

New York State Economics Association

**Proceedings
of the
New York State
Economics Association**

Volume 4

**64th Annual Meeting
Rochester, New York
September 23-24, 2011**

Editor: Richard Vogel

New York State Economics Association
Founded 1948

2010-2011 Officers

President:

Jeffrey Wagner, Rochester Institute of Technology

Vice-President:

Richard Vogel, SUNY-Farmingdale

Secretary:

Michael McAvoy, SUNY-Oneonta

Treasurer:

David Ring, SUNY-Oneonta

Editor of the *New York Economic Review*:

William P. O'Dea, SUNY-Oneonta

Web Coordinator:

Wade L. Thomas, SUNY-Oneonta

Board of Directors:

Cynthia Bansak, Saint Lawrence University

Darius Conger, Ithaca College

Robert P. Culp, Dalton State College

Richard Deitz, Federal Reserve Bank of New York – Buffalo Branch

Arindam Mandal, Siena College

Patrick Meister, Ithaca College

Manimoy Paul, Siena College

F. Pan Shu, SUNY-Potsdam

Della Lee Sue, Marist College

Published, June 2012

Table of Contents

Revitalizing Economic Education	
<i>Anna Cummings and William T. Ganley</i>	3
The World is in Overshoot, Will We Let it Collapse? Erosion Loops in our Biosphere and the Role of Effective Policy as a Counteractive Measure	
<i>Ana P. Diaz</i>	10
Who Enters the Foreclosure Process?	
<i>Eric Doviak and Sean MacDonald</i>	20
Who Defaults on their Home Mortgage?	
<i>Eric Doviak and Sean MacDonald</i>	36
Do International Capital Flows Worsen Macroeconomic Volatility in Transition Economies?	
<i>Scott W. Hegerty</i>	62
Can Crowd Out Hamper Stimulus Programs in Recessions?	
<i>John J. Heim</i>	71
A Method for Separating Income and Substitution Effects of Exchange Rate Changes	
<i>John J. Heim</i>	80
Degrowth and the Social Structure of Accumulation	
<i>Kent Klitgaard</i>	88
Factors Determining Consumer Sentiment – Evident from Household Survey Data	
<i>Kajal Lahiri and Yongchen Zhao</i>	97
Students' Economic Ideology and Deficit Reduction Choices	
<i>Michael McAvoy and William P. O'Dea</i>	107
A Merger Effect on Different Airline Groups: Empirical Study on the Delta-Northwest Merger in 2008	
<i>Sora Park</i>	117
The Rate of Returns to Schooling: A Case Study of Urban China	
<i>Xu Zhang</i>	137

“Revitalizing Economic Education”

Anna Cummings* and William T. Ganley†

ABSTRACT

Course redesign can be utilized to attain a variety of goals in undergraduate education: cost savings, improved learning outcomes and a reduction of the failure and withdrawal rate. This paper is about a successful redesign of a basic economics course: The Economic System. It is a course for non-majors offered for general education credit in the Social Sciences at Buffalo State College. The Course Redesign was part of a broader collaboration between SUNY and the National Center for Academic Transformation. This paper will analyze how this transformation took place, what was attained and how it could be used as a model for other courses and programs. The authors will highlight the unique features of economic education for college freshmen.

INTRODUCTION

The survey of economics course for non-majors has often been the grab-bag course in Economics departments. Typically it is a service course, often a general education course and quite often the last course thought about in schedule planning. There are departments in the economics community that use this type of course as a marketing tool to recruit majors to the field of economics. Yet many – maybe most – do not use the course in this manner.

In fact, these courses in basic economics are often the only course in economics students will ever take. This makes them the core foundation of economic education and economic literacy. It is not unusual to assign a brand new faculty member, a part-time lecturer or someone hired as the semester starts to teach the course. At best, an experienced faculty member who is willing to teach the course is given the assignment on a regular basis. Economics departments usually spend a significant amount of time on what the principles and intermediate theories courses ought to cover. Much less time is usually devoted to what this basic Economics 101 or survey course should include as a core curriculum. This is the state of basic economic education in higher education in the U.S. in the 21st century.

BACKGROUND

The course redesign project was initiated through a partnership between the National Center for Academic Transformation and the State of New York. It was created in 2008. The basic concept was to facilitate new thinking about the learning process for undergraduate education, especially at the

*Teaching Assistant Coordinator, Economics 101

†Professor of Economics and Finance, Buffalo State College, Ganleywt@buffalostate.edu

introductory course level. All SUNY campuses were invited to participate in a dialogue to identify course and academic programs that could benefit through an NCAT guided evaluation of how to improve student learning outcomes in courses which had a record of low grades, high failure or high withdrawal rates. Simultaneously, the cost effectiveness of the delivery of courses was considered a second feature of course redesign. In other words, were there new instructional techniques that would improve student learning outcomes? Finally, the cost of course delivery would be considered in terms of how to use instructional resources more effectively in introductory courses.

The National Center for Academic Transformation is an internationally recognized nonprofit organization with over twenty years of experience in the use of information technology to transform the learning process in higher education (Graves and Twigg 2006, p. 2). Originally funded by a grant from the Pew Charitable Trusts, the Center initiated and consulted with 30 colleges and universities on course redesign between 1999 and 2000 (Twigg, 1999, pp. 3-4). Since the early programs, the Center has been engaged with hundreds of higher education institutions to redesign courses in order to improve student learning outcomes, lower the cost of instruction and more effectively engage students in the learning process. NCAT has provided a variety of redesign models, all of which have assessment models incorporated into the redesign planning and implementation. The NCAT redesign approach is focused on large enrollment introductory courses that experience high failure and withdrawal rates (Powers 2007, p. 2).

Each SUNY campus announced the NCAT-SUNY partnership and solicited the interest of academic programs and departments. At Buffalo State College, the Office of Academic Affairs identified a series of courses in a variety of departments that might benefit from the NCAT Redesign approach. What was proposed was the following: each campus should encourage course redesign proposals from eligible programs and departments. The goal was to identify significant course redesign proposals on various campuses with successful proposals awarded two year, \$40,000 grants.

Buffalo State College's Office of the Provost constructed a list of courses that could potentially benefit from the NCAT Course Redesign process. The appropriate academic deans held discussions with department chairs with courses on the Provost's list. The Department of Economics and Finance had a general education course, Economics 101: The Economic System, which was a survey of economics for non-majors with a high freshman enrollment. Econ 101 is not a Principles of Economics course, although economic terms, concepts and principles are taught as part of the course. After a department discussion about the potential of a redesign of this course, the Chairperson of the Department at the time, Ted Schmidt, indicated to the Dean of the School of Natural and Social Sciences that a proposal to redesign Econ 101 would be submitted.

THE REDESIGN PROPOSAL

Traditionally two or three sections of The Economic System were offered every semester, with 100-150 students per section. Three faculty members rotated teaching the course. They used similar formats,

the same text and agreed that the course should use economic history and the history of economic ideas to teach students fundamental concepts, economic literacy and an understanding of economic problems and policies, past and present. Some online study guides, quizzes and in-class games had been developed by the individual instructors.

After discussion with the administrator of the instructional technology used at the College, Meghan Pereira, the key elements of the redesign were identified. The decision was to propose a course that would follow the NCAT hybrid model with a substantial part of the course online, a once-a-week traditional lecture and the utilization of undergraduate teaching assistants. The proposal argued that the online components of the course ultimately would be guided and supervised by the ULAs (Undergraduate Learning Assistants). The online components would include chapter quizzes, practice exams and study guides. The ULAs would carry out in-class activities, such as a “Jeopardy” game and others that had been created by one of the regular instructors. The class instructor would use traditional classroom lectures in the once-a-week class lecture. In addition, an online text would be written.

The proposal stated that the first year of the implementation would be used to create additional on-line components, such as online forums, which would pose broad-based economic questions and issues. In that first year the ULAs would experiment with in-class activities. It was understood that by the time the full redesign was ready the course would be 50% online and 50% in-class, which is the definition of a hybrid course at Buffalo State College.

NEW COURSE SYSTEM: ECON 101

Economics 101 and its new course system is a precursor to what can be done with classes for the future. This new course system takes traditional in-class lectures and combines them with online activities. There are two key components to how this system works. The first is ANGEL, Buffalo State College’s online learning tool and bridge between teacher and student, and the second are the Undergraduate Learning Assistants. These two components combined are the driving force for Economics 101.

This class hybrid uses ANGEL as the medium for these online activities, which help facilitate student learning. On ANGEL the students have numerous weekly assignments. One such assignment is an online discussion forum where they answer topic questions and respond to other students regarding the current week’s material. On ANGEL there are also required weekly chapter quizzes and optional review quizzes before each exam to be used as a study aid. The quizzes can be taken as many times as needed. The Power Points from the in-class lectures are also posted on ANGEL for the students to review or print out as needed. In addition to the online activities, ANGEL is also used for communication. If they have any questions, the students can talk to other students in the class, the professor or an Undergraduate Learning Assistant through email. There is also an open forum that is monitored by the ULAs where students can post their questions. This is a great feature of ANGEL because it allows

students to see other questions already asked and their answers. This leads us to the second key component.

The second key component of this system is the use of knowledgeable Undergraduate Learning Assistants, otherwise known as ULAs. ULAs are students chosen from previous semesters that did well in the class. This is important, because they know the class material well and are better able to convey these ideas to the students. This is also a great way for the teacher to receive feedback regarding what activities were beneficial to the students while taking the class. This allows the instructor and ULAs to revise anything that was not beneficial. The ULAs go through a training process before the semester starts on using ANGEL, class procedures/policies and FERPA. After the semester begins, each ULA is assigned a small group of students to work with. It is then that the ULA's job is to monitor the online activities and the students' progress in their group. Other duties include running the in-class reviews and tutoring. They are a large part of why this system works.

ANGEL and the ULAs work in unison to make this new course system possible. In this hybrid course, the instructor has a better idea of how the students are performing. The traditional teaching style in a huge lecture hall with 300 students cannot only be overwhelming for the teacher but for the students as well. It is almost impossible to know every student by name. It is easier for students who are not doing well to fall behind and feel lost, and by the time the instructor notices, it is usually too late in the semester. In this system, each ULA is given a small group of students. This is important for freshmen especially, because they have someone they can go to for help with a shorter response time. In this system there is always someone available that can help, if not the teacher or the ULA, then ANGEL with its online study material.

In this new class system, ANGEL and the ULAs are the keys to success, along with the professor, of course. Economics 101 combines traditional class lectures with online activities to facilitate student learning to form a hybrid class. This has many positives, including more study aids for the students and quick response time from teaching assistants for any questions. Both of these components work in unison with the professor to ensure the class material is being conveyed properly. This course system can be used in many other settings and classrooms, for not only economics but other subjects as well. This is a first of its kind, but will be beneficial as a template for other professors in years to come.

ULAS: THEIR ROLE AND RELEVANCE

Undergraduate Learning Assistants are the key to the new course system. All the ULAs have previously taken Econ 101. This is important, so they know what is expected of students in the course. They also know the material and have a perspective on which activities were beneficial and which were moot. It is also the ULAs' job to use this knowledge to cease those activities that were not valuable to the students in the class. ULAs then create another activity to take its place or they revise it for the upcoming semester.

The ULA's main job is to help facilitate student learning through these online activities, but since this is a hybrid course they also do face-to-face learning. To make this a little more personal for the students, each teaching assistant is given a small group of students. If there are 200 students in the class and 10 teaching assistants, then each assistant is given 20 students to work with personally. The students then go to their TA for tutoring help during office hours, or for assistance with ANGEL so that they can complete their online assignments. In addition, it is the ULA's job to keep track of how their students are doing and to see who is and is not doing well on tests or completing their assignments. Also, the ULA determines who is doing extremely well and considers the possibility that these individuals could be TAs for the next semester.

There are other face time activities as well. In the beginning of the semester, the teaching assistants hold an ANGEL instruction class for the students. This is where they cover what is required for the class and how to use ANGEL to complete their assignments. The assistants go over how to complete discussion forums, from where to find them, when they are open, how to post an initial response answering the topic questions from that week and how to reply to other students. They are taught how to complete quizzes and how to communicate with their TA through the communication tab by sending an email. A large number of the students are freshmen and ANGEL is new to them, so this is a perfect time to show them how it is used in Econ 101.

Another important duty is the in-class test reviews before each examination, with the help of the turning point clicker system. We use it to take attendance in every lecture and it is key to these review sessions. The teaching assistants make up questions using the text and class lectures for the students to use as a review. The ULAs then run a report at the end of the session. They use this report to see how the students did and then put up the review on ANGEL for the students to use as a study guide. The ULAs also proctor every exam. These are all opportunities for students to meet the teaching assistants, which make them more approachable if the students do need extra help.

The ULAs have many responsibilities, including monitoring of online activities, monitoring their group of students and being available for extra help as well as face-time activities. The assistants are a big part of why this new course system works. Their relevance speaks for itself. Without them, 300 students can be overwhelming for both teacher and students. In this system, the ULAs, with the help of ANGEL, bring it all together by being a support system for the students in the class and for the teacher. They are there to catch problems the students might have before it is too late and to help the teacher convey the course material in a way that their peers would understand.

ECO 101 AS A MODEL

The Buffalo State Econ 101 Course has the potential to serve as a model for other relatively high enrollment introductory courses. Many instructors create written or online materials for their courses. These may range from written narratives, exercises, problem sets, quizzes, study guides or review materials. A lot of this material could be put online to construct the framework for a hybrid course.

Furthermore, in many introductory courses, such as principles of economics, publishers have produced similar course materials used with the adoption of one of their textbooks. In most cases these materials are used to create an ongoing framework for a course.

In recent years many publishers have developed the capacity to customize courses in order to blend materials created by instructors with publisher course materials. Everything from the preparation and grading of course examinations, grade books, access to economic data and information sources, online and YouTube videos, homework problems and game simulations already exist, from publishers. Instructors have the option to use the publishers' materials directly or blend them with their own materials.

The use of Undergraduate Learning Assistants is one of the critical components. Our experience in their recruitment, training and supervision has been an evolutionary one. The very first year, just getting a sufficient number of ULAs was a major challenge. It was not until the second year that a well-defined training program was created, with the role of the ULAs effectively designed with a specification of their roles. One aspect of the Buffalo State experience is useful for other programs: many students decide they would like to serve more than one semester as a ULA. Our Course Redesign Program has had students interested in serving for as many as three semesters. This gave us the opportunity to have the experienced ULAs assist in the training of new student teaching assistants. From the experienced ULAs, a lead ULA was selected to serve as coordinator of the entire ULA team. This allows the instructor the opportunity to communicate more frequently with the lead TA who reports back on the team's progress and problems.

CONCLUSIONS

The Course Redesign of the Economic System (Economics 101) attained its major goals: (1) increase the class size with an actual improvement in learning outcomes for students; (2) a greater engagement for the students; (3) instructional cost reductions and (4) long-term sustainability of course redesign for Econ 101.

Over the course of an academic year the number of sections of Econ 101 went from five to two. Overall enrollments for an academic year are at 550 to 580 students, with the learning outcomes maintained as measured by test scores. The online activities monitored and supervised by the ULAs require each student to participate in the class through quizzes, games and forums. The Economics and Finance Department has been able to offer additional graduate courses and upper division courses for undergraduate majors. The three faculty members who were originally three faculty teams to redesign the course will continue to offer the course in its present format. The Department of Economics and Finance will consider the need to find new faculty members who would maintain this approach. This will be done through the standard recruitment process in future years. In general, there is a strong faculty consensus that supports the Course Redesign Program.

There were some unanticipated results from the Course Redesign. The educational-learning experience of the students who serve as Undergraduate Learning Assistants has been outstanding. Most

ULAs want to continue beyond the first semester. There is also a true *esprit de corps* among the ULA Team, so that most socialize outside the classroom. These students are exposed to the learning process in a very unique manner, with substantial interaction with faculty and professional staff in Instructional Resources. Many of the ULAs have ultimately decided to minor or even major in Economics and Finance.

Other departments at Buffalo State College have begun to explore course redesign options. Noteworthy among them have been the Biology and Computer Information Systems Departments. This past Spring semester, the College and the National Center for Academic Transformation cosponsored a Course Redesign Conference for the Social Sciences. There were faculty, Instructional Resources staff and administrators from several dozen colleges and universities in the Middle Atlantic States, as well as many faculty and administrators from Buffalo State College. The results from the Buffalo State Course Redesign and a course redesign basic psychology course at Frostburg State College in Maryland were very successful. Among many results from the day-long conference was a sense that Buffalo State College, in a few short years, has become a leader in revitalizing economic education for the average student.

REFERENCES

- Graves, William and Carol A. Twigg. 2006. "The Future of Course Redesign and the National Center for Academic Transformation." *Innovate*. 2(3): 1-5.
- Powers, Elia. "Introductory Course Makeovers." *Inside Higher Education*.
- Twigg, Carol A. 1999. "Improving Learning and Reducing Costs: Redesigning Large-Enrollment Courses." Pew Learning and Technology Program: Troy, New York.

**The World is in Overshoot, Will We Let it Collapse?
Erosion Loops in our Biosphere
and the Role of Effective Policy as a Counteractive Measure**

Ana P. Diaz*

ABSTRACT

Based on the analysis of the book “Limits to Growth: The 30-Year Update” by Meadows et. al (2004) which suggests our world is heading towards overshoot and collapse, this paper looks at the role of several erosion loops (soil depletion, and social instability) and gathers data that support the overshoot and collapse claim, and the impacts of increasing fossil fuel prices on the economy. Also, based on previous international accords in order to ameliorate the impact of the human population and its consumption to the environment and their success/ failure, the study analyzes the possibility of relying on policy-making as an effective and timely method to counteract the effects of a growing population and a diminishing resource base.

INTRODUCTION

Every day, our biosphere degrades as it bears the toll of growing human population and its sustenance needs. Before, when relative scarcity was the lens through which the economics profession approached the allocation problem, we had found ways to breach the seemingly restricted amount of capital that was at our disposition by becoming more efficient in production. Hence, we incurred in specialization by trading with other nations, and technological developments. Despite the scarcity approach, the Earth has still been bountiful regardless of how “unlimited” we considered our wants to be. Since the advent of the three-field-crop rotating system, the industrial revolution, and the discovery of hydrocarbons, our capitalistic society has been able to spring forward and go from sustaining a population of roughly one billion people in 1800, to somewhat above seven billion. These advances considerably amplified our carrying capacity (otherwise known as the amount of resources available for a population to sustain itself) and a new society developed. We were able to become more materialistically driven than we had in previous agricultural settings, or during our hunter-gatherer times¹, thus creating the ethos of the *homo economicus*.

