based on a publicly available information set of past prices and dividends, $\Omega_t = \{p_{t-1}, y_{t-1}, \ldots\}$. Note that $z_t$ in equation (2.1) is multiplied by the excess returns per share $R_{t+1} = p_{t+1} + y_{t+1} - R p_t$. This implies that the conditional variance of wealth $V_t(W_{t+1})$ is given by $z_t^2$ times the conditional variance of excess returns per share, $V_t(R_{t+1}) = V_t(p_{t+1} + y_{t+1}).$

Following Chiarella and He (2002), I assume that risk aversion $a_h$ differs among different types of traders. The literature suggests that smart-money investors have a higher risk aversion than noise traders, and that such discrepancies in risk aversion influence the price dynamics. Hence, type $h$’s demand for shares $z_{ht}$ solves

$$\text{Max } \Phi = \{E_{ht} W_{t+1} - \left(\frac{a_h}{2}\right)V_{ht}(W_{t+1})\} \quad (2.2)$$

with F.O.C.

$$\frac{d \Phi}{dz_{ht}} = E_{ht} (p_{t+1} + y_{t+1} - Rp_t) - a_h V_{ht}(R_{t+1}) z_{ht} = 0$$

Hence, in equilibrium type $h$’s demand for the risky asset is given by

$$z_{ht}^* = \frac{E_{ht} (p_{t+1} + y_{t+1} - Rp_t)}{a_h V_{ht}(p_{t+1} + y_{t+1} - Rp_t)} = \frac{E_{ht}(R_{t+1})}{a_h V_{ht}(R_{t+1})} \quad \text{[Type h's demand for the risky asset]} \quad (2.3a)$$
For simplicity I assume that beliefs about the conditional variance of excess returns \( V_{ht}(R_{t+1}) \) are constant and the same for all types \( h \). We can thus write \( V_{ht}(R_{t+1}) = \sigma^2 \) so that equation (2.3a) becomes

\[
\tilde{z}_{ht}^* = \frac{E_{ht}(R_{t+1})}{\sigma^2} \quad \text{[Type } h \text{'s demand for the risky asset]} \quad (2.3b)
\]

To establish an equilibrium condition, let

- \( z_{st} = \) Supply of shares per investor;
- \( n_{ht} = \) Fraction of traders of type \( h \).

Hence equilibrium at time \( t \) implies

\[
\sum_{h} n_{ht} \left[ \frac{E_{ht}(R_{t+1})}{a_{ht}V_{ht}(R_{t+1})} \right] = z_{st} \quad \text{[Equilibrium of supply and demand of shares]} \quad (2.4a)
\]

alternatively,

\[
\sum_{h} n_{ht} z_{ht} = z_{st} \quad \text{[Equilibrium of supply and demand of shares]} \quad (2.4b)
\]

With homogeneous expectations \( \Sigma n_{ht} = 1 \) and, using a Walrasian approach to aggregated demand, we can thus write the equilibrium condition (2.4a) as

\[
\frac{E_{ht}(p_{t+1} + y_{t+1} - Rp_t)}{a_{ht}V_{ht}(R_{t+1})} = z_{st} \quad \text{[Equilibrium condition with homogeneous expectations]} \quad (2.5)
\]

Equation (2.5) yields the pricing equation

\[
Rp_t = E_{ht}(p_{t+1} + y_{t+1}) - a_{ht}V_{ht}(R_{t+1})z_{st} \quad \text{[Pricing equation]} \quad (2.6)
\]

We can set \( z_{st} = 0 \) without any loss of generality (Gaunersdorfer 2000) to obtain

\[
Rp_t = E_{ht}(p_{t+1}) + \bar{y} \quad \text{(2.7a)}
\]

where \( \bar{y} = E_{t}(y_{t+1}) \) are constant since \( \{y\} \) is assumed iid. Rewriting equation (2.7a) as explicitly conditional upon the information set \( \Omega_t = \{p_t, p_{t-1}, \ldots, y_t, y_{t-1}, \ldots\} \)

\[
Rp_t = E_{ht}(p_{t+1} | \Omega_t) + \bar{y} \quad \text{(2.7b)}
\]

enables us to define the fundamental value of the risky asset as \( \bar{p}_t = \bar{p} \). This solution must satisfy \( \bar{R}\bar{p} = \bar{p} + \bar{y} \), for which \( \bar{p} = \bar{y} / (R - 1) \) is the unique solution to satisfy the ‘no-bubbles’ condition.
Before continuing building the model it will be convenient to work with deviations from the fundamental value \( \bar{P} \) in order to ensure tractability. Hence, I define \( x_t = p_t - \bar{p} \).