Climate change, hunger, social unrest, and other phenomena that are taking place in our world right now point to the fact that the Earth has been taken advantage of for its generosity. The Earth gave us a hand and we took her whole arm, and now we are seeing and feeling the consequences on a global

*Department of Economics and Management (BA program), class of 2012, Wells College, Aurora, NY 13026.
(correspondence via e-mail adiaz@wells.edu)

scale, through hunger, political conflict, and diseases, to name a few. Nature always prevails at least in the long run, for as a species we are still subject to the same survival tests as other living organisms. Anyone with an elementary background in college-level ecology would know, populations that exceed their carrying capacity sooner or later are brought under control. That we have managed to escape this fate for the longest time due to the increased (and cheap) energy brought about by hydrocarbons does not mean the situation will not catch up to us. It does mean, to our own disadvantage, that the fall that will bring us under our carrying capacity will be far greater than any the human race has ever experienced before. Since the 1970s, when the authors of *Limits to Growth* (Meadows et al, 2004) warned us of the consequences of our actions if business as usual continued, people have been working to change the way we behave towards our planet. However, publicity for this topic withered away, as the claims made by the authors in the 1980s and 1990s seemed ridiculous and contradictory with the recovering economy.

This time, nonetheless, we are experiencing the whiplash of our foolishness and hubris. Like the eye of the hurricane, those years only served to distract us away from the issues at hand and make us believe that it had all been a bad scare by some crazy tree-loving scientists, themselves unknowledgeable about the magic of the market. This trend has slowed down recently. Ecological concerns have gained much terrain in the political realm than that which they had 10 or 20 years ago, when the Kyoto Protocol was signed by countries in a (rather weak) effort to reduce carbon dioxide emissions. Despite the eye-opening data, skeptics of climate change are still among us (!), and most governments are still idle bystanders when it comes to taking action, for fear of letting their populations' expectations of growth and economic success be marred by the quest to save the earth we live in, and preserve the resources we use to ensure our sustenance. Governments around the world, and especially important actors such as the United States and China, have not led the world to a consensus in order to orderly cushion our collapse.

Emphasis should be put on cushioning our collapse and not straying us away from it. Lester Brown reminds us, in his book *World on the Edge*, previous complex societies collapsed after they had severely affected the environment that they lived in. He finds examples in the Mayans, the Romans and the Sumerians, who all suffered from an ecological crisis alongside the other individual factors that lead to their demise. This same phenomenon has started to show in our own society. Whether you chose to acknowledge this view as pessimistic or realistic, if we continue business as usual and don't stand up to the material changes that are happening around us by re-shaping our paradigm about economic activity, contemporary society's return to our original carrying capacity will be considered a major catastrophe.

In their book *Limits to Growth: The 30-Year Update*, Meadows et. al, (2004) pose four possible scenarios that show the relationship of the human population with its carrying capacity, the complex resource base the human population uses (described below). After discussing all four models, the authors boil it down to two realistic scenarios for the future of humanity and its resource base: the overshoot and oscillation scenario, and the overshoot and collapse scenario. The World3 Model runs point in the direction of an overshoot scenario, and it is my contention that we might be leaning towards a collapse rather than an oscillating relationship between population and carrying capacity.

In an exponentially growing economy like ours, its expansion creates stresses within the environment in which it is encapsulated. These stresses act out as part of negative feedback loops, trying to bring the economy back into alignment with the constraints of the system. They are signals that the economy is putting too much pressure on its resource base. When ignored, these stresses become the sure precursors of erosion loops. One of the stresses mentioned in Meadows et al (2004) is an increase in the energy investment that it takes to develop a new hectare of farmland. Although I am positive more stresses are taking place on the earth because of our abuses, contributing to more erosion loops, I have decided to center on this particular one, given that as Lester Brown has mentioned, the ability to procure sustenance is has been every civilization's weakest link.

The discussion given to this paper is carried out in the following order. First, I discuss the World3 Model and its relevant components. Following, I describe the impacts of falling soil yields on the financial sector, and the consequences on social stability. Third, I look at the pros and cons of policymaking as a strategy for a cushioned transit to a less energy-intensive economy. The last section is a conclusion.

THE WORLD3 MODEL, A ROLLER COASTER BETWEEN SKYROCKETING GROWTH, NATURAL RESOURCE PEAKS, AND EROSION LOOPS

The World3 model was created by the authors of *Limits to Growth* in an attempt to get to broad predictions about "how the expanding global population and material economy interact with and adapt to the earth's limited carrying capacity over the coming decades" (Meadows et al, 2004, p. 137). It is a mathematical model, and utilizes differential equations to show a simplified interaction between different sets of feedback loops of stocks, such as industrial capital, population, pollution, and cultivated land, to name a few. As all these feedback loops interact together², the model focuses on the time that it takes for events to take place, the delays in flows and the slow unfolding of physical processes (Meadows et al 2004, p. 133).³

THE FOUR SCENARIOS

To understand the claim that currently our economy is heading towards the overshoot and collapse scenario, I would advise the reader to refer to the summary of the scenarios presented by the authors of *Limits to Growth* (Meadows et al, 2004, p. 158). The authors, after running World3 several times, concluded that we are far beyond our excess capacity at this point, and that stresses are already being felt all over the globe. The scenario I am concerned with, overshoot and collapse, is characterized by population levels significantly exceeding their carrying capacity, to the point where the carrying capacity is permanently damaged. This leads to a permanent diminishing of the resource base available to the population. The unavailability of resources then leads to a quick and sharp decrease in population, eventually leveling off over time to find itself within its carrying capacity. Granted, this is true if the situation of the planet continues as it is, with humanity depleting resources at an increasing rate and no policy actions taken to shift our consumption and waste patterns. On the other hand, the overshoot and

oscillation scenario, as Meadows et al. (2004) state, “can occur only if the environment suffers insignificant damage during periods of overload or can repair itself quickly enough to recover fully during periods of underload”. Insignificant damage due to changes in the exploitation of the carrying capacity is closely tied with timely policy action, and these occurrences are but few, as will be discussed further on.

EROSION LOOPS—THE BLACK SHEEP WITHIN POSITIVE FEEDBACK LOOPS

In order to better elucidate how erosion loops work, a brief explanation of what feedback loops are, coupled with exponential growth is needed. Feedback loops are a method utilized by those who work with systems dynamics. They represent different processes and the stocks within them. Positive feedback loops are self-reinforcing loops that generate exponential growth or exponential decay. Negative feedback loops, on the other hand, are goal-seeking loops that reverse the direction of change or try to pull the system back into equilibrium (Meadows et al 2004, p. 143).

Exponential growth is the driving force behind feedback loops. The difference between exponential growth and linear growth is the amount of growth in a given period. For things that grow linearly, the amount changing every period is constant. On the other hand, things that grow exponentially will have an increasing amount each period that depends on the accumulation of the factor in question. Exponential growth frequently happens in populations. For example, the positive feedback loop for populations is birthrates, while the negative feedback loop is deaths per year (or life expectancy). The more people there are on the planet, the more births there are going to be. However, deaths per year regulate how many people there will be on the planet. Within the World3 Model, feedback loops for industrial capital, agriculture and pollution are considered as well (Meadows et al 2004, p. 144).

There is one clear reason as to why most of the thousands of runs of the World3 model return overshoot and collapse scenarios: the presence of erosion loops in the system. Erosion loops indicate stress on a system, for the positive feedback loops’ reinforcing effects are not being counteracted adequately by negative feedback loops. One of the erosion loops modeled in World3 is the intensive work of the land, which produces more food in the short term at the expense of the soil and its maintenance in the long-term (Meadows et al 2004, p. 165). Thus, the negative feedback loop that would keep population in check, the lack of available food by regular solar flow means, is overpowered by the positive feedback loop of increased food availability through industrial outputs for agricultural inputs, such as fertilizers. That leads to industrial output pollution, which in turn causes stress on the environment, coupled with the unavailability of the soil to recover itself properly for better crop yield. Eventually, as food supplied declines and demand increases, more fertilizer is added to the soil, reinforcing the process of the positive feedback loop.

The most successful element to aid the multiple negative feedback loops that try and bring humanity under its carrying capacity to cushion collapse would be effective policy-making, incentivized by concerned and active citizens (Heinberg, 2005, p. 243). If for each sector that is suffering from stresses that are creating erosion loops, governments worldwide could develop and implement policies to

incentivize the private sector and consumers into protecting our resources, we would stand a chance to regenerate the earth's carrying capacity⁴. This, topic, however, will be dealt with later in the paper.

THE LAND PRODUCTIVITY EROSION LOOP: THE "WEAK LINK"

Liebig's Law stresses that a reaction or the growth of a system is limited by its least abundant factor. The human population, as any other population on this planet, is circumscribed by the amount of food and sustenance that it can procure for itself. As I mentioned above, one of the erosion loops present in World3 has to do with land productivity and investment on soil maintenance. As Malthus would argue in his Theory of Rent, increased population growth rates involve exploiting land more intensively and extensively due to diminishing marginal returns on the land or less productivity of the soil (Hunt, 2002, pp. 95-96). This presupposes that more energy (in the form of fertilizers, pesticides, mechanical harvesting, pumps for irrigation, etc.) would have to be invested in the growing of food to keep feeding the population⁵. We can see this phenomenon happening already. An analysis of data provided by the United Nation's FAO (Food and Agricultural Organization) shows that world consumption of fertilizers increased greatly in the past four decades: From a little over 30 million tons consumed in 1961, to more than 140 million tons in 2002 (FAOSTAT, 2004). Thus, world agricultural yield has increased from over 13 thousand Hg/Ha (Hectograms per Hectare) in 1961 to a little over 35 thousand Hg/Ha in 2009, and the area harvested increased by over 50 million hectares in that same period. That stated, the increase in fertilizer use would indicate, as Malthus pointed out, that as populations increased, lesser quality soil is worked (represented by the additional 50 million hectares harvested). Similarly, world fertilizer use per capita increased from one-hundredth of a ton in 1961 to a little more over two-hundredths of a ton in 2009, and world agricultural yield per hectare per capita in the same period of time decreased from around 0.4E-06 to 0.5E-06 (FAOSTAT 2004, US CENSUS BUREAU 2011). As lesser quality soils are used, we resort to more fertilizers and more land, both of which experience diminishing marginal returns over time.

An important element to take into account here is that we have not yet reached, much less gone beyond, a tipping point where the amount of energy we put into growing things would make no difference in yield, at least for a while. Soil continues to be productive, albeit at a decreasing rate. However, what this does indicate is that *ceteris paribus*, soil productivity is bound to collapse. We will eventually run out of cheap energy to throw into the soil⁶, as EROI for oil extraction has been declining in the past several years, and soil productivity levels without its miracle nutrient source will rapidly fall to their natural rate (maybe even less), as will its process of regeneration. Fallback to pre-industrial revolution agricultural levels will occur as we run out of these additives coming from the cheap energy derived from the easily accessible portion of the fossil fuels that are now massively exploited. Absolute soil fertility decreases and decline in agricultural yield will lead to war, famine and disease as the human population continues to grow, dramatic decreases in human population, and thus, the overshoot collapse scenario described by Meadows et al will realize itself.

CHEAP ENERGY AND THE FINANCIAL CONSEQUENCES OF FALLING SOIL YIELDS

Cheap energy (or resources), including oil and water, for example, is what has allowed us to develop into such a plentiful society that has defied the scarcity approach advanced by Smith and others. Not only that, but we have been able over the centuries to substitute high energy return on investment fuels for low EROI fuels (e.g. oil and natural gas for coal, coal for wood, wind, animal and human muscle power, since the 1700s. That process has come to an end since the high energy return on investment (EROI) fuels has been depleted. Since the industrial revolution, we have not dealt with absolute scarcity, or in other words, limited means. Economic welfare has been built around economic growth, which has been provided to such a wonderful degree through the use of fossil fuels. The Earth has been generous and given us plenty, and our human genius has allowed us to enhance this bounty. It is in this way that as a species we were able to reach 7 billion people in the world, as of late 2011. Liebig's Law, once again, implies that mentioned before, the only thing that restricts a population's reproduction rate is its carrying capacity, but we increased our carrying capacity as we took control of fossil fuels.

Population is estimated to increase to around 9 billion by 2050 and we are not only to provide sustenance but also economic stability for the coming generations. However, with increasing oil prices derived from higher costs of extraction, and the slowdown of economic processes around the globe as shortages in oil become more persistent, this will not be possible. Increases in oil prices will not only cause a decrease in the annual yields of the soil as it becomes to buy industrial fertilizers, but it will also affect every single aspect of the economy, and thus, the financial system. As Richard Heinberg explains, "Our current financial system was designed during a period of consistent growth in available energy, with its designers operating under the assumption that continued economic growth was both inevitable and desirable," (2005, p. 188) and so the spike in costs of what has been the engine of this unlimited growth will indubitably bring about financial turmoil. Given the above, if what lies ahead of us is a decrease in economic activity as oil becomes scarcer and more expensive; there is very little chance, if any, that the financial system will survive an energy decline in the economy, despite whether it is gradual or rapid. Add to this equation the human element of social instability wrought by precarious economic conditions, and the subsequent unwillingness of investors to put their money where these investments could prove unsuccessful or their profits taken away from them by recently-turned nationalist governments, the panorama for economic stability is grim, especially for poorer nations.

SOCIAL INSTABILITY: THE EROSION LOOP NOT MODELED

The question of cheap energy as a means to sustaining great populations, like the ones we have in societies all over the world today brings us to another erosion loop that was not modeled in World3: social instability (Meadows et al, 2004, p. 166). To go back to the land productivity erosion loop mentioned before, we can connect the fact that people who are not well fed, and see their resource based depleted, resort to violent means to assure a share of the shrinking pie.

Proof of the interconnectedness between erosion loops lies in the recent revolutions taking place in the Middle East, where their people ousted several governments who had been in power for decades. Food prices for cereals and grains had been rising for the past couple years, and for the populations that find themselves highly dependent upon these sources for food, their subsistence was clearly jeopardized. A closer look at the producer price per ton of cereals in these countries provides us with concrete evidence for the revolts, even though we only have figures until 2008 and the revolution took place in early 2011⁷. In Tunisia, in only one year (2007-2008), the price to produce one ton of wheat rose by 31%, and the price to produce one ton of barley or triticale rose by 55.8% (FAOSTAT, 2011). Its 23 year-old government met its end January 14, 2011 under protests of corruption and unemployment by the people⁸. Unemployment (labor wise) presupposes that there are less or no means in the household to support it and its most basic need: food. Coupled with rising prices, and thus decreasing purchasing power, revolution was bound to take place sooner than later. Food production prices in Egypt suffered similar spikes: a ton of Barley produced in 2008 cost 74.2% more than in 2007, and the production of a ton of wheat cost 130% more (FAOSTAT, 2011). The fact that the Egyptian government suffered the same fate as the Tunisian only a month later comes as no surprise. Relating the earlier example of increased use of fertilizer and its rising cost due to higher oil prices with the soaring food prices in the Middle East (and the rest of the world), we see then how the biophysical erosion loop affects the social one.

This of course, is not to mention the challenge resource-plentiful but poorer countries will face from richer and more militarily capable countries such as the U.S. and China. China, with its burgeoning economy is already leasing land in countries in Africa to ensure both the yields of the land and mining of minerals. The U.S. has had a resource-dominance policy for years now, both indirectly (through corporate profits) and directly, by ensuring resources through military control and bases all over the world. When food scarcity becomes an even more acute problem, it is these poor nations' peoples which will take the hardest toll. Other, more capable nations will take a hold of their resources one way or another, unless a policy of global food security is undertaken by all parties in an effort to cushion the transition into a less energy intensive economy (Heinberg, 2005, p. 251).

CREATING A PATH TOWARDS THE SOLUTION: PRIVATE ENTERPRISE VS. POLICY-MAKING?

Usually, it is believed effective policy-making and implementation is really the only tool we possess to make change in our communities, countries and the world. While this is not entirely true, it does play a major role in assuring that there is a systematic and leveled approach to combat societal collapse, ensuring that we oscillates back into equilibrium without severely diminishing our carrying capacity. In order to bring about change in our laws in order to reduce our energy consumption, we are going to need more people involved in grassroots organizations that reach out to their politicians and demand that the situation is addressed as soon as possible. If politicians in the international community are coaxed by their active constituents to, instead of offering optimistic but unrealistic goals for the economy, reach an

agreement on how to manage the resource shortage we will face, we will have as a species, a better (and less violent) chance of adapting to the new situation.

In our efforts at global accords, there is both good news and bad news. The good news, is that as a global community, we have been able to reach agreements when the threat to the human population has been imminent. A perfect example is provided by Meadows et al (2004) where they describe the process by which CFCs, chemical compounds which were destroying the Earth's ozone layer, were banned. However, It took the international community several years before the problem of the depleting ozone layer was uncovered in 1974 to when they reached an agreement in 1987 in Montreal and countries started implementing policies. This one of a kind event saved the future of the ozone layer, for the Montreal agreement and its successors impeded CFC concentrations to more than 40 ppb in 2075 (Meadows et al, 2004, p. 195). Concentrations are now more stable and are predicted to stay that way for the coming century.

Unfortunately, the bad news is that this is not always the case, and policy delays are the biggest stumbling rock we face as we try to prevent overshoot and collapse. The latest climate change conferences, where tackling carbon dioxide emissions was the key point in efforts to fight global warming, have advanced very slowly. Since the signing of the Kyoto Protocol in 1992, little progress has been made, China was exempt from limitations and the U.S. never ratified the treaty, these two nations are the world's largest emitters of heat-trapping gases. The last meeting of the UN, which took place in Bangkok only has brought good intentions, but no real agreements⁹. The big difference, it seems, is that CFCs have a more benign substitute, HFCs (the bonding of hydrogen seems to have a smaller impact on ozone depletion) while as of this time there is no viable substitute for fossil fuels. Despite many years of greenwashing most diplomats know this, but refuse to acknowledge the severity of the situation for political reasons. No one wants to be the bringer of bad news, or tell people that the economy is going to shrink and that standards of living are going to have to radically change in order for us to not fall into complete chaos. Similarly, interests from corporations and the financial sector which thrive on relentless growth, and which would in turn be severely downsized (if not liquidated), clash with that of the citizenry and are pressed upon politicians that have strong ties to them due to financial contributions to their campaigns or because they are a part of their "constituency". Overall, painting an optimistic and unrealistic vision of the future and telling people to hope for it is far easier for those in power than to roll their sleeves up and actually start fomenting change before the inevitable future rolls around the corner.

CONCLUSION

The data presented currently support the overshoot and collapse scenario. Evidence is present all over the planet, as erosion loops just as the ones we described exacerbate each other as time goes by and no real change is being made. We are experiencing diminishing crop yields per hectare, producing more social unrest. Unfortunately, at this point policy-makers are in a catch-22, as they see the signs of overshoot affecting their countries, but are unable to switch away from cheap energy sources that keep

their populations happy and fed in the short term. The entrenched paradigm of the *homo economicus*, and what we have known to be telling characteristics of a successful economy (growth and material acquisition) will be difficult to strip away from a societies that have many advantages thanks to the technology fossil fuels have provided us. More_so, it will be difficult to tell developing countries that they are to stop that process of developing for growth, and that the goal of more material accumulation and a better standard of living enjoyed by the First world is not feasible for them as well if we wish to be alive in the next couple hundred years. This fact would create intense discontent among these countries, leading in the worst cases to intense international political turmoil and war over the remaining resources. Thus, our efforts since 1949 of creating a global community would most likely be either severely altered or completely destroyed.

The above prospect looks grim, no doubt, and one can somewhat sympathize with these world leaders and their fear of what taking action would ensue in the international community. Nonetheless, paying attention and acting on the ecological needs of our planet despite of the political, social and economic turmoil that those actions might bring is necessary. There is no easy way to go back within our carrying capacity, and most likely we will not find more cheap energy to sustain or overconsumption lifestyle. We need effective action to be taken by the international community, just as they did to tackle the problem of the ozone layer hole in the 1980's. Changes *can* be made, industries *can* shift to other practices and economies actually do survive, albeit in a different fashion. The faster we take action and compromise with ourselves and our neighbors to understand what a less intensive economy entails and how to live within it; the inevitable transition away from fossil fuels will be more bearable and less violent.

ENDNOTES

1. A discussion of how needs were met by hunter-gather societies, and their complete disregard for material possessions due to the hindrance they posed on mobility can be found in a great book by Marshall Sahlins (1972) *Stone Age Economics*.
2. A diagram of how the stocks and their feedback loops interact together can be found in pages 144-145 (Meadows et al 2004).
3. If the reader were interested in understanding the model's intricacies, I would highly recommend they read chapter 4 of *Limits to Growth*.
4. Meadows et al (2004) refer to the case of CFCs and how the international community was able to protect the ozone layer from suffering more degradation, after the hole in Australia was discovered and the effects of it were laid out to be extremely harmful for life on Earth as we know it.
5. Malthus wrote at a time where fertilizers were not in use, so he did not know how fertilizers could enhance natural soil productivity and curb diminishing marginal returns for a while. This fact, however, is inconsequential as nature also catches up with fertilizers, and there are also diminishing marginal returns to the use of fertilizers in the soil.

6. Fertilizer nutrients are in their majority derived from oil, whose price has fluctuated with the years, mostly due to politics, such as the OPEC's oil embargo in the 1970's. However, as the era of peak oil approaches, oil's price will progressively increase due to greater extraction and refining costs. Not only is oil a non-renewable resource, but the best quality oil is used up first, which leaves us diminishing marginal returns and the certainty that at some point it will not be profitable to employ it to enhance agricultural product.
7. There is indication that prices did not drop substantially from 2008-2011, given that rising global food prices are still of much concern for the UN (see "World Food Situation" at www.fao.org.) Moreover, the fact that revolution actually took place a year and a half later should be indication enough that change for the better did not happen.
8. Elaine Ganley and Bouazza Ben Bouazza. "Tunisians drive leader from power in mass uprising". 14 January 2011. The Boston Globe. <http://timelines.boston.com/2011/1/14/tunisian-government-ousted> (last accessed April 13, 2011)
9. United Nations Framework Convention on Climate Change. *UN Climate Change Chief urges countries to push ahead with their work in 2011 as countries agree agenda for 2011*. Bangkok, April 8, 2011. http://unfccc.int/files/press/press_releases_advisories/application/pdf/pr20110408_bkk_close.pdf (last accessed April 15th, 2011)

REFERENCES

- Brown, Lester R. 2011. "World on the Edge: How to Prevent Environmental and Economic Collapse." Earth Policy Institute.
- FAOSTAT. faostat.fao.org. (Last Accessed April 13, 2011)
- Ganley, Elaine and Ben Bouazza. "Tunisians drive leader from power in mass uprising." 14 January 2011. The Boston Globe <http://timelines.boston.com/2011/1/14/tunisian-government-ousted>
- Heinberg, Richard. "The Party's Over: Oil, War and the Fate of Industrial Societies." New Society Publishers, 2005.
- Hunt, E.K. 2002. "History of Economic Thought: A Critical Perspective." 2nd Edition. Armonk, New York: M.E. Sharpe
- Meadows, Donella, Randers Jorgen and Dennis Meadows. 2004. "Limits to Growth: The 30-Year Update." White River Junction, VT: Chelsea Green Publishing Co.
- United Nations Framework Convention on Climate Change. 2012 "UN Climate Change Chief urges countries to push ahead with their work in 2011 as countries agree agenda for 2011." Bangkok, April 8, 2011 <http://unfccc.int/>
- U.S. Census Bureau International Database www.census.gov (last accessed January 20, 2012)

Who Enters the Foreclosure Process?

Eric Doviak* Sean MacDonald†

ABSTRACT

Since February 2010, detailed information on every home mortgage default and foreclosure in New York State must be filed with the New York State Banking Department (NYSBD). The data enables us to identify the financial characteristics that make a defaulted borrower more (or less) likely to enter the foreclosure process.

Our analysis of the NYSD data suggests that borrowers in default who took larger loans are more likely to progress to foreclosure. It also suggests that reducing principal balances may reduce the foreclosure rate, but might have an adverse effect on the mortgage industry.

Given the frequent criticism of the Home Affordable Modification Program (HAMP), it is no surprise that defaulted borrowers whose mortgages were modified via HAMP progress to a *lis pendens* filing a higher rate than defaulted borrowers without a modification or with a non-HAMP modification. After controlling for delinquency length (and other factors) however, we find that the HAMP program may have been effective in helping defaulted borrowers avoid foreclosure.

1 INTRODUCTION

Under a state law that was enacted on December 15, 2009, mortgage servicers must send a “pre-foreclosure filing” (PFF) notice to delinquent borrowers at least 90 days prior to filing for foreclosure on a primary residence in the State of New York. The notice informs homeowners that their loan is in default, lists the amount necessary to cure the default and lists measures that they can take to avoid foreclosure, such as negotiating a loan modification with their lender and consulting with a non-profit housing counselor (New York State Banking Department, 2009).

Since February 13, 2010, mortgage servicers are also required to file the notices with the New York State Banking Department (NYSBD), which collected an extraordinary level of detail on the loans. Among the many data fields collected are: the property address, the names of the borrowers, the current monthly payment, the delinquent contractual payments, the interest rate, whether the loan is a fixed-rate or adjustable-rate mortgage, the date and the amount of the original loan, the lien type, the loan term, whether the loan has been modified or not and whether an investor’s approval is necessary to modify the loan. If the loan progresses to a *lis pendens* filing (i.e. the first step in the foreclosure process – the filing

*Brooklyn College, City University of New York – eric@doviak.net

†New York City College of Technology, City University of New York – smacdona1d@citytech.cuny.edu

of the complaint), then servicers are also required to follow up on their initial filing with information on the entity filing for foreclosure.

The detail captured in the PFF data makes three forms of analysis possible. First, we can match the defaulted loans to publicly available data on originations from the Home Mortgage Disclosure Act (HMDA). By combining the HMDA and PFF data, we can see which borrowers were more likely to default. Second, we can compare the loans that entered the foreclosure process to those that did not. For reasons described below, we call this the “Short PFF” dataset. Finally, our “Full PFF” dataset allows us to compare defaulted loans across the years in which they were originated.

This paper discusses the findings of our analysis of the Full PFF and Short PFF datasets. We discuss our analysis of the combined HMDA-PFF dataset in a separate paper (Doviak and MacDonald, 2011).

Not surprisingly, the financial characteristics of a loan are the best predictors of progression from default to a *lis pendens* filing. Larger loan amounts, larger monthly payments and adjustable interest rates all make a defaulted borrower more likely to enter the foreclosure process. Although the PFF data does not contain the loan-to-value ratio, we can surmise that borrowers who took out larger loans were left with less equity (or even negative equity) in his/her home after home prices tumbled during the recent economic and financial crisis. Consequently, borrowers who took out larger loans may have had greater incentive to walk away from the loan and shift the loss onto the lender.

Another shortcoming of the PFF data is that it does not contain information on the borrower’s current or past income. Nonetheless, we can surmise that larger monthly payments would make it more difficult for the borrower to remain current on the loan.

All else equal, a higher interest rate also increases the borrower’s monthly payments, thus making the loan more difficult to repay. In this regard, it is important to note that mortgages without a fixed interest rate were more likely to progress from default to a *lis pendens* filing than those with a fixed interest rate. Once again, this is not surprising. When interest rates reset upward, the increasing cost of these adjustable-rate mortgages left borrowers unable to afford the mortgage.

What is surprising is the Home Affordable Modification Program (HAMP) may have been more successful in preventing foreclosure than its critics suggest. Like many other studies, we find that HAMP-modified mortgages progress to a *lis pendens* filing a higher rate than mortgages in default that were not modified or were modified outside of the HAMP program. The higher rate of progression may be attributable to the longer length of time HAMP-modified mortgages have been delinquent however. After controlling for delinquency length (and other factors), we find that HAMP-modified mortgages were less likely to progress to a *lis pendens* filing, which indicates that the HAMP program may have been effective in helping defaulted borrowers avoid foreclosure.

Prior to discussing those findings, we will first discuss the literature on mortgage modification and foreclosure prevention in section 2. Then, in section 3, we describe the PFF data in more detail and how we prepared it for analysis.

Section 4 is the heart of this paper. It explores the question of “Who enters the foreclosure process?” by comparing the defaulted loans that did not enter the foreclosure process to those that did. To control for the many different factors that can affect a the probability that a loan will progress from default to foreclosure, section 5 provides a very basic regression analysis that attempts to resolve some of the puzzles that we find in the comparisons and continues to explore the racial and ethnic dimensions of the

foreclosure crisis.

Section 6 concludes with a discussion of how lenders can use our findings to reduce the losses that they suffer when a loan enters the foreclosure process.

2 REVIEW OF THE FORECLOSURE PREVENTION LITERATURE

The subprime mortgage foreclosure crisis, which began in 2006 and escalated rapidly throughout 2007 and 2008, was one of the first indicators of the forthcoming bursting of the nation's housing bubble. It also foreshadowed the onset of the financial and credit crises that unfolded during 2008 – 2009, resulting in a persistently high rate of unemployment that contributed to the national wave of foreclosures that began in 2008.

Unemployment alone does not explain the recent surge in defaults and foreclosures however. The character of the loans originated between 2004 and 2008 also played a significant role, so regulators responded by adopting several major loan modification programs. This section provides an overview of loan modification programs adopted since 2008 and discusses their success in stemming the foreclosure crisis.

The primary programs – the Making Home Affordable program¹, the Home Affordable Modification Program (HAMP)² and the FHA's HOPE for Homeowners Program³ – have relied upon the voluntary participation of lenders and servicers. These programs have employed a number of incentives to encourage participants to modify the loans of homeowners at risk of default and foreclosure. Overall, they have reached a relatively small percentage of the borrowers in need because the participation of lenders and loan servicers is voluntary and because many lenders and loan servicers have resisted efforts to persuade them to reduce principal balances.

According to CoreLogic (2011), at the end of the fourth quarter of 2010, approximately 11 million U.S. households – representing an estimated 23 percent of all homes with a mortgage – owed more on their homes than their home was worth. As of May 2011, more than two million borrowers were seriously delinquent on their loans (90 days or more past due), while another two million homeowners were in some stage of the foreclosure process as of May 2011 (RealtyTrac, 2011).

In an effort to avoid such outcomes, the US Treasury Department established HAMP as part of the Making Home Affordable program in October 2009. HAMP aims to modify the principal balance and interest rate on the home mortgages of borrowers who have good repayment histories but are experiencing financial hardship (U.S. Dept. of the Treasury, 2009).

Analysis of the Treasury Department's own published reports on HAMP's performance indicates that, as of June 2011, a total of 772,559 homeowners were participating in active trial modifications or active permanent modifications (U.S. Dept. of the Treasury, 2011). Thus, slightly more than three-quarters of a million distressed homeowners nationwide are or have been assisted by the Making Home Affordable program. More distressing however is the fact that almost half (46.4 percent) of the 1.6 million trial

¹The Making Home Affordable program, which was created under the Emergency Economic Stabilization Act (EESA) of 2009, provided \$29.9 billion for participating institutions.

²The Home Affordable Modification Program (HAMP) is part of the Making Home Affordable program.

³The FHA's HOPE for Homeowners Program was enacted under the 2008 Housing and Economic Recovery Act.

modifications started under the Making Home Affordable program were cancelled as of June 2011 (U.S. Dept. of the Treasury, 2011).

These programs may have had some success in reducing the foreclosure rate. According to RealtyTrac (2011), in May 2011, foreclosure filings nationally declined to their lowest level since December 2006, reaching a 53-month low. The decline also reflected a 33 percent decrease from May 2010.

Other programs have been far less successful.

Using its authority to provide mortgage insurance for refinancing the mortgages of borrowers at risk of default and foreclosure, the FHA introduced its HOPE for Homeowners Program in October 2008. Four months after the program began however, the FHA had only received 451 applications and closed a mere 25 loans, far short of the 400,000 homeowners that the program had been expected to help (Naylor, 2009).

While changes were made to attract participation, its prospects are not very bright. One key problem is that participation is voluntary. Lenders do not have to agree to a modification that writes down any portion of the loan's value. A second problem is that the borrowers most in need are the one least like to be served by the program. For example, borrowers with a lower credit score or more erratic payment history must meet more stringent qualifying criteria. Moreover, payment-to-income ratios can run as high as 38 percent (U.S. Office of the Comptroller of the Currency, 2010). Such requirements deter both borrowers and lenders from participating.

More importantly, the continued decline in home values places lenders at risk of holding a loan that is greater than the home's value and thus eliminating the possibility of securitizing the loan. Home prices declined an average of 8.2 percent from first quarter 2010 through first quarter 2011 (Humphries, 2011). The resulting low appraisal values have hindered millions of homeowners – not just those at risk – from refinancing their loans.

From the perspective of a borrower in a negative equity position, shifting losses onto the lender through the foreclosure process may be a regrettable but financially sensible course of action. Consequently, reductions in principal balance may be necessary to avert foreclosure. This was a primary finding of the State Foreclosure Prevention Working Group (SFPWG)⁴, which found that loan modifications involving "significant payment reduction has succeeded in creating more sustainable loan modifications" (2010).

In particular, the SFPWG, which analyzed a longitudinal dataset from nine loan servicers, found that more recent modifications which included significant reductions in principal balance tend to have lower re-default rates than their counterparts. However, modifications with a significant reductions in principal balance represent just 20 percent of the loan modifications that they studied. In most modifications, the loan amount increased as service charges and late payments were rolled into the loan.

To convince the borrower to accept the modification, lenders and servicers preferred to reduce the monthly payment. In the SFPWG's dataset, 89 percent involved a reduction in monthly payments. Of those, the modification reduced the monthly payment by more than 10 percent in 77 percent of cases (State Foreclosure Prevention Working Group, 2010).

⁴The State Foreclosure Prevention Working Group, which began publishing findings on mortgage delinquency and loss mitigation trends in early 2008, includes representatives of the attorneys general from twelve states, state bank regulators from three states and the Conference of State Bank Supervisors.

Table 1: Distribution of Pre-Foreclosure Filings by Year of Origination

	total	percent
1976-1989	2,502	1.3%
1990-1999	13,692	7.3%
2000	2,414	1.3%
2001	4,390	2.4%
2002	7,470	4.0%
2003	16,706	9.0%
2004	18,669	10.0%
2005	28,506	15.3%
2006	35,947	19.3%
2007	31,771	17.0%
2008	16,019	8.6%
2009	6,957	3.7%
2010	1,323	0.7%
total	186,366	100.0%
<i>Data: Full PFF</i>		

3 THE NEW YORK STATE PRE-FORECLOSURE FILING DATA

As mentioned in the introduction, in February 2010, the New York State Banking Department (NYSBD) began collecting data on home mortgages in default. When the borrower defaults on his/her primary residence, his/her mortgage servicer sends him/her a “pre-foreclosure filing” (PFF) notice and transmits an extraordinary level of detail on the mortgage to the NYSBD. If the borrower does not cure the default within 90 days, the servicer may commence the foreclosure process with a *lis pendens* filing. If it chooses to do so, it must also inform the NYSBD of the *lis pendens* filing.

Because the PFF dataset contains information on both defaults and foreclosures, the analysis in this paper compares defaulted loans that did not progress to foreclosure to those that did. Given the 90-day window between the date that the PFF notice was sent and the commencement of foreclosure proceedings, such an analysis requires examination of a subset of the data. We refer to this subset as the “Short PFF” dataset to distinguish it from the “Full PFF” dataset, which contains all of the filings.

This paper also makes reference to the Full PFF dataset because it is useful for comparing defaulted loans by year of origination. For example, table 7 shows that defaulted loans originated between 2004 and 2007 were more likely to be adjustable rate mortgages than loans originated in other years.

Prior to making those comparisons however, we first explain how we prepared the PFF dataset for statistical analysis in subsection 3.1. After providing that explanation, we discuss our comparisons in section 4 and we provide a very basic regression analysis in section 5.