I denote rational expectations of excess returns
\[
\rho_t = E_t R_{t+1} = E_t x_{t+1} - Rx_t = x_{t+1} - Rx_t
\]
[Rational expectations of excess returns] (2.8a)

and type \( h \) conditional expectations
\[
\rho_{ht} = E_{ht} R_{t+1} = E_{ht} x_{t+1} - Rx_t = f_{ht} - x_{t+1} + \rho_t = f_{ht} - Rx_t
\]
[h’s Conditional expectations] (2.8b)

Type \( h \)'s maximization problem in (2.2) can then be written
\[
\max_{z} \Phi = \left\{ E_{ht} R_{t+1} z - \left( \frac{a_h}{2} \right) z^2 V_{ht}(R_{t+1}) \right\} = \left\{ \rho_{ht} z - \left( \frac{a_h}{2} \right) z^2 V_{ht}(R_{t+1}) \right\}
\]
(2.9)

With F.O.C.
\[
\frac{d\Phi}{dz} = \rho_{ht} - \frac{a_h}{2} V_{ht}(R_{t+1}) = 0
\]

Hence, analogous to (2.3), \( h \)'s equilibrium demand for the risky asset is
\[
z(\rho_{ht}) = \frac{\rho_{ht}}{\frac{a_h}{2} V_{ht}(R_{t+1})}
\]
[Type \( h \)'s equilibrium demand for the risky asset] (2.10)

We can now define type \( h \)'s risk-adjusted realized profits as
\[
\pi_{ht} = \pi(\rho_t, \rho_{ht}) = \rho_t z(\rho_{ht}) - \left( \frac{a_h}{2} \right) z(\rho_{ht})^2 V_{ht}(R_{t+1})
\]

\[
= \rho_t \frac{\rho_{ht}}{\frac{a_h}{2} V_{ht}(R_{t+1})} - \left( \frac{a_h}{2} \right) \left( \frac{\rho_{ht}}{\frac{a_h}{2} V_{ht}(R_{t+1})} \right)^2 V_{ht}(R_{t+1})
\]
(2.11)

where \( R_{t+1} \) is the \textit{realized excess return} over the period \( t \) to \( t+1 \), defined as:
\[
R_{t+1} = p_{t+1} + y_{t+1} - R \bar{p}_t = x_{t+1} - p_{t+1} + y_{t+1} - Rx_t - R \bar{p}_t
\]
\[
= x_{t+1} - Rx_t + p_{t+1}^* + y_{t+1} - E_t \left( p_{t+1} + y_{t+1} \right) + E_t \left( p_{t+1} + y_{t+1} \right) - R \bar{p}_t
\]
\[
\quad = x_{t+1} - Rx_t + \delta_{t+1}
\]
(2.12)

For simplicity, I shall set \( \alpha_{KB} = 1/\left( \sigma^2 + \sigma_o^2 \right) \)
\[
\pi_t = \frac{\rho_t^2}{2}
\]
(2.19b)
Given our assumption of constant variances, this specification seems to be a justified simplification since under rational expectations any prediction error is only due to random events \((\sigma^2 + \sigma^2_e)\) which here can be assumed as small as we like.

The fractions of different types of traders at the end of period \(t\), \(n_{ht}\) is determined by \(\pi_{h,t-1}\). We can subtract \(\pi_{t-1}\) from \(\pi_{h,t-1}\) without altering the discrete choice fractions \(n_{ht}\). This allows us to define the ‘performance measure’ \(U_{ht}\) used to update \(n_{ht}\) as the difference in risk-adjusted profits from type \(h\)’s beliefs and rational expectations beliefs:

\[
U_{ht} = \Delta \pi \left( \rho_{h-1}, \rho_{h,t-1} \right) = \pi_{h,t-1} - \pi_{t-1}
\]

\[
= -\frac{1}{2a_h (\sigma^2 + \sigma^2_e)} \left( \frac{(f_{h,t-1} - x_t)^2}{\text{Squared Prediction Error}} - \frac{(x_t - R_{h,t-1})^2}{\text{Squared Realized Excess Return}} \right)
\]

\[
= -\frac{1}{2a_h (\sigma^2 + \sigma^2_e)} \left( \frac{(x_t - R_{h,t-1})^2}{2} \right)
\]

(2.19)

Fundamentalists use a sophisticated and costly predictor

\[
\tilde{U}_{ht} = U_{ht} - C + \alpha x_t^2
\]

(2.20a)

where \(\alpha, C \geq 0\), whereas all other traders use the simple

\[
\tilde{U}_{ht} = U_{ht}.
\]

(2.20b)

We see that the utility (meaning realized profits) of using this sophisticated predictor is lowered by the cost \(C\). However, due to the term \(\alpha x_t^2\), as the market price \(p_t\) is driven further away from the fundamental price \(\bar{p} \left( |x_t| \right)\) its utility increases, meaning that a larger fraction \(n_{ht}\) choose the sophisticated predictor. Hence, \(\alpha\) works as an exogenous ‘stabilizing force,’ driving prices back to the fundamental as \(|x_t|\) becomes large. This mechanism allows the predictor-choice dynamics to take market conditions into account, as well as predictor performance.

The model’s dynamics in terms of equilibrium prices and fractions can be summarized by the following adaptive belief system:

\[
R_{x_t} = \frac{\sum_k n_{ht} \tilde{U}_{ht}}{\sum_k a_h n_{ht}} \tag{2.21}
\]

[Equilibrium condition in deviations form]

\[
n_{ht} = \exp \left[ \beta \tilde{U}_{h,t-1} \right] / Z_t \tag{2.22}
\]

[Fractions of type \(h\)]
We see that equilibrium prices in period $t$ are determined by the ‘performance measure’ $\hat{U}_{k_{i-1}}$. Before we can analyze the model, we need to specify the types of traders as given by $f_{ht}$. I will consider three types, where type 1 traders are fundamentalists and types 2 and 3 are trend chasers and contrarians, respectively. Without further due, I now introduce the traders:

- Type 1: Fundamentalist
  \[
  \begin{align*}
  \text{Mean:} & \quad f_{1t} = 0 \\
  \text{Variance:} & \quad \sigma^2 = 0 \\
  \text{Risk aversion coefficient:} & \quad a_1
  \end{align*}
  \]

- Type 2: Pure trend chaser
  \[
  \begin{align*}
  \text{Mean:} & \quad f_{2t} = g_2 \hat{x}_t + b_2 \quad [g_2 > 0] \\
  \text{Variance:} & \quad \sigma^2 = 0 \\
  \text{Risk aversion coefficient:} & \quad a_2
  \end{align*}
  \]

- Type 3: Contrarian
  \[
  \begin{align*}
  \text{Mean:} & \quad f_{3t} = g_3 \hat{x}_t + b_3 \quad [g_3 < 0] \\
  \text{Variance:} & \quad \sigma^2 = 0 \\
  \text{Risk aversion coefficient:} & \quad a_3
  \end{align*}
  \]

### Model Analysis

Considering the system with types 1 and 2 gives rise to the following system:

\[
\begin{align*}
\hat{x}_t &= \frac{g_2}{R} \frac{a_1 a_2 (\sigma^2 + \sigma^2_{\theta})}{d_1 n_{yt} + d_2 n_{2t}} - n_{2t} \hat{x}_t \\
\hat{n}_{yt} &= \exp \left[ \beta \left( -\frac{1}{2 a_1 (\sigma^2 + \sigma^2_{\theta})} \left[ \hat{x}_{t-1}^2 - (x_{t-1} - Rx_{t-2})^2 \right] - \frac{(x_{t-1} - Rx_{t-2})^2}{2} + \alpha x_{t-1}^2 - C \right) \right] / Z_t, \\
\hat{n}_{2t} &= \exp \left[ \beta \left( -\frac{1}{2 a_2 (\sigma^2 + \sigma^2_{\theta})} \left[ (g_2 \hat{x}_{t-2} - x_{t-1})^2 - (x_{t-1} - Rx_{t-2})^2 \right] - \frac{(x_{t-1} - Rx_{t-2})^2}{2} \right) \right] / Z_t.
\end{align*}
\]

(2.23a)  
(2.23b)  
(2.23c)

It is convenient to take the difference in fractions. Let therefore

\[
\begin{align*}
m_t &= n_{yt} - n_{2t} = 1 - 2n_{2t} \\
\Rightarrow n_{2t} &= \frac{1 + m_t}{2} \quad \text{and} \quad n_{yt} = \frac{1 - m_t}{2}
\end{align*}
\]

By defining $a = a_2/a_1$ we can then express the system (2.23) as
For reasons of space I here restrict attention to beliefs functions with one lag, \( L = 1 \) and only consider fundamentalists vs. trend chasers. Furthermore, given constant vari-
ances, I simplify the notation by letting \( \sigma^2 \) denote \( (\sigma^2 + \sigma_e^2) \).

Case 1 – Fundamentalists vs. Pure Trend Chasers: \( L = 1 \)

From (2.16) we now have \( \hat{x}_t = x_{t-1} \). Note that when \( L = 1 \), \( f_{x_{t-1}} = g_2 \hat{x}_{t-1} = g_2 x_{t-1} \). This implies the following system:

\[
x_t = \frac{g_2}{R} \frac{\alpha a^2 (1 - m_t)}{a + 1 + (a - 1) m_t} x_{t-1} \\
m_t = \tanh \left[ -\frac{\beta}{4 a_1 \sigma^2} \left( \frac{R x_{t-2} - x_{t-1}}{2} - \frac{(g_2 \hat{x}_{t-2} - x_{t-1})^2}{a} \right) \right] \\
\left( \frac{\beta a x_{t-1}^2 - \beta C}{2} \right) 
\]

(2.24)

whose properties can be summarized in our first lemma.