Table 2: Rate of Lis Pendens by Month when PFF Letter Sent

	Lis Pendens	PFF Letters	rate
January	9	1,573	0.6%
February	1,745	8,171	21.4%
March	1,416	12,625	11.2%
April	1,725	12,847	13.4%
May	1,807	14,135	12.8%
June	1,474	27,810	5.3%
July	657	20,139	3.3%
August	405	16,779	2.4%
September	68	13,837	0.5%
October	24	15,395	0.2%
November	119	13,438	0.9%
December	17	27,623	0.1%
total	9,466	184,372	5.1%
<i>Data: Full PFF</i>			

3.1 Preparing the Data for Analysis

Prior to performing an analysis of the PFF data, we had to remove duplicate filings because servicers who missed the three-business day deadline or submitted incorrect information would “re-file” the loan. Some servicers also submitted one filing for each borrower on the loan.

The duplicates were fairly easy to identify however, because servicers almost always included their loan numbers with the filing, so the combination of the servicer’s identity and the loan number enabled us to uniquely identify each loan⁵. In cases where a servicer submitted one filing for each borrower, we compared the borrower’s first and last name to the names of other borrowers on the loan to see if there was a co-applicant or not.

Because servicers re-filed a loan to correct mistakes, we assumed that the filing which was submitted last contained the correct information. However if one of the duplicates contained information on a lis pendens filing, we retained that information.

Using this method, we found a total of 214,705 unique loans and 33,859 duplicates in the PFF dataset. From there, we removed records that contained obvious errors (e.g. loans that were originated in the future) and records of 90-day letters that were not mailed in the year 2010. This reduced the PFF dataset to 211,962 clean records.

To ensure comparability across loans, we chose to focus on first-lien mortgages. This reduced the PFF dataset to 186,366 records, but it was a necessary step because a first-lien mortgage is very different from a home equity line of credit (HELOC). The former is frequently taken for the purpose of purchasing a home, while the latter is often used for home improvement.

⁵In cases where the servicer did not include a loan number, we used the property address instead of the loan number.

Table 3: Lis Pendens Filing by Loan Amount (in thousands)

	no lis pendens	lis pendens	percent
under 50	6.3%	3.4%	5.9%
50 to 99	19.7%	14.1%	18.8%
100 to 249	29.9%	26.2%	29.3%
250 to 399	26.3%	33.1%	27.4%
400 to 499	9.4%	12.4%	9.9%
500 and up	8.2%	10.8%	8.6%
total	36,865	7,152	44,017
<i>Data: Short PFF</i>			

Table 4: Lis Pendens Filings by Monthly Payment

	no lis pendens	lis pendens	percent
under 1,000	26.2%	17.1%	24.7%
1,000 to 1,499	14.6%	12.3%	14.2%
1,500 to 1,999	13.3%	13.1%	13.3%
2,000 to 2,499	13.3%	15.5%	13.7%
2,500 to 2,999	12.5%	15.3%	13.0%
3,000 to 3,999	13.0%	17.6%	13.7%
4,000 and up	7.0%	9.2%	7.4%
total	36,865	7,152	44,017
<i>Data: Short PFF</i>			

Finally, we wanted to know which of the loans that were in default progressed to foreclosure. Given the required 90-day period between filing a pre-foreclosure notice and the commencement of foreclosure proceedings (with a lis pendens filing) – we also had to restrict our attention to the loans that were filed prior to July 1, 2010 when preparing the Short PFF dataset.

Ideally, we would have chosen a later cut-off date, but the fact that Bank of America, JP Morgan Chase and GMAC Mortgage suspended foreclosure actions in late-September 2010 left us with little choice. As table 2 shows, there was a sharp drop in the number of loans that progressed to a lis pendens filing among loans which received a pre-foreclosure filing from July 2010 onward.

Another difficulty arose because the law which required mortgage servicers to file the pre-foreclosure filing notices did not explicitly require servicers to notify the NYSBD when the mortgage progressed to a lis pendens filing. The NYSBD strongly pressed servicers to update the filings however and most servicers complied.

Nonetheless, the suspension of foreclosure action by three of the largest servicers and the small degree of non-compliance with the NYSBD's two-step filing process convinced us that we should limit the

Table 5: Lis Pendens Filings by Interest Rate

	no lis pendens	lis pendens	percent
under 4.000	4.5%	3.5%	4.4%
4.000 to 4.999	4.3%	4.1%	4.3%
5.000 to 5.999	21.9%	20.6%	21.7%
6.000 to 6.999	34.4%	39.6%	35.3%
7.000 to 7.999	17.5%	17.1%	17.4%
8.000 to 9.999	12.1%	11.1%	11.9%
10.000 and up	5.3%	4.1%	5.1%
total	36,865	7,152	44,017
<i>Data: Short PFF</i>			

Table 6: Lis Pendens Filing by Loan Detail

	no lis pendens	lis pendens	total
Fixed Rate	84.4%	15.6%	35,117
Adj. Rate	82.6%	17.4%	7,309
Pay. Op. Adj. Rate	78.5%	21.5%	451
Interest Only	73.5%	26.5%	1,140
percent	83.8%	16.2%	44,017
<i>Data: Short PFF</i>			

Short PFF dataset to loans submitted by servicers with at least 40 total filings and notified the NYSBD of a lis pendens filing on at least five percent of their loans. This step ensured that the information in the dataset would accurately reflect whether the loan progressed to a lis pendens filing or not.

4 WHO ENTERS THE FORECLOSURE PROCESS?

As mentioned in the introduction, the financial characteristics of a loan – such as the loan amount, monthly payment and whether the interest rate is fixed or adjustable – are the best predictors of progression from default to foreclosure.

For example, table 3 shows that – among borrowers who defaulted – 56 percent of the ones who entered the foreclosure process with a lis pendens filing borrowed more than \$250,000, whereas only 44 percent of the borrowers who did not go into foreclosure borrowed more than \$250,000.

Unfortunately, the PFF dataset does not have information on the loan-to-value ratio, so we do not know if and how far the recent collapse in home prices pushed these borrowers “underwater.” Nonetheless, borrowers who took out larger loans would have had greater incentive to shift their losses onto their

Table 7: Loan Detail by Year of Origination

	Fixed Rate	Adj. Rate (AR)	Pay. Option AR	Int. Only	total
1976-1989	51.1%	48.8%	0.0%	0.0%	2,502
1990-1999	89.1%	10.8%	0.1%	0.0%	13,692
2000	90.6%	9.2%	0.0%	0.1%	2,414
2001	93.9%	6.0%	0.0%	0.0%	4,390
2002	93.2%	6.6%	0.0%	0.1%	7,470
2003	92.4%	7.3%	0.1%	0.2%	16,706
2004	83.5%	15.3%	0.4%	0.8%	18,669
2005	73.8%	22.5%	1.2%	2.5%	28,506
2006	69.0%	26.2%	1.9%	2.9%	35,947
2007	77.7%	17.3%	1.8%	3.2%	31,771
2008	93.9%	4.6%	0.6%	0.9%	16,019
2009	97.6%	2.3%	0.0%	0.2%	6,957
2010	92.4%	7.3%	0.0%	0.3%	1,323
percent	81.2%	16.1%	1.0%	1.7%	186,366
<i>Data: Full PFF</i>					

Table 8: Lis Pendens Filings by Additional Borrower

	no lis pendens	lis pendens	total
no co-borrower	83.3%	16.7%	26,501
co-borrower	84.4%	15.6%	17,516
percent	83.8%	16.2%	44,017
<i>Data: Short PFF</i>			

lenders by walking away from the loan if the drop in home prices left them with less equity (or negative equity).

Another possible reason why defaulted borrowers with large loan amounts are more likely to enter the foreclosure process is because – all else equal – they would have to make larger monthly payments. In fact, the distributions are very similar. Table 4 shows that, among borrowers who defaulted, 58 percent of the ones who entered the foreclosure process had a monthly payment of \$2,000 or more, whereas only 46 percent of the defaulted borrowers who did not progress to a lis pendens filing had a monthly payment in excess of \$2,000.

Surprisingly however, there is no clear relationship between a defaulted borrower's current interest rate and his/her chances of progressing to a lis pendens filing (as shown in table 5). The lack of a clear relationship may be attributable to the fact that we're looking at borrowers who have already defaulted. A high interest rate may be a good predictor of default, but not a good predictor of progression to a

Table 9: Lis Pendens Filings by Amount of Delinquent Payment

	no lis pendens	lis pendens	percent
under 1,000	4.1%	0.9%	3.5%
1,000 to 2,499	21.3%	9.6%	19.4%
2,500 to 4,999	26.0%	19.2%	24.9%
5,000 to 7,499	17.7%	18.0%	17.8%
7,500 to 9,999	8.5%	10.8%	8.9%
10,000 to 19,999	11.6%	20.4%	13.0%
20,000 to 49,999	7.1%	15.0%	8.4%
50,000 and up	3.6%	6.0%	4.0%
total	36,865	7,152	44,017
<i>Data: Short PFF</i>			

Table 10: Lis Pendens Filings by Modification

	no lis pendens	lis pendens	total
No modification	83.9%	16.1%	34,962
HAMP modification	81.3%	18.7%	4,335
Non-HAMP modification	85.2%	14.8%	4,720
percent	83.8%	16.2%	44,017
<i>Data: Short PFF</i>			

foreclosure filing.

The characteristic of interest rates that does predict progression from default to foreclosure is whether the interest rate is fixed or adjustable. Defaulted mortgages with adjustable interest rates progress to a lis pendens filing at slightly higher (but statistically significant) rate, while mortgages with payment option adjustable interest rates and interest only mortgages progress at a much higher rate (as shown in table 6).

Borrowers' difficulty in repaying loans with an adjustable rate mortgages also helps to explain why loans originated between 2004 and 2007 constitute 62 percent of all pre-foreclosure filings (as shown in table 1). Of the loans that went into default, those that were originated between 2004 and 2007 were more likely to be adjustable rate mortgages those originated in other years (as shown in table 7).

Interestingly however, adding a co-borrower to the loan did not necessarily reduce the chances that a loan would progress from default to foreclosure. As table 8 shows, the percentage of defaulted loans that progressed to a lis pendens filing was approximately the same for loans with a co-borrower and loans without a co-borrower (although the difference is statistically significant). The regression model in section 5 however suggests that loans with a co-borrower are less likely to progress to foreclosure.

Not surprisingly, defaulted borrowers who have to make a larger delinquent payment (i.e. the missed

Table 11: Lis Pendens Filings by Length of Delinquency

	no lis pendens	lis pendens	percent
less than 60 days	58.0%	31.8%	53.7%
61-90 days	15.1%	16.3%	15.3%
91-120 days	6.5%	9.9%	7.0%
over 120 days	20.5%	42.0%	24.0%
total	36,865	7,152	44,017
<i>Data: Short PFF</i>			

Table 12: Modifications by Delinquency Length

	No mod.	HAMP	non-HAMP	percent
less than 60 days	54.9%	28.9%	68.0%	53.7%
61-90 days	16.1%	13.4%	11.0%	15.3%
91-120 days	6.6%	11.7%	5.8%	7.0%
over 120 days	22.4%	46.0%	15.2%	24.0%
total	34,962	4,335	4,720	44,017
<i>Data: Short PFF</i>				

monthly payments plus late fees, etc.) are more likely to progress from default to a lis pendens filing. As shown in table 9, 70 percent of defaulted borrowers who entered the foreclosure process owed \$5,000 or more, whereas only 49 percent of defaulted borrowers who did not progress to a lis pendens filing owed \$5,000 or more.

Finally, table 10 indicates that defaulted borrowers whose mortgages were modified via the Home Affordable Modification Program (HAMP) progress to a lis pendens filing at a higher rate than defaulted borrowers whose mortgages were either not modified at all or modified outside of the HAMP program.

This is not necessarily evidence that the HAMP program was unsuccessful however. Table 12 shows that mortgage servicers tended to send a pre-foreclosure filing notice (i.e. our indicator of default) to borrowers in the HAMP program at a much later stage of delinquency and table 11 shows that borrowers who receive a pre-foreclosure filing notice at a later stage of delinquency are far more likely to progress to a lis pendens filing.

Consequently, the high rates of progression to a lis pendens filing among defaulted borrowers in the HAMP program may be attributable to the late stage at which the NYSBD was notified of the default. In section 5, we'll revisit this question and show that defaulted borrowers in the HAMP program are much less likely to progress to a lis pendens filing than defaulted borrowers whose mortgages were either not modified at all or modified outside of the HAMP program.

Table 13: Probit Models, dependent variable: Lis Pendens Filing

	model #1	model #2	model #3	model #4
In(Orig. Loan Amount)	0.0795 *** (0.0171)		0.0592 * (0.0275)	0.0658 * (0.0273)
In(Amt. Delinq. Pay.)	0.0456 *** (0.0123)	0.0424 ** (0.0131)	0.0411 ** (0.0131)	0.0386 ** (0.0131)
In(Monthly Pay.)		0.0779 *** (0.0182)	0.0283 (0.0294)	0.0267 (0.0292)
Delinq. 61-90 days	0.3429 *** (0.0219)	0.3444 *** (0.0220)	0.3443 *** (0.0220)	0.3483 *** (0.0220)
Delinq. 91-120 days	0.5230 *** (0.0290)	0.5258 *** (0.0291)	0.5260 *** (0.0291)	0.5315 *** (0.0291)
Delinq. over 120 days	0.6607 *** (0.0253)	0.6664 *** (0.0262)	0.6674 *** (0.0262)	0.6716 *** (0.0262)
Current Int. Rate	-0.0049 (0.0049)	-0.0090 (0.0048)	-0.0060 (0.0050)	-0.0050 (0.0050)
Adj. Rate	0.0190 (0.0212)	0.0242 (0.0211)	0.0199 (0.0212)	
Pay. Op. Adj. Rate	0.0178 (0.0708)	0.0524 (0.0705)	0.0275 (0.0715)	
Interest Only	0.1984 *** (0.0431)	0.2121 *** (0.0428)	0.2005 *** (0.0432)	
Not Fixed Rate Mortgage				0.0468 * (0.0196)
modified via HAMP	-0.1350 *** (0.0254)	-0.1358 *** (0.0255)	-0.1358 *** (0.0255)	-0.1404 *** (0.0254)
modified, not HAMP	0.0058 (0.0255)	0.0090 (0.0255)	0.0063 (0.0255)	0.0050 (0.0255)
Add'l Borrower on Loan	-0.0678 *** (0.0157)	-0.0678 *** (0.0157)	-0.0682 *** (0.0157)	-0.0711 *** (0.0157)
Pay. inc. Escrow	0.1697 *** (0.0200)	0.1558 *** (0.0202)	0.1650 *** (0.0206)	0.1700 *** (0.0204)
Loan Investor Owned	-0.1576 *** (0.0186)	-0.1567 *** (0.0186)	-0.1567 *** (0.0186)	-0.1507 *** (0.0185)
In(No. Filings by Servicer)	0.0474 *** (0.0067)	0.0470 *** (0.0068)	0.0468 *** (0.0068)	0.0464 *** (0.0067)
AIC	36,545	36,549	36,546	36,558

*** $p < 0.001$, ** $p < 0.010$, * $p < 0.050$, . $p < 0.100$

Standard errors in parenthesis. All models also contain an intercept term and dummies for region and year of origination. Those coefficients are not shown.

Data: Short PFF

5 SIMPLE ECONOMETRIC MODEL

With few exceptions, our findings that the financial characteristics of a home mortgages are good predictors of whether a loan progresses from default to foreclosure are not surprising. The rate of progression from default to a lis pendens filing is higher among defaulted borrowers who took out larger loans, who must make larger monthly payments and who face adjustable interest rates.

There were two puzzles however. One puzzle was why the rate of progression to a lis pendens filing bore no relation to the size of the interest rate. Another puzzle was why the difference in the rate of progression was so small between defaulted borrowers with and without a co-borrower.

Finally, we saw that a larger proportion of defaulted borrowers whose mortgages were modified via the HAMP program progressed to a lis pendens filing, but this difference may be attributable to the fact that their loans were reported to the NYSBD at a later stage of delinquency, so table 10 is not necessarily an indictment of the HAMP program.

In an attempt to resolve some of these puzzles, this section presents a very basic regression analysis in which we use probit models to estimate the probability that a defaulted loan will progress from default to a lis pendens filing. This allows us to examine the effect of one variable while holding others constant. The analysis presented here makes no effort to place the variables in a theoretical framework. Nor does it make much effort to check for robustness across specifications. Such work is left to future research.

Prior to discussing the probit models of the probability that a borrower will progress from default to a lis pendens filing, it is worth noting that the overall rate at which loans in the Short PFF dataset progressed from default to a lis pendens filing is only about 16 percent. The low overall rate is attributable to two factors. First, table 11 shows that about half of them were delinquent for less than 60 days when the servicer filed the pre-foreclosure filing notice. These loans were less likely to progress to a lis pendens filing than those that had been delinquent for a longer period of time. Second, table 9 shows that about 48 percent of the borrowers who defaulted owed less than \$5,000. Not surprisingly, their loans were less likely to progress to a lis pendens filing than borrowers who owed more.

In choosing variables to include in the probit models, we took note of the high degree of correlation among the original loan amount, the monthly payment and the amount of delinquent payment. Consequently, we excluded the original loan amount and the monthly payment from two of the regression models. This helps show that the original loan amount, the monthly payment and the amount of delinquent payment are all good predictors of progression from default to foreclosure.

Interestingly however, the regression results suggest that – after controlling for other factors – defaulted borrowers with an adjustable rate mortgage or a payment option adjustable rate mortgage do not progress to foreclosure at a significantly higher rate than defaulted borrowers with a fixed rate mortgage. However, the difference between defaulted borrowers with a fixed rate mortgage and defaulted borrowers with an interest only loan is statistically significant. Those with an interest only loan are more likely to progress to foreclosure.

The regression results also help resolve the two puzzles mentioned above. A defaulted borrower's probability of progressing to foreclosure is positively correlated with the interest rate that he/she pays on the loan and the coefficient is statistically significant.

The probit models also help resolve the puzzle that we observed in table 8 – the finding that loans

with a co-borrower progress to foreclosure at approximately the same rate as those without co-borrower. After controlling for other factors, defaulted loans with a co-borrower are negatively correlated with the probability of progression to a lis pendens filing.

Finally, the regression results indicate that the HAMP program was more successful than critics have argued. In the Short PFF dataset, defaulted borrowers with a HAMP-modified mortgage progressed to foreclosure at a higher rate, but the regression results suggest that this difference is attributable to the fact that they received a pre-foreclosure filing notice at a later stage of delinquency (as we saw in tables 10, 11 and 12). After controlling for other factors, defaulted borrowers with a HAMP-modified mortgage were significantly less likely to progress to foreclosure than their counterparts without a modification or with a non-HAMP modification.