**LEMMA 2.1:** Existence and stability of steady states of (2.25)

Let \( m^* = \tanh \left( -\frac{\beta C}{2} \right) \), \( m^* = 1 - \frac{2 a R}{g_2 a_2 \sigma^2 + R (a - 1)} \), and \( x^* \) be the positive solution (if it exists) of

\[
tanh \left[ -\frac{\beta}{4 a_1 \sigma^2} \left( R (2 - R) - \frac{(g_2 - 2) g_2 + (2 - R) R}{a} \right) \left( x^* \right)^2 + \frac{\beta a (x^*)^2 - \beta C}{2} \right] = m^*. 
\]

(2.26)

a) For \( 0 < g_2 (a_2 \sigma^2) < R \), \( E_1 = (0, m^*) \) is the unique, globally stable steady state (fundamental steady state) of (2.25);

b) For \( R < g_2 (a_2 \sigma^2) < R (a + 1) \), there are two possibilities:

- If \( m^* < m^* \), then \( E_1 \) is the unique, globally stable steady state of (2.25);
If $m^* > m^\#$, then (2.25) has three steady states $E_1$, $E_2$, and $E_3$, where $E_1$ is unstable.

c) For $g_2(a_\sigma^2) > R(a+1)$, (2.25) has three steady states $E_1, E_2(x^*, m^*)$, and $E_2(-x^*, m^*)$, where $E_1$ is unstable.

Note that when trend chasers extrapolate only weakly $(0 < g_1(a_\sigma^2) < R)$, $E_1$'s global stability is independent of relative risk attitudes $a$. When there is strong risk-adjusted extrapolation $(R < g_2(a_\sigma^2) < (a+1)R)$, however, the stability of $E_1$ depends on $a$.

Finally, when risk-adjusted extrapolation becomes very strong $(g_3(a_\sigma^2) > (a+1)R)$, $E_1$ becomes unstable and bifurcates two additional steady states $E_2$ and $E_3$. Focusing on regime b), $R < g_2(a_\sigma^2) < (a+1)R$, we now study how changes in $a$ causes bifurcations. Let $a^*$ satisfy

$$m^\# = \tanh \left( \frac{-\beta C}{2} \right) = 1 - \frac{2a^* R}{g_2a_\sigma^2 + R(a^*-1)} = m^*$$

Then $m^* < m^\#$ iff $a > a^*$.

**Lemma 2.2**: Pitchfork Bifurcation at $a = a^*$.

Assume $R < g_2(a_\sigma^2) < (a+1)R$, and let $a^*$ be the solution of (2.27). Then, for $a > a^*$, $E_1$ is the unique equilibrium; for $0 < a < a^*$, (2.25) has three equilibria $E_1, E_2$, and $E_3$. Therefore (2.25) has a pitchfork bifurcation at $a = a^*$.

**Lemma 2.3**: Second Bifurcation at $a = a^{**}$.

Let $E_1(x^*, m^*)$ and $E_2(-x^*, m^*)$ be the non-fundamental steady states in Lemma 2.1. Assume $R < g_2(a_\sigma^2) < (a+1)R$ and $C > 0$, and let $a^*$ be the pitchfork bifurcation value in Lemma 2.2. There exists $a^{**} < a^*$ such that $E_2$ and $E_3$ are stable for $a \in (a^{**}, a^*)$ and unstable for $a < a^{**}$. For $a = a^*$, $E_2$ and $E_3$ exhibit a Hopf bifurcation.

BH obtained similar dynamics when studying changes in the intensity of choice between predictors, $\beta$. Focusing on the case $R < g_2(a_\sigma^2) < (a+1)R$, we now study the behavior of our system for various $\beta$.

**Lemma 2.4**: Stability and bifurcations w.r.t. $\beta$.

Let $R < g_2(a_\sigma^2) < (a+1)R$ and $C > 0$. There exists $0 < \beta < \beta^{**}$, such that

1. For $0 \leq \beta < \beta^*$, $E_1$ is globally stable;
2. At $\beta = \beta^*$, a pitchfork bifurcation occurs, creating the two non-fundamental steady states $E_2 = (x^*, m^*)$ and $E_3 = (-x^*, m^*)$. Based on Gaunersdorfer’s analysis I assume:
   a) For $\alpha < \alpha^*$ the pitchfork bifurcation is subcritical, i.e. there exist two unstable non-fundamental steady states for $\beta < \beta^*$.
   b) For $\alpha > \alpha^*$ the pitchfork bifurcation is supercritical, i.e. there exist two stable non-fundamental steady states for $\beta^* < \beta < \bar{\beta}$.

3. For $\beta^* < \beta < \beta^{**}$, $E_1$ is unstable, while $E_2$ and $E_3$ are stable;
4. At $\beta = \beta^{**}$, a Hopf-bifurcation of $E_2$ and $E_3$ occurs;
5. For $\beta > \beta^{**}$, $E_1$, $E_2$ and $E_3$ are all unstable.

From system (2.23) we see that if $C > 0$, the fraction of fundamentalists decreases to zero as $\beta \rightarrow +\infty$. This is rational since there is no point in paying any cost for a strategy that yields no extra profits. Figures 2.1-2.3 shows the resulting dynamics of setting $R = 1.1$, $\varepsilon_2 = 1.2$, $\alpha = 0.5$, $C = 1$, $\sigma^2 = 1$, $a_1 = 1$, $a_2 = a = 1.1$, with initial values $(1.2, 0.7, -0.2)$, and letting $\beta$ vary.

**Figure 2.1: Trend Chasers vs. Fundamentalists:**

The top diagram shows the time series for $\beta = 10$. The two bottom diagrams show the corresponding attractor in the $x(t)$ vs. $m(t)$ and $x(t)$ vs. $x(t-1)$ planes.
FIGURE 2.2: Trend chasers vs. Fundamentalists:

The top diagram shows the time series for $\beta = 15$. The two bottom diagrams show the corresponding strange attractors in the $x(t)$ vs. $m(t)$ and $x(t)$ vs. $x(t-1)$.

FIGURE 2.3: Trend chasers vs. Fundamentalists:

The top diagram shows the time series for $\beta = 20$. The two bottom diagrams show the corresponding strange attractors in the $x(t)$ vs. $m(t)$ and $x(t)$ vs. $x(t-1)$.
Fig. 2.1 shows that the system is periodic for $\beta = 10$. As $\beta$ increases the system becomes chaotic and develops strange attractors, as shown in fig. 2.2 for $\beta = 15$. For $\beta = 20$ the time series in fig. 2.3 indicates that there are repeated volatile bubbles followed by fundamental prices levels. BH (1998) found similar dynamics and concludes that for $\beta$ large, the system must be close to having a homoclinic point. This implies the existence of many horseshoes, which leads to topological chaos (Gaunersdorfer 2000). For $\beta > 65$ the system converges to zero, which makes economic sense since there is no point in incurring a cost when everyone switches “infinitely” quickly between predictors.

III. Conclusion

This paper emphasizes the significance of expectations heterogeneity in explaining price volatility, as a complementary deterministic contribution to the prevalent stochastic framework. GARCH models, for example, replicate real data well, but provide no economic explanation of volatility. I have shown that by modeling returns based on the traditional martingale component and a deterministic part we may gain important insight into the underlying causes of price volatility.

By drawing on three models I have created an enhanced model that incorporates differentiated risk aversion among traders and in which predictor choice is determined by both risk-adjusted predictor performance and market conditions. Although the model does not generate GARCH effects, we observed volatility clustering. The asset-pricing model framework used above allows for several modifications. An important extension of my model would be the addition of several different types of traders, ideally with low levels of noise added to their beliefs functions to simulate random influences on traders’ beliefs. Furthermore, the magnitude of the stabilizing force should ideally be endogenously determined. The above gives a taste of the potential of multidisciplinary approaches to economics. What recent advances in nonlinear dynamics show is that order and chaos are not diametrically opposed. The former is hidden in the latter. The same goes for economics. Fundamentals and heterogeneous expectations coexist. After all, traders do not play dice.
References


APPENDIX A: Sixth Annual Carroll Round Presentation Schedule

Session 1A Chair: Philip J. Cross (Assistant Professor, Georgetown University)
Andrew O’Brien-Penney (Georgetown University)
What is the Effect of the Abundance of Natural Resources on Income Distribution?
Yana Morgulis (University of Chicago)
Impact of Innovative Micro-Credit Policy on Rural Credit Markets: the Case of Thailand’s Million Baht Village Fund
R. Priya Mathew (Washington University in St. Louis)
The Effect of Microfinance on Income Inequality: Evidence from Ghana

Session 1B Chair: Anders F. Olofsgard (Assistant Professor, Georgetown University)
Lucia Franzese (Georgetown University)
Joining the European Union: The Real Effect on Foreign Direct Investment
Heleri Rande (New York University)
Has Foreign Bank Entry Led to Sounder Banks in Transition Economies? The Estonian Experience
Jennifer Xi (Dartmouth College)
The Conflicting Effects of Fiscal and Political Decentralization on Corruption: A Cross Country Study

Session 1C Chair: Holger C. Wolf (Associate Professor, Georgetown University)
Yi Sun (Georgetown University)
Exchange Rate Policy and Economic Growth
Nedko Kyuchukov (Dartmouth College)
Foreign Money Flows Where Privatization Thrives

Freddy Tsai (University of British Columbia)
The Effect of the Sarbanes-Oxley Act on U.S. Mergers & Acquisition Activity