6 CONCLUSION

Given our finding that large loan original amounts and high monthly payments are good predictors that a defaulted borrower will progress to a lis pendens filing, one could conclude that reducing the principal balances on home mortgages would substantially reduce the foreclosure rate.

Reducing principal balances may be impractical, however. In cases where borrowers have negative equity, this would require lenders to absorb potentially very large losses on their portfolio of mortgages. Secondly, an across-the-board reduction in principal balance would benefit a large number of borrowers who otherwise would not default on their mortgages.

If the modifications were well-structured however, so that the balance-sheet effect of the lower probability of progressing to foreclosure offsets the losses that the lender would suffer by taking the loan to foreclosure, then reducing principal balances might have the desired effect of reducing losses in the mortgage industry.

Assuming that such a structure could be found, it may depend on information that the NYSBD's pre-foreclosure filing data does not contain, such as the borrower's income or the purpose of the loan. The publicly-available data from the Home Mortgage Disclosure Act (HMDA) does contain this information however, so in future work, we plan to incorporate the information from the HMDA dataset into the Short PFF dataset to see how those factors affect a borrower's probability of progressing from default to foreclosure and explore other options that may help the industry reduce its losses. In that analysis, we will also attempt to quantify the savings that the industry would achieve from such modifications.

7 ACKNOWLEDGEMENTS

We would like to thank the New York State Banking Department for making the Pre-Foreclosure Filing data available to us and for supporting our research. The views expressed in this paper are our own opinions and do not necessarily reflect the opinions of the New York State Banking Department.

REFERENCES

- CoreLogic. "US Housing and Mortgage Trends", May 2011. URL http://www.corelogic.com/about-us/researchtrends/asset_upload_file722_8290.pdf.
- E. Doviak and S. MacDonald. "Who Defaults on their Home Mortgage?". 24 Sept. 2011.
- S. Humphries. "No Respite From Housing Recession in First Quarter". Zillow Real Estate Research, 8 May 2011. URL <http://www.zillow.com/blog/research/2011/05/08/no-respite-from-housing-recession-in-first-quarter/>.
- B. Naylor. "Homeowners Rescue Program Shows Slim Benefits". National Public Radio, 3 Feb. 2009. URL <http://www.npr.org/templates/story/story.php?storyId=100163398>.
- New York State Banking Department. "2009 Mortgage Foreclosure Law - Overview", 2009. URL <http://www.banking.state.ny.us/mf12009.htm>.
- RealtyTrac. "Foreclosure Activity Continues Free-Fall in May". RealtyTrac press release, 16 June 2011. URL <http://www.realtytrac.com/content/news-and-opinion/foreclosure-activity-continues-free-fall-in-may-6651>.
- State Foreclosure Prevention Working Group. "Redefault Rates Improve for Recent Loan Modifications". Memorandum on Loan Modification Performance, Aug. 2010. URL <http://www.csbs.org/regulatory/Documents/SFPWG/DataReportAug2010.pdf>.
- U.S. Dept. of the Treasury. "Final Notice of Proposed New Privacy Act System of Records for the Home Affordable Modification Program". *Federal Register*, 74(207):55621–55623, 28 Oct. 2009. URL <http://edocket.access.gpo.gov/2009/pdf/E9-25897.pdf>.
- U.S. Dept. of the Treasury. "Making Home Affordable Program Performance Report Through June 2011", 5 Aug. 2011. URL <http://www.treasury.gov/initiatives/financial-stability/results/MHA-Reports/Documents/June%202011%20MHA%20Report%20FINAL.pdf>.
- U.S. Office of the Comptroller of the Currency. "FHA HOPE for Homeowners Program". Community Developments, Apr. 2010. URL http://www.occ.gov/static/community-affairs/fact-sheets/FHA_HOPE_for_Homeowners_Program_FS.pdf.

Who Defaults on their Home Mortgage?

Eric Doviak* Sean MacDonald†

ABSTRACT

Since February 2010, detailed information on every home mortgage default and foreclosure in New York State must be filed with the New York State Banking Department (NYSBD). Pairing the NYSBD's data with data on originations from the Home Mortgage Disclosure Act (HMDA) enables us to identify the race and ethnicity of borrowers who defaulted on their home mortgages (in New York State).

Like many previous studies, we find strong racial and ethnic disparities in lending practices, but we do not find conclusive evidence that HMDA-measurable forms of discrimination increased a borrower's probability of default. After controlling for other factors, we find that the interest rates charged to black and Latino borrowers tended to be higher than the ones charged to their white and non-Latino counterparts. This may be one reason why blacks and Latinos tend to default at a higher rate, but other factors, such as the tendency of black and Latino borrowers to take out larger loans than their white and non-Latino counterparts, may also have contributed to the higher default rate among black and Latino borrowers.

1 INTRODUCTION

Under a state law that was enacted on December 15, 2009, mortgage servicers must send a “pre-foreclosure filing” (PFF) notice to delinquent borrowers at least 90 days prior to filing for foreclosure on a primary residence in the State of New York. The notice informs homeowners that their loan is in default, lists the amount necessary to cure the default and lists measures that they can take to avoid foreclosure, such as negotiating a loan modification with their lender and consulting with a non-profit housing counselor (New York State Banking Department, 2009).

Since February 13, 2010, mortgage servicers are also required to file the notices with the New York State Banking Department (NYSBD), which collected an extraordinary level of detail on the loans. Among the many data fields collected are: the property address, the names of the borrowers, the current monthly payment, the delinquent contractual payments, the interest rate, whether the loan is a fixed-rate or adjustable-rate mortgage, the date and the amount of the original loan, the lien type, the loan term, whether the loan has been modified or not and whether an investor's approval is necessary to modify the loan. If the loan progresses to a *lis pendens* filing (i.e. the first step in the foreclosure process – the filing

*Brooklyn College, City University of New York – eric@doviak.net

†New York City College of Technology, City University of New York – smacdona1d@citytech.cuny.edu

of the complaint), then servicers are also required to follow up on their initial filing with information on the entity filing for foreclosure.

The detail captured in the PFF data makes three forms of analysis possible. First, we can match the defaulted loans to publicly available data on originations from the Home Mortgage Disclosure Act (HMDA). By combining the HMDA and PFF data, we can see which borrowers were more likely to default. Second, we can compare the loans that entered the foreclosure process to those that did not. Third, we can use the “Full PFF” dataset to compare defaulted loans across the years in which they were originated.

The first two datasets – which compare originations to defaults and compare defaults to lis pendens filings – effectively generate a quasi-longitudinal analysis, in the sense that we track the universe of New York State home mortgages from origination to default to foreclosure. We use the term “quasi-longitudinal” however, because the PFF data only provides information on borrowers who defaulted in 2010. We do not have data on the borrowers who did not default or defaulted in other years. Moreover, there is no way to perfectly match the PFF data to the HMDA data. We believe that our matching strategy generates a reasonably accurate result, but there is no way to verify the match.

This paper discusses the findings of our analysis of the combined HMDA and PFF datasets. In the interest of brevity and to avoid confusion between the two datasets, we discuss our analysis of the PFF dataset (which tracks loans from default to foreclosure) in a separate paper (Doviak and MacDonald, 2011).

After matching the PFF data to the 2004-2008 HMDA data, we found that the loan amount was the best predictor of default. This finding is not surprising. All else equal, a borrower who took out a larger loan was left with less equity (or even negative equity) in his/her home after home prices tumbled during the recent economic and financial crisis. Consequently, the borrower who took out a larger loan may have had greater incentive to walk away from the loan and shift the loss onto the lender or may have faced greater obstacles in obtaining a loan modification.

What may be surprising to some readers is the extent to which blacks and Latinos borrowed more than their white and non-Latino counterparts. As discussed in section 4, about half of white borrowers and half of non-Latino borrowers borrowed less than \$200,000, whereas about a quarter of blacks and a quarter of Latinos borrowed less than \$200,000. This is one important reason why the foreclosure crisis has had a disproportionately larger impact on black and Latino communities.

Another interesting finding from the combined HMDA-PFF data is that “middle-income” borrowers (those with incomes between \$80,000 and \$200,000¹) seem to have received disproportionately more pre-foreclosure filing notices than low-income and high-income borrowers. Section 4 will also show that black and Latino borrowers were disproportionately middle-income, yielding another reason why the foreclosure crisis has had a disproportionately larger effect on their communities.

The combined HMDA-PFF data also shows that borrowers who took out “high-cost” loans (i.e. loans in which the interest rate at origination is three or more percentage points above the yield on the comparable U.S. Treasury bond) were more likely to default. This is not surprising. One would expect lenders to charge a higher rate to borrowers who are more likely to default as compensation for the additional risk. Moreover, a higher interest rate increases the borrower’s monthly payments, thus making the loan more difficult to

¹This odd definition of “middle-income” corresponds to the income brackets that had a greater than average rate of default.

repay.

What may be surprising to the reader is the extent to which blacks and Latinos were more likely to take out high-cost loans. This is a third reason why the foreclosure crisis has had a disproportionately larger effect on black and Latino communities.

Previous studies have argued that blacks and Latinos faced discriminatory lending practices. However, we must first check to see if there some were fundamental factors which help explain why they were more likely to take out high-cost loans.

To perform such a check, we controlled for other factors by regressing their loan's rate spread (i.e. the difference between the interest rate and the yield on the comparable U.S. Treasury bond) on a large number of variables, such as income, loan amount, whether there was a co-borrower on the loan, the purpose of the loan, etc. and dummy variables for race and ethnicity. As discussed in section 5, the regressions suggest that the interest rates on loans originated to blacks and Latinos were higher than those originated to their white and non-Latino counterparts and that the differences were statistically significant. The HMDA data does not contain important variables (such as the borrower's credit score and the loan-to-value) however, so the evidence of discrimination is not conclusive.

Assuming that blacks and Latinos did face discrimination however, the question of whether that discrimination increased their probability of default remains open. So in a second step, we regressed the probability of default on the predicted rate spread, a large number of other variables and dummy variables for race and ethnicity. In those regressions, the coefficient on the predicted rate spread was not statistically significant from zero in one of the regressions, but the coefficients on the dummy variables for black race and Latino ethnicity were positive and statistically significant.

Consequently, we cannot conclude that the HMDA-measurable form of discrimination (i.e. the rate spread) increased their probability of default. Instead, we have to accept at face value the finding of the regression analysis – that blacks and Latinos default at a significantly higher rate after controlling for other factors.

We do not believe however that the melanin level in a person's skin makes him/her more or less likely to default on his/her home mortgage. Instead we believe that black race and Latino ethnicity must be acting as a proxy for some missing variable that does increase their probability of default, such as differences in socio-economic characteristics, racial and ethnic disparities in the effect of the recent economic recession and/or forms of discrimination that we cannot measure with the HMDA data.

Because the foreclosure crisis has had a disproportionately large effect on blacks and Latinos, one cannot understand the mortgage foreclosure crisis without understanding its racial and ethnic dimensions. For this reason, we discuss the literature on discrimination in mortgage lending in section 2 and we will discuss the racial and ethnic disparities that we observe in the data in section 4.

In section 3, we describe the PFF data in more detail and explain how we paired it with the HMDA data. Section 4 compares the loan origination data to the pre-foreclosure filing data and attempts to explain why blacks and Latinos tend to default at a higher rate than their white and non-Latino counterparts. Section 5 provides a very basic regression analysis that attempts to resolve some of the puzzles that we find in the comparisons and continues to explore the racial and ethnic dimensions of the foreclosure crisis. Section 6 concludes.

2 REVIEW OF THE DISCRIMINATION LITERATURE

The subprime mortgage foreclosure crisis, which began to unfold in 2006, was one of the first indicators of the forthcoming bursting of the nation's housing bubble. It was also perhaps the first sign that the risky home mortgages that formed the basis of subprime lending were the major culprit in the rapidly escalating delinquency and foreclosure rates leading up to the banking and financial crisis of 2008.

To provide some perspective, subprime lending was virtually non-existent at the peak of the previous real estate boom in 1989-90. Subprime loans then increased from 5 percent of total mortgage originations in the U.S. in 1994 to almost 20 percent in 2005 (Doms et al., 2007).

A number of recent studies have examined the relationship between the subprime foreclosure crisis, predatory lending and the concentration of subprime loans within predominantly black and Latino and other minority communities. These studies have sought to explore the underlying causes of the subprime foreclosure crisis, from residential segregation within many of the nation's metropolitan areas (Rugh and Massey, 2010), the lack of alternatives to subprime lenders in predominantly minority communities (U.S. Dept. of Housing and Urban Development, 2000b), changes in home prices and home price volatility (Doms et al., 2007) and the inability to exclude the mortgage on a primary residence from protection following the passage of the Bankruptcy Reform Act of 2005 (Morgan et al., 2011).

As mentioned previously, this review focuses on the analyses that explore the racial and ethnic dimensions of high-cost lending and whether discriminatory lending practices contributed to the subprime mortgage foreclosure crisis. We acknowledge that discrimination may have occurred, but the empirical evidence of discrimination suffers from the limitations of the HMDA data, which omits important variables, such as the borrower's credit score, the borrower's other assets and the loan-to-value ratio.

Consequently, it is easy to show that high-cost lending was most prevalent in predominantly minority communities, but it is difficult to take the next step and use the HMDA data to show that such lending is evidence of discrimination.

Bocian et al. (2006) does overcome some of the HMDA data's limitations however. After pairing the 2004 HMDA data with a proprietary dataset of 177,000 subprime loans, they found that blacks and Latinos received a disproportionate share of high-cost loans after controlling for other factors, including the borrower's FICO score and the loan-to-value ratio. The major limitation of their study however is the fact that it does not contain the universe of originations. Instead the study focuses on the dataset of subprime loans that was available to them, so it is not generalizable to the broader market and it does not explain why borrowers took a subprime loan as opposed to a prime loan.

When employing the HMDA data to study the broader market researchers are generally confined to finding a correlation between racial segregation and the probability of receiving a high-cost loan. For example, Squires et al. (2009) use the 2000 Census data to construct a dissimilarity index to obtain a measure of the ten most segregated and the ten least segregated metropolitan areas in the U.S. They then compare the indices derived to the percentage of high-cost loans originated. Using 2006 HMDA data and the 2006 American Community Survey, they employ a multivariate OLS model (to control for several MSA-level variables) and find that racial segregation is a significant predictor of the percentage of high-cost loan originations in an MSA. Their results suggest that a 10 percent increase in black segregation was associated with a 1.4 percent increase in high-cost loans.

Other studies have also found a link between the racial composition of a neighborhood and its share of subprime lending in that neighborhood. For example, in a joint study conducted by several community organizations, Bromley et al. (2008) focused on subprime lending activity across seven large metropolitan areas in the U.S. in 2006. Data on the number of high-risk loan originations conducted by a sample of 35 subprime lenders during that year indicated that these lenders accounted for an estimated 20 percent of the market share of subprime loans in predominantly minority neighborhoods within these metropolitan areas. Further, more than 40 percent of the loans made by high-risk lenders in these metropolitan areas were in neighborhoods where the share of minority residents was 80 percent or more. Subprime lenders' market share was also positively correlated with the percentage of minority residents within a given census tract.

The U.S. Dept. of Housing and Urban Development (2000b) also finds a disproportionate concentration of subprime lending in predominantly minority – and particularly – African-American communities. In their study, which focuses primarily on subprime refinance lending, the number of subprime refinance loans originated in the New York metropolitan area between 1993 and 1998 increased by an estimated 350 percent. The study also found that subprime loans were three times more likely to be originated in lower-income neighborhoods in the New York metropolitan area than in higher-income neighborhoods, and more than four times more likely in predominantly black than in predominantly white neighborhoods.

It's particularly interesting to note that their study was published in 2000, which indicates that subprime lending expanded rapidly into minority communities long before the subprime mortgage meltdown began in 2006. According to Laderman (2001), one factor which contributed to the expansion of subprime mortgage lending in the early 1990s was the increasing frequency with which mortgages were securitized. Securitization reduced the risk associated with lending to subprime borrowers and it enabled large sums to be assembled for the purpose of subprime lending. Another factor that Laderman cites was deregulation. Prior to passage of the Depository Institutions Deregulation and Monetary Control Act in 1980, limits were imposed on the interest rates that lenders could charge. Once the caps were lifted, lenders could raise interest rates high enough to absorb the risk associated with lending to subprime borrowers.

In a separate but related report, the U.S. Dept. of Housing and Urban Development (2000a) found that the pattern of originating subprime loans to minorities transcended income level and that this pattern established itself long before the subprime loan market reached its peak during the early 2000s. Instead, borrowers in high-income black neighborhoods were twice as likely as those in low-income white neighborhoods to take out a subprime loan. Specifically, the study found that just six percent of borrowers in high-income white neighborhoods had subprime loans while 39 percent of borrowers in upper-income black neighborhoods had subprime loans. This figure was more than twice the 18 percent rate for borrowers in low-income white neighborhoods.

Such findings are disturbing. The lack of information on credit scores in the HMDA data may explain some of the disparities in the rate spreads among individual borrowers, but it is hard to see how this could be applicable across neighborhoods. In other words, it is easy to imagine individual cases where a high-income black borrower's credit score is lower than a low-income white borrower's credit score; however it is difficult to see how the average credit score of a high-income black neighborhood could be lower than the average credit score of a low-income white neighborhood.

Given that blacks and Latinos took a disproportionately high share of subprime loans, one would also

expect a disproportionately high rate of foreclosure in black and Latino communities. This is precisely what two other studies have found.

Rugh and Massey (2010) attempt to link the correlation between the high-cost lending and the patterns of residential segregation to the subprime foreclosure crisis. To find the link, they obtained the total number of foreclosures between 2006 and 2008 from RealtyTrac's foreclosure database and computed the foreclosure rate as the number of filings per household unit. They then used the 2004-2006 HMDA data to compute the share of high-cost loans² in each MSA. To derive a measure of regulatory oversight, they also computed the share of loans within the MSAs that were originated by institutions covered under the Community Reinvestment Act (CRA).