Session 2A Chair: Robert Cumby (Professor, Georgetown University)
Bennett Surajat (Carleton College)
A Study of Monetary Transmission Mechanism Convergence and Monetary Policy Effectiveness among Core EMU Members from 1991-2006
Erik Eggum (University of Warwick)
Fitting Martingale with Horseshoes for a Deterministic Walk toward Chaos
Sören Radde (Universität Bayreuth)
Oil Price Shocks and Monetary Policy Revisited - An Energy-Augmented New Keynesian DSGE Model for the Analysis of the Role of Systematic Monetary Policy in the Oil Price-Macroeconomy Relationship

Session 2B Chair: Robin A. King (Visiting Assistant Professor, Georgetown University)
Mohammad Huq (Georgetown University)
Heterodox Stabilization in Brazil and Argentina in the 1980’s
Angelica da Rocha (University of Warwick)
The impact of Stock Market Liquidity on Economic Growth: an Empirical Analysis
Tanja Groth (University of St. Andrews)
Indirect Impact of Obesity on OECD Economic Growth

Session 2C Chair: Rodney Ludema (Associate Professor, Georgetown University)
Allison Phillips (Georgetown University)
Fruitless Endeavors: Agricultural Protectionism and North-South Preferential Trade Agreements
APPENDICES

Matthew Adler (Oberlin College)
The Effect of Preferential Trade on Multilateral Trade Liberalization: Canada and the North American Free Trade Agreement

Matthew Pech (Dartmouth College)
Physical, Informational, and Human Distance: A Comparison of the Determinants of Bilateral Portfolio and Foreign Direct Investments

Session 3A Chair: James Albrecht (Professor, Georgetown University)

Marion Aouad (Princeton University)
Not Quite Monopoly Money: Local Currency Systems- The Factors Leading to their Emergence and Sustainability

Abdulla Humaidan (University of Warwick)
Altruistic Giving and Ethnic Reciprocity: Applied to Returning Misplaced Mail

Elena Spatoulas (University of Michigan)
An Extension of Cartel Theory

Stephen Brinkmann (Georgetown University)
Determinants of Sovereign Credit Risk Premia in Commodity-Dependent Latin American Countries Discussant: Marion Aouad (Princeton University)

Session 3B Chair: Peter Norman (Visiting Associate Professor, University of North Carolina at Chapel Hill)

Cynthia Yim (Princeton University)
The Impact of Economic Growth on the Environment

Zachary Mahone (New York University)
Aid Flow Characteristics and Growth

Ashley Halpin (Dartmouth College)
The Determinants of Latin American Immigrants' Settlement Locations in the Southwestern United States

Jennifer Noh (Georgetown University)
Migration and FDI: Is There a Bilateral Complementary Relationship between Foreign Tertiary Student Enrollment in the U.S. and U.S. FDI Outflows?

Session 3C Chair: Susan Vroman (Professor, Georgetown University)

Adrienna Huffman (Washington University in St. Louis)
How do Formal and Informal Restrictions on Women Affect Economic Performance?

David Wolff (Dartmouth College)
The Determinants of International Terrorism: Its Perpetrators and Their Targets

Nicholas Hartman (Georgetown University)
Intrahousehold Bargaining and Divorce Norms in Indonesia
APPENDICES

APPENDIX B: PAST SPEAKERS

First Annual Carroll Round (April 5-7, 2002)
Roger W. Ferguson, Federal Reserve Board of Governors
Donald L. Kohn, Federal Reserve Board of Governors
Lawrence B. Lindsey, Assistant to the President and National Economic Council
Edwin M. Truman, Institute for International Economics
John Williamson, Institute for International Economics

Second Annual Carroll Round (April 11-13, 2003)
R. Glenn Hubbard, Council of Economic Advisers and Columbia University
Donald L. Kohn, Federal Reserve Board of Governors
John Williamson, Institute for International Economics

Third Annual Carroll Round (April 15-18, 2004)
Donald L. Kohn, Federal Reserve Board of Governors
John F. Nash, Jr., Princeton University (1994 Nobel Laureate)
Peter R. Orszag, The Brookings Institute

Fourth Annual Carroll Round (April 22-24, 2005)
Ben S. Bernanke, Federal Reserve Board of Governors
William Easterly, New York University
Maurice Obstfeld, University of California at Berkeley
Edwin M. Truman, Institute for International Economics

Fifth Annual Carroll Round (April 28-30, 2006)
Kemal Dervis, United Nations Development Programme
Thomas C. Schelling, University of Maryland (2005 Nobel Laureate)

Sixth Annual Carroll Round (April 19-22, 2007)
Grant D. Aldonas, Center for Strategic and International Studies
François Bourguignon, Chief Economist and Senior Vice President of the World Bank
Randall Kroszner, Federal Reserve Board of Governors
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)  

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)

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)  
Daphney 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)  

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)

Third Annual Carroll Round  
(April 15-18, 2004)  
Meredith L. Gilbert, chair (COL ’04)  
Héber M. Delgado-Medrano (SFS ’06)  

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 C: *Former Carroll Round Steering Committees*

Sixth Annual Carroll Round
(April 19-22, 2007)
Stephen A. Brinkmann, chair (SFS ’07)
Lucia Franzese (SFS ’07)
Nicholas A. Hartman (SFS ’07)
Ian P. Hinsdale (COL ’09)
Alexander P. Kostura (SFS ’09)
Jennifer M. Noh (SFS ’07)
Amy M. Osekowsky (SFS ’07)
Allison E. Phillips (SFS ’07)
Sun Yi (SFS ’07)

APPENDIX D: *Members of the Advisory Panel*

Meredith L. Gilbert, Yale University
Christopher L. Griffin, Yale University
Andrew T. Hayashi, University of California at Berkeley
Mitch Kaneda, Georgetown University
Robert S. Katz, Acumen Fund
J. Brendan Mullen, The Advisory Board
Scott E. Pedowitz, Corporate Executive Board
Erica C. Yu, University College London
APPENDIX E: *Past Participants*

**FIRST ANNUAL CARROLL ROUND (APRIL 5-7, 2002)**

Azhar Adbul-Quader  
Santosh Anagol  
William Brady  
Daniel Braun  
Jacqueline Bueso  
Karla Campbell  
Benn Eifert  
Courtney Fretz  
Carlos Galvez  
Aniruddha Gopalakrishnan  
Christopher Griffen  
Casey Hanson  
Joshua Harris  
Andrew Hayashi  
Marco Hernandez  
Katia Hristova  
Maria Jelescu  
Fadi Kanaan  
Avinash Kaza  
Vinay Kumar  
Anisha Madan  
Kathryn Magee  
Ryan Michaels  
Jack Moore  
Brendan Mullen  
Andrei Muresianu  
Scott Orleck  
Scott Pedowitz  
Jonathan Prin  
Jeremy Sandford  
Deborah Slezak  
Conan Wong  

Columbia University  
Stanford University  
Georgetown University  
Oberlin College  
University of Pennsylvania  
University of Virginia  
Stanford University  
University of Pennsylvania  
Stanford University  
Duke University  
Georgetown University  
Lehigh University  
Georgetown University  
Georgetown University  
Massachusetts Institute of Technology  
Illinois-Wesleyan University  
Massachusetts Institute of Technology  
Yale University  
Stanford University  
Duke University  
Illinois-Wesleyan University  
Georgetown University  
Georgetown University  
Stanford University  
Georgetown University  
Brown University  
Duke University  
Georgetown University  
University of Pennsylvania  
Illinois-Wesleyan University  
Illinois-Wesleyan University  
Brown University
APPENDICES

APPENDIX E: Past Participants

Second Annual Carroll Round (April 11-13, 2003)

Nada Abdelnour, Georgetown University
Amanda Barnett, Emory University
Andrea Bell, Wellesley College
Patrick Byrne, University of Colorado
David Chao, Cornell University
Sylvia Ciesluk, Lehigh University
Adam Doverspike, Georgetown University
Benn Eifert, Stanford University
Adam Engberg, Georgetown University
Alexandra Fiorillo, Connecticut College
Eric Fischer, Georgetown University
Zlata Hajro, Wellesley College
Samina Jain, Georgetown University
Avinash Kaza, Stanford University
Eric Kim, George Washington University
Seth Kundrot, Georgetown University
Lada Kyi, Rice University
Lee Lockwood, Northwestern University
Sunil Mulani, New York University
Holly Presley, Vanderbilt University
Duncan Roberts, University of California at Berkeley
Lu Shi, Georgetown University
Shanaz Taber, Barnard College
Jiang Wei, University of Michigan
APPENDIX E: Past Participants

Third Annual Carroll Round (April 15-18, 2004)

Jeffrey Arnold Dartmouth College
Julia Berazneva Mt. Holyoke College
Mehmet Cangul Georgetown University
Richard Carew University of Virginia
Ashley Coleman Vanderbilt University
Dilyana Dimova Stanford University
Fernando Galeana Stanford University
M. Blair Garvey Emory University
Meredith Gilbert Georgetown University
Adam Greene Oberlin College
Asim Gunduz University of Virginia
Marc Hafstead Northwestern University
Andrew Hayashi University of California at Berkeley
Katherine Howitt McGill University
Sohini Kar Columbia University
Josh Lewis McGill University
Satish Lohani Illinois Wesleyan University
Alexis Manning Illinois Wesleyan University
Sara Menker Mt. Holyoke College
Elizabeth Mielke Vanderbilt University
Stratos Pahis Dartmouth College
Alicja Pluta Georgetown University
Adam Raymakers Dalhousie University
Caroline Schmutte Dartmouth College
Matt Sekerke Johns Hopkins University
John Soleanianov Columbia University
Kai Szakmary Columbia University
Brandon Wall Yale University
Kenneth Ward University of Chicago
Susan Work Georgetown University
### Fourth Annual Carroll Round (April 22-24, 2005)