Employing an OLS multiple regression model, Rugh and Massey regress the number and rate of foreclosures in the nation's 100 largest MSAs on two measures of segregation: residential unevenness and spatial isolation. Their regression results suggest that residential segregation and the share of high-cost loans are both positively correlated with the number and rate of foreclosures across U.S. metropolitan areas.

One frustrating omission in their published paper however is the lack of a regression of the high-cost lending share on measures of racial and ethnic segregation. If segregation enabled lenders to target minorities for high-cost loans (as Rugh and Massey claim), then they should have regressed the high-cost lending share on measures of segregation. If the coefficient were positive and statistically significant, then their claims of racial and ethnic targeting would have a firmer foundation.

Gerardi and Willen (2008) also examine the relationship between foreclosures and subprime lending in urban and minority communities. By matching the 1998-2006 HMDA data to deed registry data in the State of Massachusetts, they generate a dataset that contains the universe of mortgages, foreclosures, purchases and sales. In their analysis of the data, they find that a disproportionate share of subprime loans were originated to blacks and Latinos, but these loans proved unsustainable when home prices fell. The records of property sales in their dataset indicate that a "sudden and severe fall in the share of minority home ownership" began in 2005 due to a significant increase in foreclosures among minority homeowners.

The studies reviewed above show that blacks and Latinos took a disproportionately high share of high-cost and subprime loans, but the evidence that this trend reflects discrimination suffers from the limitations of the HMDA data. Nonetheless, the studies do help explain our finding that blacks and Latinos defaulted on their mortgages at a higher rate than their white and non-Latino counterparts.

3 THE NEW YORK STATE PRE-FORECLOSURE FILING DATA

As mentioned in the introduction, in February 2010, the New York State Banking Department (NYSBD) began collecting data on home mortgages in default. When the borrower defaults on his/her primary residence, his/her mortgage servicer sends him/her a "pre-foreclosure filing" (PFF) notice and transmits an extraordinary level of detail on the mortgage to the NYSBD. If the borrower does not cure the default within 90 days, the servicer may commence the foreclosure process with a *lis pendens* filing. If it chooses

²Rugh and Massey use the term "subprime" to describe high-cost loans.

to do so, it must also inform the NYSBD of the lis pendens filing.

Given our discussion of the racial and ethnic disparities in mortgage lending during the previous decade, one interesting way to analyze the PFF data is to match the loans to the HMDA originations data and compare the borrowers who defaulted in 2010 to those who did not default.

Because the PFF dataset contains information on both defaults and foreclosures, another interesting way to analyze the data is to compare the defaulted loans that did not progress to foreclosure to those that did. As mentioned previously, we discuss these comparisons in a separate paper (Doviak and MacDonald, 2011).

Prior to comparing originations to defaults however, we first explain how we prepared the PFF dataset for statistical analysis in subsection 3.1. Then, in subsection 3.2, we explain how we matched the PFF data to the HMDA data. After providing those explanations, we discuss our comparisons in section 4 and we provide a very basic regression analysis in section 5.

3.1 Preparing the Data for Analysis

Prior to performing an analysis of the PFF data, we had to remove duplicate filings because servicers who missed the three-business day deadline or submitted incorrect information would “re-file” the loan. Some servicers also submitted one filing for each borrower on the loan.

The duplicates were fairly easy to identify however, because servicers almost always included their loan numbers with the filing, so the combination of the servicer’s identity and the loan number enabled us to uniquely identify each loan³. In cases where a servicer submitted one filing for each borrower, we compared the borrower’s first and last name to the names of other borrowers on the loan to see if there was a co-applicant or not.

Because servicers re-filed a loan to correct mistakes, we assumed that the filing which was submitted last contained the correct information. However if one of the duplicates contained information on a lis pendens filing, we retained that information.

Using this method, we found a total of 214,705 unique loans and 33,859 duplicates in the PFF dataset. From there, we removed records that contained obvious errors (e.g. loans that were originated in the future) and records of 90-day letters that were not mailed in the year 2010. This reduced the PFF dataset to 211,962 clean records.

To ensure comparability across loans, we chose to focus on first-lien mortgages. This reduced the PFF dataset to 186,366 records, but it was a necessary step because a first-lien mortgage is very different from a home equity line of credit (HELOC). The former is frequently taken for the purpose of purchasing a home, while the latter is often used for home improvement.

Our analysis pays particular attention to the 130,912 first-lien mortgages that were originated in the years 2004-2008. Table 1 shows that these five years account for 70 percent of all PFF filings on first-lien mortgages.

We chose to work with the years 2004-2008 because we wanted to compare the PFF data to the data on originations from the Home Mortgage Disclosure Act (HMDA). We chose 2004 as the first year,

³In cases where the servicer did not include a loan number, we used the property address instead of the loan number.

Table 1: Distribution of Pre-Foreclosure Filings by Year of Origination

	total	percent
1976-1989	2,502	1.3%
1990-1999	13,692	7.3%
2000	2,414	1.3%
2001	4,390	2.4%
2002	7,470	4.0%
2003	16,706	9.0%
2004	18,669	10.0%
2005	28,506	15.3%
2006	35,947	19.3%
2007	31,771	17.0%
2008	16,019	8.6%
2009	6,957	3.7%
2010	1,323	0.7%
total	186,366	100.0%
<i>Data: Full PFF</i>		

because the variables available in the pre-2004 HMDA data were quite limited. At the time of this writing, the 2009 HMDA data was available to us, but we chose not to work with it because lending practices changed dramatically after the subprime mortgage crisis crippled the world financial system in late 2008. Loans originated in 2009 were very different from loans originated in previous years, so – for this analysis – we wanted to focus on loans originated in the years leading up to and including the crisis. In a forthcoming analysis, we will compare lending patterns in the periods before and after the crisis and how those differences affect the rate at which borrowers default on their home mortgages.

3.2 Matching the Pre-Foreclosure Filing Data to the HMDA Originations Data

The HMDA originations data contains the FIPS county code and census tract number of each property. This is a particularly valuable piece of information because census tracts have a small population (between 2,500 and 8,000 people) which is fairly homogeneous in terms of socio-economic characteristics and living conditions (U.S. Census Bureau, 2000).

So our first step in matching the PFF data to the HMDA data was to use the address information to identify the census tract of each property in the PFF dataset. To identify the census tracts, we used Erle's (2005) "Geo-Coder-US-1.00" Perl module in conjunction with the U.S. Census Bureau's (2007) TIGER/Line Files.

After using Erle's Perl module to create a database of New York State addresses from the TIGER/Line Files, we queried the database to obtain the latitudes and longitudes of the property addresses in the PFF dataset. Once we had the coordinates, we compared them to a database of census tract coordinates that

Table 2: Pre-Foreclosure Filings by Loan Amount

	no PFF	received PFF	percent
under 50	4.9%	2.8%	4.8%
50 to 99	16.5%	13.4%	16.3%
100 to 249	36.1%	27.7%	35.4%
250 to 399	25.8%	33.7%	26.4%
400 to 499	8.3%	12.7%	8.6%
500 and up	8.4%	9.7%	8.5%
total	1,544,118	130,722	1,674,840

Data: Combined HMDA-PFF

Table 3: Pre-Foreclosure Filings by Applicant Income

	no PFF	received PFF	percent
under 40	10.9%	9.9%	10.8%
40 to 59	18.0%	15.6%	17.8%
60 to 79	19.2%	18.3%	19.1%
80 to 99	15.8%	17.3%	15.9%
100 to 119	10.9%	12.9%	11.1%
120 to 159	11.9%	14.0%	12.0%
160 to 199	5.0%	5.4%	5.0%
200 and up	8.4%	6.6%	8.2%
total	1,465,078	123,878	1,588,956

Data: Combined HMDA-PFF

we generated from the U.S. Census Bureau's (2005) "Cartographic Boundary Files."

Using this method, we were able to identify the census tracts for 96 percent⁴ of the addresses in the PFF database. To avoid losing the information that the other four percent contain, we identified each of the census tracts within the property's five-digit zip code and counted the number of times each census tract corresponded to that zip code. We then randomly assigned the property to one of those census tracts (using the number of occurrences as weights).

Once the Census Tracts of each property had been identified and we had purged the duplicates, matching the pre-foreclosure filing data to the HMDA originations data was fairly simple. We divided owner-occupied⁵, first-lien mortgages in the HMDA data and first-lien mortgages in the PFF data into buckets by year of origination, census tract and co-applicant status. On average, there were 34 loans in

⁴238,830 of the 248,556 (non-unique) addresses

⁵Mortgage servicers only file pre-foreclosure filing notices when the property is a primary residence, so when matching the PFF data to the HMDA data, we focus on mortgages originated for owner-occupied properties.

Table 4: Pre-Foreclosure Filings by Loan Cost

	no PFF	received PFF	total
non-high cost	92.8%	7.2%	1,364,557
high cost	89.4%	10.6%	310,283
percent	92.2%	7.8%	1,674,840

Data: Combined HMDA-PFF

Table 5: High Cost Loans by Applicant Income

	non-high cost	high cost	percent
under 40	10.1%	13.9%	10.8%
40 to 59	17.8%	18.0%	17.8%
60 to 79	19.0%	19.4%	19.1%
80 to 99	15.6%	16.9%	15.9%
100 to 119	10.8%	12.1%	11.1%
120 to 159	12.1%	11.9%	12.0%
160 to 199	5.3%	4.1%	5.0%
200 and up	9.3%	3.8%	8.2%
total	1,290,774	298,182	1,588,956

Data: Combined HMDA-PFF

each HMDA bucket and 3 loans in each PFF bucket, so to figure out which HMDA origination corresponded to the pre-foreclosure filing, we compared the loan amounts and chose the closest match.

4 WHO DEFAULTS ON THEIR HOME MORTGAGE?

4.1 Financial Characteristics

Using the combined HMDA-PFF data, we find that the best predictor that a borrower would default is the amount borrowed. As table 2 shows, 56 percent of the borrowers who received a pre-foreclosure filing took loans in excess of \$250,000, whereas only 43 percent of the borrowers who did not default borrowed more than \$250,000.

It would be particularly insightful to compare the amounts that borrowers owe to the value of their homes. Unfortunately, HMDA does not provide the loan-to-value ratio or any information on the down payment, so we cannot make such a comparison. Nonetheless, if individuals who borrowed less have a larger equity stake in their homes, then these findings would illustrate the general principle that borrowers who have a larger equity stake in their home are less likely to default and enter the foreclosure process.

Repaying a mortgage also depends on the ability to pay, of course. But it's particularly striking to note

Table 6: High Cost Loans by Loan Amount

	non-high cost	high cost	percent
under 50	4.4%	6.6%	4.8%
50 to 99	15.9%	18.0%	16.3%
100 to 249	37.1%	28.1%	35.4%
250 to 399	25.8%	29.0%	26.4%
400 to 499	8.1%	10.8%	8.6%
500 and up	8.7%	7.6%	8.5%
total	1,364,557	310,283	1,674,840
<i>Data: Combined HMDA-PFF</i>			

Table 7: High Cost Loans by Additional Applicant

	non-high cost	high cost	total
no co-applicant	77.8%	22.2%	952,877
co-applicant	86.3%	13.7%	721,963
percent	81.5%	18.5%	1,674,840
<i>Data: Combined HMDA-PFF</i>			

that “middle-income” borrowers received pre-foreclosure filings at a higher rate than low and high-income borrowers. Specifically, as table 3 shows, the distribution of income among borrowers who defaulted was more heavily weighted in the \$80,000 to \$199,999 range than the comparable distribution of borrowers who did not default⁶.

Why middle-income borrowers default at a higher rate than low-income borrowers is puzzling. The regression models discussed in section 5 suggest however that borrowers with higher incomes are less likely to receive a pre-foreclosure filing after controlling for other factors, such as: loan amount, predicted rate spread, changes in county-level employment and changes in the MSA-level home price index.

Another good predictor of default is the interest cost of the loan. Table 4 shows that borrowers who took “high-cost” loans were more likely to receive a pre-foreclosure filing. When viewed in a risk-premium context, this finding should not be surprising. Borrowers who are more likely to default will have to compensate the lender for the additional risk by paying a higher interest rate.

However, there is also a risk that the additional cost of the loan will make the borrower more likely to default and go into foreclosure. In particular, a borrower’s monthly payment is an increasing function of the interest rate, so a higher interest rate reduces a borrower’s ability to repay the loan.

Lenders do not set interest rates exogenously however. Since a borrower’s income and loan amount

⁶Using a range of \$80,000 to \$199,999 is an odd way to define “middle income.” The range corresponds to the break points in table 3.

Table 8: Pre-Foreclosure Filings by Additional Applicant

	no PFF	received PFF	total
no co-applicant	91.1%	8.9%	952,877
co-applicant	93.6%	6.4%	721,963
percent	92.2%	7.8%	1,674,840
<i>Data: Combined HMDA-PFF</i>			

Table 9: High Cost Loans by Applicant Race

	non-high cost	high cost	total
Asian	89.7%	10.3%	89,998
Black/Afr. Am.	64.9%	35.1%	166,380
White	84.2%	15.8%	1,161,960
not provided	76.8%	23.2%	234,393
percent	81.5%	18.5%	1,674,840
<i>Data: Combined HMDA-PFF</i>			

affect his/her probability of default, all else equal one would expect lenders to compensate for the additional risk by charging a higher interest rate to low-income borrowers and borrowers who take out a larger loan.

In line with this reasoning, we find that low-income borrowers are more likely to receive a high-cost loan than borrowers with higher income. Table 5 shows 80 percent of high-cost loans were originated to borrowers with income below \$120,000, whereas only 73 percent of loans that were not high-cost loans were originated to such borrowers.

Surprisingly however, there does not appear to be any systematic relationship between loan amount and the likelihood of the loan being a high-cost loan. Table 6 shows that loan amounts below \$100,000 were more likely to be high-cost loans and loan amounts in the \$250,000 to \$499,999 range were also more likely to be high-cost loans.

It is difficult to understand why small loan amounts (i.e. those under \$100,000) were more likely to be high-cost loans and why large loan amounts (i.e. those over \$500,000) were less likely to be high-cost loans. Regression analysis does not even help to explain this puzzle. As discussed in section 5, borrowers who took out larger loan amounts tended to receive lower interest rates on their mortgages after controlling for other factors even though the larger loan amounts made them more likely to default.

Another important factor in explaining interest rates is whether there is a co-borrower on the loan or not. As table 7 shows, 22 percent of loans without a co-applicant were high-cost loans, whereas only 14 percent of loans with a co-applicant were high-cost loans. This may be attributable to the fact that a second borrower is a (potential) second source of income, which helps to mitigate the risk that the loan

Table 10: High Cost Loans by Applicant Ethnicity

	non-high cost	high cost	total
Hispanic/Latino	71.9%	28.1%	134,937
Not Hispanic/Latino	82.8%	17.2%	1,263,971
not provided	77.5%	22.5%	232,693
percent	81.5%	18.5%	1,674,840
<i>Data: Combined HMDA-PFF</i>			

Table 11: Pre-Foreclosure Filings by Applicant Race

	no PFF	received PFF	total
Asian	92.8%	7.2%	89,998
Black/Afr. Am.	88.0%	12.0%	166,380
White	92.8%	7.2%	1,161,960
not provided	91.7%	8.3%	234,393
percent	92.2%	7.8%	1,674,840
<i>Data: Combined HMDA-PFF</i>			

will go into default. As table 8 shows, 9 percent of loans without a co-borrower received a pre-foreclosure filing, whereas only 6 percent of loans with a co-borrower received a pre-foreclosure filing.

4.2 Race and Ethnicity

In section 2, we reviewed evidence of racial and ethnic discrimination in lending practices. The HMDA data captures one form of such discrimination – the difference in the rate spread between loans originated to minorities and loans originated to whites. As tables 9 and 10 show, blacks and Latinos received a disproportionately high share of high-cost loans. Asians, by contrast, received a disproportionately low share.

Tables 11 and 12 show that blacks and Latinos also received a disproportionately high share of pre-foreclosure filings, so one also has to wonder if racial and ethnic discrimination in lending practices contributed to the disproportionately high share of defaults among blacks and Latinos.

One way to address this question is to ask if fundamental differences between minorities and non-minorities justify the difference in rate spreads. If so, then the next question to ask is if those fundamental differences could have caused blacks and Latinos to default at disproportionately higher rates.

The first fundamental factor that we'll consider is income. If minority borrowers tended to have lower income than their non-minority counterparts, then one could justify the difference in rate spreads on the basis of income.

Table 12: Pre-Foreclosure Filings by Applicant Ethnicity

	no PFF	received PFF	total
Hispanic/Latino	89.0%	11.0%	134,937
Not Hispanic/Latino	92.4%	7.6%	1,263,971
not provided	92.0%	8.0%	232,693
total	92.2%	7.8%	1,674,840

Data: Combined HMDA-PFF

Table 13: Applicant Income by Applicant Race

	Asian	Black/Afr. Am.	White	not provided	percent
under 40	4.0%	8.0%	12.2%	8.7%	10.8%
40 to 59	11.7%	16.5%	18.9%	16.1%	17.8%
60 to 79	16.3%	23.0%	18.7%	19.4%	19.1%
80 to 99	17.3%	20.1%	15.1%	16.0%	15.9%
100 to 119	14.4%	13.6%	10.4%	11.0%	11.1%
120 to 159	17.6%	12.4%	11.5%	12.5%	12.0%
160 to 199	8.3%	3.7%	4.9%	5.5%	5.0%
200 and up	10.5%	2.7%	8.4%	10.8%	8.2%
total	85,965	156,030	1,105,913	220,741	1,588,956

Data: Combined HMDA-PFF

Such a hypothesis only finds partial support in the data. Table 13 shows that 26 percent of Asian borrowers and 18 percent of white borrowers had income over \$140,000, while only 11 percent of black borrowers did. The distribution of income by ethnicity shows a similar pattern. As table 14 shows, 18 percent of non-Latino borrowers had income over \$140,000, while only 14 percent of Latinos did.

The fact that there is more weight in the upper region of the distribution of income among non-minority borrowers than there is in the distribution of income among non-minority borrowers lends some support to the hypothesis that differences in income help explain why blacks and Latinos received a disproportionate share of high-cost loans.

However, the lower region of the income distributions refutes the hypothesis. It appears to have been easier for low-income whites and non-Latinos to obtain a mortgage. Specifically, 31 percent of white borrowers had income below \$60,000, while only 25 percent of black borrowers did. Similarly, 30 percent of non-Latinos had income below \$60,000, while only 19 percent of Latinos did.

Consequently, it would be hard to justify the disproportionate share of high-cost loans that blacks and Latinos received on the basis of income differentials.

A second fundamental factor to consider is the amount of the original loan. Differences in loan

Table 14: Applicant Income by Applicant Ethnicity

	Hispanic/Latino	Not Hispanic/Latino	not provided	percent
under 40	5.8%	11.6%	8.9%	10.8%
40 to 59	12.9%	18.5%	16.3%	17.8%
60 to 79	20.6%	18.9%	19.2%	19.1%
80 to 99	21.4%	15.4%	15.8%	15.9%
100 to 119	15.9%	10.6%	10.9%	11.1%
120 to 159	14.8%	11.7%	12.4%	12.0%
160 to 199	4.8%	5.0%	5.5%	5.0%
200 and up	3.8%	8.2%	11.0%	8.2%
total	125,440	1,203,686	219,669	1,588,956

Data: Combined HMDA-PFF

Table 15: Loan Amount by Applicant Race

	Asian	Black/Afr. Am.	White	not provided	percent
under 50	1.0%	3.2%	5.7%	2.7%	4.8%
50 to 99	6.3%	8.4%	19.1%	12.1%	16.3%
100 to 249	26.3%	28.3%	37.2%	35.2%	35.4%
250 to 399	33.3%	40.5%	23.0%	29.9%	26.4%
400 to 499	18.0%	12.8%	7.1%	9.4%	8.6%
500 and up	15.1%	6.8%	7.8%	10.7%	8.5%
total	89,998	166,380	1,161,960	234,393	1,674,840

Data: Combined HMDA-PFF

amounts help explain why blacks and Latinos received a disproportionate share of pre-foreclosure filings, but they do not necessarily explain why they received a disproportionate share of high-cost loans.

Specifically, minorities tended to borrow much more than their non-minority counterparts. According to table 15 shows, 53 percent of white borrowers borrowed less than \$200,000 whereas only 28 percent of blacks did. Interestingly however, Asians appear to have borrowed even more than blacks (only 25 percent borrowed less than \$200,000), but had the lowest rate of high-cost loans. Turning to ethnicity, table 16 shows that 50 percent of non-Latinos borrowed less than \$200,000, whereas 25 percent of Latinos borrowed less than that amount.

The finding that blacks and Latinos tended to borrow more helps explain why they received a disproportionately high share of pre-foreclosure filings, but it does not explain why they took high-cost loans at a higher rate than their white, Asian and non-Latino counterparts. Asians borrowed more, but took fewer high-cost loans. Moreover, as mentioned previously, the regression analysis in section 5 also refutes the hypothesis that borrowers who took out larger loan amounts would receive lower interest rates. The

Table 16: Loan Amount by Applicant Ethnicity

	Hispanic/Latino	Not Hispanic/Latino	not provided	percent
under 50	2.1%	5.4%	2.9%	4.8%
50 to 99	7.1%	17.8%	12.7%	16.3%
100 to 249	26.4%	36.1%	35.6%	35.4%
250 to 399	41.7%	24.4%	29.0%	26.4%
400 to 499	13.6%	8.1%	9.0%	8.6%
500 and up	9.2%	8.1%	10.7%	8.5%
total	134,937	1,263,971	232,693	1,674,840
<i>Data: Combined HMDA-PFF</i>				

opposite is true. All else equal, the rate spreads on larger loans tend to be lower.

5 ECONOMETRIC MODELS OF RATE SPREADS AND DEFAULTS

Section 4 describes several questions raised by the PFF data and the combined HMDA-PFF dataset. The most striking questions are why blacks and Latinos were more likely to take high-cost loans and why they are more likely to default on their mortgages. But there were other questions too. One is the lack of a clear relationship between the amount of the original loan and the whether the loan was a high-cost loan or not. Another was why middle-income borrowers were more likely to default than both low-income and high-income borrowers.

In an attempt to resolve some of these puzzles, this section presents a very basic regression analysis, so that we can examine the effect of one variable while holding others constant. The analysis presented here makes no effort to place the variables in a theoretical framework. Nor does it make much effort to check for robustness across specifications. Such work is left to future research.

The regression analysis presented here simply attempts to use the available variables to predict the rate spread on a borrower's loan. The predicted rate spread is then used as an instrument in a second-stage probit regression to estimate a borrower's probability of defaulting on the loan.

5.1 Econometric Methods

One problem confronting any econometric analysis of the HMDA data is how to work with the rate spread. The HMDA data only provides a value of the rate spread when the difference between the interest rate on the mortgage and the yield on the comparable U.S. Treasury exceeds three percentage points⁷. Consequently, when addressing the question of why black and Latino borrowers were more likely to take out a high-cost loan, we have to find a way to work with the rate spread.

⁷More precisely, the HMDA data provides a value for the rate spread of a first-lien mortgage when it exceeds three percentage points. For other lien statuses, the HMDA data provides a value for the rate spread when it exceeds five percentage points. Our analysis focuses exclusively on first-lien mortgages.

Table 17: Two-Stage: Tobit predicts Rate Spread, then Probit predicts PFF

	Model #1				Model #2			
	Tobit		probit		Tobit		probit	
Intercept	-0.0513	***	-2.1133	***	0.0037		-2.1071	***
	(0.0004)		(0.1183)		(0.0054)		(0.1715)	
Pred. Rate Spread			0.4093	.			0.3302	
			(0.2434)				(0.3173)	
ln(Loan Amount)	-0.0005	***	0.2511	***	-0.0005	***	0.2486	***
	(0.0001)		(0.0252)		(0.0001)		(0.0366)	
ln(App. Income)	-0.0014	***	-0.2067	***	-0.0009	***	-0.2054	***
	(0.0001)		(0.0251)		(0.0001)		(0.0365)	
Co-Applicant	-0.0053	***	-0.1044	***	-0.0049	***	-0.1059	**
	(0.0001)		(0.0243)		(0.0001)		(0.0352)	
Conv'l Loan	0.0156	***			0.0158	***		
	(0.0002)				(0.0002)			
Home Purchase	0.0114	***			0.0112	***		
	(0.0001)				(0.0001)			
Home Improve.	0.0075	***			0.0073	***		
	(0.0001)				(0.0001)			
Hispanic/Latino	0.0092	***	0.1705	***	0.0064	***	0.1702	**
	(0.0001)		(0.0424)		(0.0001)		(0.0616)	
Asian	-0.0017	***	-0.0447		-0.0034	***	-0.0456	
	(0.0002)		(0.0510)		(0.0002)		(0.0742)	
Black/Afr. Am.	0.0136	***	0.2381	***	0.0086	***	0.2396	***
	(0.0001)		(0.0395)		(0.0001)		(0.0575)	
Race not provided	0.0060	***	0.0662	*	0.0047	***	0.0640	
	(0.0001)		(0.0334)		(0.0001)		(0.0485)	
Female	0.0019	***	-0.0174		0.0018	***	-0.0180	
	(0.0001)		(0.0249)		(0.0001)		(0.0363)	
Δ ln(County Emp.)			-1.8524	**			-1.9836	*
			(0.5722)				(0.8206)	
Δ ln(House Price Idx.)			-0.3514	.			-0.3530	
			(0.1844)				(0.2678)	
Minority Pop. Pct.					0.0001	***		
					(0.0000)			
ln(HUD Median Family Income)					-0.0059	***		
					(0.0005)			

Continued on the next page.

Table 17 (continued)

	Model #1		Model #2	
	Tobit	probit	Tobit	probit
Purch. Type = 5	0.0288 *** (0.0001)		0.0282 *** (0.0001)	
Purch. Type = 6	0.0114 *** (0.0001)		0.0112 *** (0.0001)	
Purch. Type = 7	0.0186 *** (0.0001)		0.0183 *** (0.0001)	
Purch. Type = 8	0.0030 *** (0.0001)		0.0030 *** (0.0001)	
Purch. Type = 9	0.0196 *** (0.0001)		0.0192 *** (0.0001)	
Capital	0.0058 *** (0.0001)		0.0132 *** (0.0002)	
Central	0.0065 *** (0.0002)		0.0134 *** (0.0002)	
Finger Lakes	0.0058 *** (0.0001)		0.0126 *** (0.0002)	
Long Island	0.0012 *** (0.0001)		0.0083 *** (0.0002)	
Mid-Hudson	0.0004 *** (0.0001)		0.0058 *** (0.0001)	
Mohawk Valley	0.0116 *** (0.0002)		0.0182 *** (0.0002)	
North Country	0.0119 *** (0.0002)		0.0180 *** (0.0003)	
Southern	0.0099 *** (0.0002)		0.0165 *** (0.0002)	
Western	0.0073 *** (0.0001)		0.0140 *** (0.0002)	
New York County	-0.0233 *** (0.0004)		-0.0206 *** (0.0004)	
orig. 2005	0.0110 *** (0.0001)	0.1604 *** (0.0402)	0.0111 *** (0.0001)	0.1589 ** (0.0583)
orig. 2006	0.0146 *** (0.0001)	0.3100 *** (0.0498)	0.0147 *** (0.0001)	0.3096 *** (0.0723)
orig. 2007	0.0096 *** (0.0001)	0.3678 *** (0.0542)	0.0099 *** (0.0001)	0.3642 *** (0.0785)
orig. 2008	0.0041 *** (0.0001)	0.2130 *** (0.0546)	0.0049 *** (0.0001)	0.2098 ** (0.0790)
AIC	-561,338	827,003	-572,134	826,728
*** $p < 0.001$, ** $p < 0.010$, * $p < 0.050$, . $p < 0.100$				
Standard errors in parenthesis.				
<i>Data: Combined HMDA-PFF</i>				

Table 18: Distribution of Interest Rates in Pre-Foreclosure Filing Data

	total	percent
under 4.000	11,133	6.0%
4.000 to 5.999	49,876	26.8%
6.000 to 7.999	94,870	50.9%
8.000 to 9.999	21,643	11.6%
10.000 to 11.999	7,060	3.8%
12.000 to 13.999	1,430	0.8%
14.000 and up	354	0.2%
total	186,366	100.0%
<i>Data: Full PFF</i>		

Table 19: Pre-Foreclosure Filings by Predicted Interest Rate (Tobit Model #1)

	no PFF	received PFF	percent
under 4.000	18.9%	13.2%	18.4%
4.000 to 5.999	49.2%	47.0%	49.0%
6.000 to 7.999	26.0%	30.4%	26.4%
8.000 to 9.999	5.8%	9.1%	6.0%
10.000 to 11.999	0.1%	0.3%	0.1%
total	1,435,566	122,402	1,557,968
<i>Data: Combined HMDA-PFF</i>			

The simplest method is to reduce the rate spread to a binary variable (i.e. one if high-cost, zero otherwise) and employ a probit or logit model to estimate the probability that a borrower took a high-cost loan. The trouble with such a strategy is that it purges valuable information on the magnitude of the differences in rate spread among borrowers.

The alternative is to employ a Tobit model to obtain an estimate of the rate spread itself. The trouble with this strategy is that 81 percent of the loans in the combined HMDA-PFF dataset are not high-cost loans, so no value of the rate spread is reported for these loans. Therefore, instead of using the Tobit model to estimate the tail of the distribution, the Tobit model has to estimate 81 percent of the distribution.

We chose to use the Tobit model however because it provides an estimate of the rate spread which can be used as an instrument in a second-stage regression on the probability of defaulting on the home mortgage. One must use an instrument for the rate spread in the second-stage to overcome the endogeneity problem that arises when lenders charge higher interest rates to borrowers who are more likely to default.

To obtain efficient estimates of the parameters in the second-stage probability model, we used an algorithm that Adkins (2009) developed to implement Amemiya's Generalized Least Squares (AGLS).

Table 20: Pre-Foreclosure Filings by Predicted Interest Rate (Tobit Model #2)

	no PFF	received PFF	percent
under 4.000	19.4%	13.7%	19.0%
4.000 to 5.999	48.4%	45.6%	48.2%
6.000 to 7.999	26.0%	30.9%	26.4%
8.000 to 9.999	6.0%	9.6%	6.3%
10.000 to 11.999	0.1%	0.3%	0.2%
total	1,435,566	122,402	1,557,968
<i>Data: Combined HMDA-PFF</i>			

Adkins (2008) shows that the AGLS estimator yields consistent estimates of the parameters' standard errors and can be used to test the statistical significance of the parameters.

The AGLS algorithm requires estimates of the residuals from the first-stage regression, but – because the rate spread is censored at three percentage points – we could not use response residuals as we would if the first-stage regression were a standard OLS regression model. Therneau and Lumley's (2009) "survival" package for R (R Development Core Team, 2010) provides a viable alternative however. As its "survreg" function iteratively maximizes the log-likelihood function, it predicts the value of the dependent variable and calculates a correction term, called the "working residual" (Therneau, 1999), which we use in place of the response residual.

5.2 Discussion of the Regression Results

As shown in table 17, the rate spreads on owner-occupied, first-lien mortgages originated to blacks and Latinos were higher than those originated to their white and non-Latino counterparts and the differences were statistically significant, even after controlling for other variables such as income, loan amount, whether there was a co-borrower on the loan, the purpose of the loan and region of the state and year of origination.

Importantly, the racial and ethnic disparities in interest rates were large. The coefficient estimates in model #1 suggest that the interest rate on a loan originated to a black borrower was 1.36 percentage points higher than a the interest rate originated to an equivalent white borrower. Model #2 suggests a slightly smaller difference: 0.86 percentage points. Turning to Latinos, the coefficient estimates in model #1 suggest that Latinos paid 0.92 percentage points more than an equivalent non-Latino borrower, while model #2 puts the gap at 0.64 percentage points.

While this is deeply disturbing, the HMDA data omits many important variables (such as the borrower's credit score and the loan-to-value ratio), so we are reluctant to conclude that this is evidence of discrimination.

With one exception, the signs of the other coefficients in the model are not surprising. The coefficient on loan amount is the exception. It seems odd to us that borrowers who took out larger loans would pay a

lower interest rate. In the case of the HMDA data however, a large loan amount may be acting as a proxy for variables that we do not observe and thus indicate that the borrower is more creditworthy.

Before accepting our findings at face value however, one must note an important limitation of using the Tobit model to predict the rate spread: the estimates are far from perfect. By adding the average yield on a 30-year U.S. Treasury to the predicted rate spread, we can compare the Tobit models' predicted interest rates to the ones in the pre-foreclosure filing data. As tables 18, 19 and 20 show, the predicted interest rates do not have as much weight in the upper region as the interest rates in the Full PFF dataset. We believe that the predicted rate spread is correlated with the unobserved true values of the rate spread, but there is no way to check the validity of this assumption.

Turning to the second-stage model of the probability that a borrower will default, we find that the coefficient on the predicted rate spread is positive (suggesting that borrowers with higher rate spreads were more likely to default), but is only statistically significant at the 10 percent level in model #1 and is not statistically significant at all in model #2.

Both models suggest that black race and Latino ethnicity are positively correlated with the probability of default after controlling for other factors, such as income, loan amount and whether there is a co-applicant on the loan. We do not believe however that this reflects personal characteristics. Instead, we believe that the limitations of the HMDA data are causing black race and Latino ethnicity to act as a proxy for a missing variable. Given our review of the evidence of discrimination in lending practices in section 2, one possibility is that black race and Latino ethnicity are acting as a proxy for a form of discrimination that we cannot measure with the HMDA data.

Both of the models also predict that borrowers who took out larger loans were more likely to default after controlling for other factors. This finding coupled with table 2's finding that borrowers who took loans in excess of \$250,000 were much more likely to default than those who borrowed less leads us to conclude that large loan amounts are the best predictor of default.

The signs of the coefficients on other variables were in line with expectations, but it is important to note that the coefficient on applicant income was negative and statistically significant in both models. This finding helps us explain one of the puzzles that we observed in section 4: the puzzle that middle-income borrowers were more likely to default (as shown in table 3).

We could have used a quadratic term in the regression model to reproduce the result in table 3, but given the possibility that income is correlated with some of the other explanatory variables, we were reluctant to over-fit the model. Testing the statistical significance of the coefficient on a quadratic term is left to future research.

It is also interesting to note that the coefficient on the percentage change in the home price index is only statistically significant at the 10 percent level in model #1 and is not statistically significant at all in model #2. By contrast, the coefficient on the percentage change in county-level employment is statistically significant at the 5 percent level in both models. This suggests that job losses are a better predictor of default than decreasing home values.

6 CONCLUSION

Matching the New York State pre-foreclosure filing data to the HMDA originations data reveals the same racial and ethnic disparities in lending practices that other studies have found, but that finding provides very little insight into how one can reduce the rate at which borrowers default on their mortgages.

Given our finding that large loan original amounts are the best predictor that a borrower would default on his/her home mortgage, one could conclude that reducing the principal balances on home mortgages would substantially reduce the rate at which borrowers default.

Reducing principal balances may be impractical, however. In cases where borrowers have negative equity, this would require lenders to absorb potentially very large losses on their portfolio of mortgages. Secondly, an across-the-board reduction in principal balance would benefit a large number of borrowers who otherwise would not default on their mortgages.

Moreover, it is not the default per se that imposes a financial burden on lenders. It is the default that progresses to foreclosure that reduces the value of a lender's portfolio of home mortgages (when the lender is unable to recover the principal balance from the proceeds of the foreclosure sale).

We analyze the issues associated with foreclosure prevention in a second paper (Doviak and MacDonald, 2011). In that paper, we find that the original loan amount is positively correlated with the probability of progressing from default to foreclosure. Consequently, reducing principal balances is might have the desired effect of reducing losses in the mortgage industry if the modifications were well-structured, so that the balance-sheet effect of the lower probability of progressing to foreclosure offsets the losses that the lender would suffer by taking the loan to foreclosure.

Assuming that such a structure could be found, it may depend on information that the pre-foreclosure filing data does not contain, such as the borrower's income or the purpose of the loan. The HMDA data does contain this information however, so in future work, we plan to incorporate the information from the HMDA dataset into the Short PFF dataset to see how those factors affect a borrower's probability of progressing from default to foreclosure and explore other options that may help the industry reduce its losses. In that analysis, we will also attempt to quantify the savings that the industry would achieve from such modifications.

7 ACKNOWLEDGEMENTS

We would like to thank the New York State Banking Department for making the Pre-Foreclosure Filing data available to us and for supporting our research. The views expressed in this paper are our own opinions and do not necessarily reflect the opinions of the New York State Banking Department.