<table>
<thead>
<tr>
<th>Name</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lidia Barabash</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Jasmina Beganovic</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Xun Bian</td>
<td>Illinois-Wesleyan University</td>
</tr>
<tr>
<td>Michael Fuchtgott</td>
<td>Columbia University</td>
</tr>
<tr>
<td>Michael Gechter</td>
<td>Pomona College</td>
</tr>
<tr>
<td>Kevin B. Goldstein</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Michael Haase</td>
<td>University of Copenhagen</td>
</tr>
<tr>
<td>Dennis Huggins</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Michael Insel</td>
<td>Claremont McKenna College</td>
</tr>
<tr>
<td>Jonathan Kirschner</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Shiyieng Lee</td>
<td>Duke University</td>
</tr>
<tr>
<td>James Liao</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Brian Lichter</td>
<td>Washington University</td>
</tr>
<tr>
<td>Wee Lee Loh</td>
<td>Cornell University</td>
</tr>
<tr>
<td>Alice Luo</td>
<td>Duke University</td>
</tr>
<tr>
<td>Katharine Mullock</td>
<td>University of Western Ontario</td>
</tr>
<tr>
<td>Jose Mustre del Rio</td>
<td>Ohio State University</td>
</tr>
<tr>
<td>Leah Nelson</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Ee Cheng Ong</td>
<td>Wellesley College</td>
</tr>
<tr>
<td>Matthew Phan</td>
<td>Columbia University</td>
</tr>
<tr>
<td>Nina Rendelstein</td>
<td>Washington University</td>
</tr>
<tr>
<td>David Rogier</td>
<td>Washington University</td>
</tr>
<tr>
<td>Ana Maria Romero</td>
<td>Illinois-Wesleyan University</td>
</tr>
<tr>
<td>Nathan Saperia</td>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Bogdan Tereshchenko</td>
<td>Georgetown University</td>
</tr>
<tr>
<td>Olga Timoshenko</td>
<td>University of Western Ontario</td>
</tr>
<tr>
<td>Tom Vogl</td>
<td>Princeton University</td>
</tr>
<tr>
<td>Kenneth Ward</td>
<td>University of Chicago</td>
</tr>
<tr>
<td>Jonathan Wolfson</td>
<td>Washington University</td>
</tr>
<tr>
<td>Suzanne Zurkiya</td>
<td>Emory University</td>
</tr>
</tbody>
</table>

266
APPENDIX E: Past Participants

Fifth Annual Carroll Round (April 28-30, 2006)

Sarah Carroll
Ruth Coffman
Dubravka Colic
Pratik Dattani
Jennifer Dawson
Héber Delgado-Medrano
Sherri Haas
Jen Hardy
Lauren Iacocca
Salifou Issoufou
Stella Klemperer
Daniel Kurland
Corinne Low
Shanthi Manian
Michael Monteleone
John Nesbitt
Natasha Nguyen
Oyebanke Oyeyinka
Evgeniya Petrova
Emy Reimao
Svetoslav Roussanov
Vikram Shankar
Juan Carlos Suarez
Austin Vedder
David Wiczer
Geoffrey Yu
Xiaoti Zhang

Stanford University
Georgetown University
Wellesley University
University of Warwick
Illinois Wesleyan University
Georgetown University
Illinois Wesleyan University
Georgetown University
University of California at Los Angeles
University of Wisconsin at Madison
Brown University
Dartmouth College
Duke University
Georgetown University
University of Chicago
Georgetown University
University of California at Berkeley
Carleton College
Dartmouth College
Georgetown University
Columbia University
Georgetown University
Trinity University
Dartmouth College
Carleton College
Carleton College
University of Warwick
APPENDIX E: Past Participants

Sixth Annual Carroll Round (April 19-22, 2007)

Matthew Adler
Marion Aouad
Stephen Brinkmann
Erik Eggum
Lucia Franzese
Tanja Groth
Ashley Halpin
Nicholas Hartman
Adrienna Huffman
Abdulla Humaidan
Mohammad Huq
Nedko Kyuchukov
Zachary Mahone
R. Priya Mathew
Yana Morgulis
Jennifer Noh
Andrew O’Brien Penney
Jessica Oliveri
Matthew Pech
Allison Phillips
Angelica da Rocha
Sören Radde
Heleri Rande
Elena Spatoulas
Yi Sun
Bennett Surajat
Freddy Tsai
David Wolff
Jennifer Xi
Cynthia Yim
Oberlin College
Princeton University
Georgetown University
University of Warwick
Georgetown University
University of St. Andrews
Darmouth College
Georgetown University
Washington University
University of Warwick
Georgetown University
Darmouth College
New York University
Washington University
University of Chicago
Georgetown University
Georgetown University
Monash University
Darmouth College
Georgetown University
University of Warwick
University of Bayreuth
New York University
University of Michigan
Georgetown University
Carleton College
University of British Columbia
Darmouth College
Darmouth College
Princeton University
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.