REFERENCES

- L. C. Adkins. "Small Sample Performance of Instrumental Variables Probit Estimators: A Monte Carlo Investigation". *Joint Statistical Meetings Proceedings, Business and Economics Statistics Section*, 4 Aug. 2008.

- L. C. Adkins. "An Instrumental Variables Probit Estimator using gretl". *Econometrics with gretl, Proceedings of the gretl Conference 2009*, pages 59–74, 28-29 May 2009.
- D. G. Bocian, K. S. Ernst, and W. Li. "Unfair Lending: The Effect of Race and Ethnicity on the Price of Subprime Mortgages", 31 May 2006. URL http://www.responsiblelending.org/mortgage-lending/research-analysis/rr011-Unfair_Lending-0506.pdf.
- C. Bromley, J. Campen, S. Nafici, A. Rust, G. Smith, K. Stein, and B. van Kerkhove. "Paying More for the American Dream: The Subprime Shakeout and Its Impact on Lower-Income and Minority Communities", Mar. 2008. URL <http://www.policyarchive.org/handle/10207/19021>.
- M. Doms, F. Furlong, and J. Krainer. "Subprime Mortgage Delinquency Rates". Working Paper 2007-33, Nov. 2007. URL <http://www.frbsf.org/publications/economics/papers/2007/wp07-33bk.pdf>.
- E. Doviak and S. MacDonald. "Who Enters the Foreclosure Process?". 24 Sept. 2011. URL <http://www.doviak.net/foreclosure/foreclosure.htm>.
- S. Erle. "Geo-Coder-US-1.00: Geocode (estimate latitude and longitude for) any US address". Comprehensive Perl Archive Network, 17 May 2005. URL <http://search.cpan.org/~sderle/Geo-Coder-US-1.00/>.
- K. S. Gerardi and P. S. Willen. "Subprime Mortgages, Foreclosures, and Urban Neighborhoods". Public Policy Discussion Papers, no. 08-6, 22 Dec. 2008. URL <http://www.bos.frb.org/economic/ppdp/2008/ppdp0806.pdf>.
- E. Laderman. "Subprime Mortgage Lending and the Capital Markets". *FRBSF Economic Letter*, 2001(38), 28 Dec. 2001. URL <http://www.frbsf.org/publications/economics/letter/2001/el2001-38.pdf>.
- D. P. Morgan, B. Iverson, and M. Botsch. "Subprime Foreclosures and the 2005 Bankruptcy Reform". *FRBNY Economic Policy Review*, 2011. Forthcoming.
- New York State Banking Department. "2009 Mortgage Foreclosure Law - Overview", 2009. URL <http://www.banking.state.ny.us/mf12009.htm>.
- R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2010. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- J. S. Rugh and D. S. Massey. "Racial Segregation and the American Foreclosure Crisis". *American Sociological Review*, 75(5):629–651, 2010.
- G. D. Squires, D. S. Hyra, and R. N. Renner. "Segregation and the Subprime Lending Crisis". EPI Briefing Paper, 4 Nov. 2009. URL <http://www.epi.org/page/-/pdf/110409-briefingpaper244.pdf>.

- T. Therneau and T. Lumley. *survival: Survival analysis, including penalised likelihood.*, 2009. URL <http://CRAN.R-project.org/package=survival>. R package version 2.35-8.
- T. M. Therneau. "A Package for Survival Analysis in S", 27 Jan. 1999.
- U.S. Census Bureau. "Census Tracts and Block Numbering Areas", 19 Apr. 2000. URL http://www.census.gov/geo/www/cen_tract.html.
- U.S. Census Bureau. "Cartographic Boundary Files", 27 June 2005. URL <http://www.census.gov/geo/www/cob/tr2000.html>.
- U.S. Census Bureau. "2006 Second Edition TIGER/Line Files", 5 Mar. 2007. URL <http://www.census.gov/geo/www/tiger/tiger2006se/tgr2006se.html>.
- U.S. Dept. of Housing and Urban Development. "Unequal Burden: Income and Racial Disparities in Subprime Lending", May 2000a. URL <http://archives.hud.gov/reports/subprime/subprime.cfm>.
- U.S. Dept. of Housing and Urban Development. "Unequal Burden in New York: Income and Racial Disparities in Subprime Lending", May 2000b. URL <http://www.huduser.org/Publications/pdf/newyork.pdf>.

Do International Capital Flows Worsen Macroeconomic Volatility in Transition Economies?

Scott W. Hegerty*

ABSTRACT

The 2008 financial crisis ultimately resulted in severe output declines in the transition economies of Central and Eastern Europe. Are capital movements responsible for the rapid growth and subsequent recession in this region? To answer this question, we examine the responses of output, consumption, and investment variability to shocks to both Foreign Direct Investment and non-FDI flows. Impulse-response and variance decomposition analysis shows that “hot” non-FDI flows contribute more to macroeconomic volatility than do more stable FDI flows, and that certain countries, particularly those with fixed exchange rates, seem to be more vulnerable to shocks than others.

INTRODUCTION

The international financial crisis that began in 2008 put paid to a number of commonly accepted economic ideas. The drawbacks of free capital movement to emerging markets would have to be considered alongside the benefits, and it was no longer a truism that macroeconomic fluctuations had been conquered in what had become known as the “Great Moderation.” Nowhere were the effects of the crisis more acutely felt than in the transition economies of Central and Eastern Europe. Those countries that maintained fixed exchange rates to the Euro, such as Estonia and Latvia, experienced Depression-level output drops after foreign capital inflows dried up. Within the CEE region, as well as within the entire European Union, only Poland escaped an economic contraction. This region thus serves as a case study of the effects of large capital flows on macroeconomic fluctuations in an important group of emerging markets.

A study of these relationships must tie together two branches of the economic literature, both of which have produced a large body of analyses but little in the way of unambiguous conclusions. The first branch studies whether Foreign Direct Investment (FDI) is inherently more stable than “hot” non-FDI flows such as portfolio investment. While Claessens *et al.* (1995) conclude that they do not, the dominant view in the literature (held by Chohan *et al.*, 1996, and Sarno and Taylor 1999), is that FDI is longer-term and less likely to be withdrawn after a shock that reduces an investment’s expected profitability. The second branch of the literature analyzes the effects of access to foreign capital on a country’s macroeconomic volatility, but as Kose *et al.* (2006) note in a review article, this relationship is ambiguous. In theory, such capital should allow a country to smooth its consumption in face of economic shocks. At the same time,

* Department of Economics and Finance, Wehle School of Business, Canisius College, Buffalo, NY 14208. This research was partially funded by a Canisius College Summer Grant.

financial openness exposes economies to instability elsewhere. Evidence of the former view is provided by for various countries by Karras (2006) and Ahmed and Suardi (2009). Razin and Rose (1994), on the other hand, find no link between trade and financial openness and the volatility of output, consumption, and investment for a set of 138 countries. Kose *et al.* (2003) uncover a “hump-shaped” pattern whereby volatility transmission is more likely for less-developed countries, and volatility smoothing is more common for developed countries.

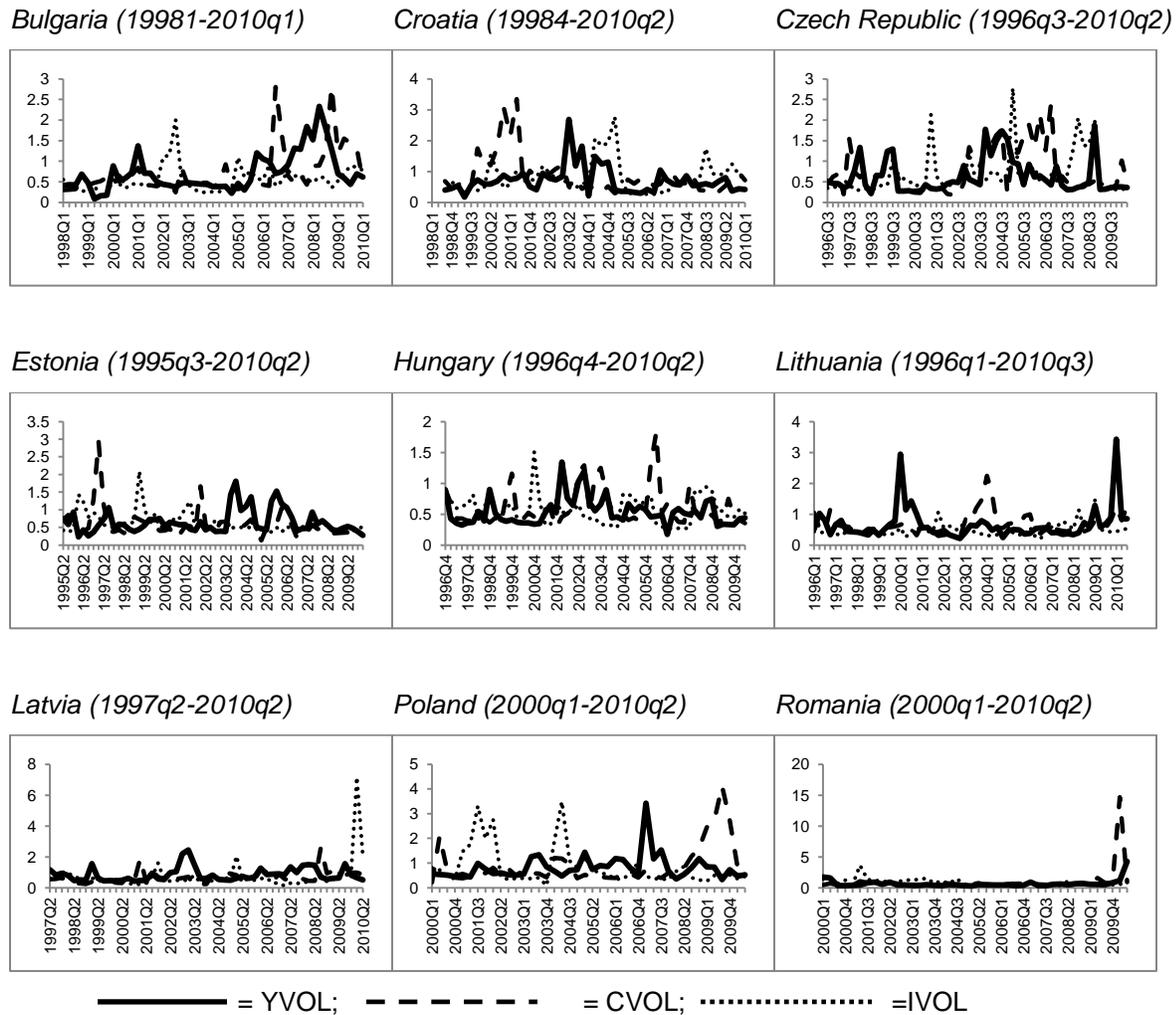
While trade and financial openness (usually proxied by *de jure* measures such as liberalization indices, or *de facto* ones such as stocks of foreign investment) are often the key variable in these analyses, relatively few studies look at capital flows directly. One exception is the study by Alper (2002), who examines the cross-correlations between real output volatility and net capital inflows for Mexico and Turkey, but finds significant variables only for Mexico. Other variables have been examined as sources of macroeconomic volatility, and these have often been shown to play more of a role than does trade or financial openness. Examples include exchange-rate flexibility and shocks to the money supply and technology (Karras and Song, 1996); country size (Karras, 2006); government size and fiscal policy (Furceri, 2007); and external shocks, usually proxied by terms of trade volatility (Hirata *et al.*, 2007 and Kim, 2007). The effects of financial openness or capital flows on macroeconomic volatility specifically in transition economies have received relatively little attention in the literature, however. This may be due to the short time series that are available for the post-communist period. For the same reason, earlier studies were unable to include many CEE countries at all. Ramey and Ramey (1995), for example, include no countries in the region save Yugoslavia.

This study examines the impact of capital flows on output, consumption, and investment volatility in a set of CEE countries. These volatility series, beginning in the mid-1990s, are shown in Figure 1. Clearly, each type of variability varies throughout the period, with country- and region-specific patterns. The Balkan countries, particularly Romania and Bulgaria, register notable spikes in the around the 2008 crisis, although macroeconomic volatility was present throughout the sample period. Likewise, other countries experienced periods of high volatility, particularly during the financial crisis. Latvian investment, Polish consumption, and Lithuanian output all stand out for sharp increases. Yet, unlike its two Baltic neighbors, Estonia seems to have seen a reduction in volatility during the crisis period. This might be expected, since Estonia was the first former Soviet republic with the requisite macroeconomic performance to be invited to join the Euro. It is clear that each country and GDP component follows its own path.

Do these countries also respond differently to FDI and non-FDI investment? This study answers this question by entering each flow separately into each nation’s estimation. By applying Vector Autoregressive (VAR) methods to data for nine transition economies, we are able to see how shocks to both flows affect macroeconomic volatility in each country. Impulse-response and variance decomposition analysis shows that “hot money” does indeed contribute more to macroeconomic volatility in the region than does FDI. Our country-specific results indicate that certain more stable countries, such as Estonia, are more insulated from these effects, while capital flows introduce more instability in the Balkans. This

paper proceeds as follows: Section II outlines the methodology and Section III provides the empirical results. Section IV concludes.

Figure 1. Macroeconomic Volatility Series, Nine Transition Economies



METHODOLOGY

This study examines the effects of increases (or decreases) in nine CEE countries’ capital flows on their volatility of output, consumption, and investment. To do this, quarterly data from the International Financial Statistics of the International Monetary Fund are used to create relevant volatility terms and an appropriate set of control variables. The length of each final dataset varies from country to country, but each are given in Figure 1. The variables, which are chosen based on the literature outlined above, are as follows:

YVOL = volatility of real GDP (nominal deflated by GDP deflator)

CVOL = volatility of real consumption (nominal, deflated by Consumer Price Index)

IVOL = volatility of real gross fixed capital formation plus changes in inventories (nominal deflated by PPI)

GVOL = volatility of real Government expenditure (nominal, deflated by GDP deflator)

REERVOL = volatility of the real effective exchange rate

REALR = real interest rate (nominal money-market rate minus inflation rate)

YGROWTH = percentage change in real GDP (over four quarters previous)

FDI = net foreign direct investment (inflows minus outflows), as a share of GDP

NONFDI = net portfolio plus net other investment, as a share of GDP.

When necessary, all variables are deseasonalized using the Census X-12 procedure. Each volatility series is created by capturing the standard errors of a rolling AR(1) estimation over eight quarters, using first-differenced log data. While the literature has failed to point toward a single best proxy for volatility, this method was chosen over an alternative (a moving standard deviation of log changes in each relevant series) because it shows less persistence after a shock. The three main macroeconomic volatility terms are included together in the VAR, since they might have a significant influence on one another. Besides FDI and non-FDI flows, variability in Government spending and the exchange rate are the focus of this study, the former to test whether fiscal policy is an effective method of stabilizing these economies, and the latter to analyze the role of external volatility on these domestic economies. Finally, growth and the real interest rate are included, since they might influence both capital flows and macroeconomic volatility.

All variables are tested for stationarity using the Phillips-Perron test, which is similar to the Augmented Dickey-Fuller test, but uses Newey-West (1987) standard errors to control for autocorrelation. Taking first differences for any I(1) variables, each variable is entered in to a VAR. This method is not only an excellent way to deal with endogenous variables; it also captures the short-term effects that sudden shocks to net capital flows might bring about. In this VAR, one lag is chosen in all cases, based on both the Schwarz criterion and the desire to reduce the number of parameters in the model. Next, impulse-response functions (IRFs) and forecast error variance decompositions (FEVDs) are obtained for each country's vector. Because these variables have been shown to affect one another—both capital flows and volatility have been shown to influence growth, for example—there is no particular “best” ordering for the variables. For this reason, the Generalized VAR approach of Pesaran and Shin (1998) is used rather than the more traditional orthogonalized method of Sims (1980). Estimations obtained using this newer approach is insensitive to the ordering of the variables. These results are as follows.

RESULTS

The results of the Phillips-Perron stationarity test, for both level and first-differenced variables, are available upon request. Capital flows are generally stationary, but other variables differ from country to country. We use these results to make sure that only stationary variables are entered into the VAR. With

the exception of capital flows, a time series is differenced if the PP test suggests that it is nonstationary in levels.

Next, we generate Generalized Impulse-Response Functions for each of the nine countries' VARs. A summary of the original functions is provided in Table 1. In addition, those GIRFs that are shown to be significant, along with standard-error bands, are shown in Figure 2. We can compare the results in three ways: By type of capital flow, by type of macroeconomic volatility, and by country or region.

Here, we find that non-FDI flows have more of an effect on all types of macroeconomic volatility in these countries. This confirms the "hot money" hypothesis in this case. At the same time, no specific pattern can be determined for responses by the three types of variability. Output, consumption (which is expected to be more easily smoothed by households) and investment (which is usually the most volatile) each show a unique response. For example, non-FDI flows to Bulgaria seem to smooth output fluctuations, but FDI has little effect. For consumption, only FDI has an impact—but it leads to increased volatility, suggesting either transmission or the inability of the domestic market to absorb these inflows.

These differences persist, especially for those countries that are more integrated with Western Europe. The Czech Republic's FDI inflows reduce output variability, while non-FDI flows have an effect only on investment variability. This suggests that the country's strong connections to German and other industry might lead to increased stability in the domestic economy. In Lithuania, which is somewhat less integrated, we find that "hot" money has led to an increase in this output variability. On the other hand, FDI seems to increase output volatility in Poland, which, like the Czech Republic, has seen a large share of German direct investment.

Of the four fixed-rate regimes (the three Baltic nations plus Bulgaria), two have seen clear increases in domestic macroeconomic instability as a result of increased capital inflows. One may argue that the two countries where the responses are statistically insignificant—particularly Estonia, whose economic success allowed it to join the Euro in 2011—are the more advanced countries in the group. Therefore, our results suggest that maintaining a fixed rate, which removes the main adjustment mechanism to deal with capital inflows, might make transition economies more vulnerable to volatility increases. These results are more pronounced for the less developed economies in the group. A floating rate seems to help here: The results for Romania suggest that capital flows can reduce volatility. A positive shock to FDI results in reduced investment volatility, while non-FDI leads to a reduction in consumption variability that is significant at a level just higher than 10 percent.

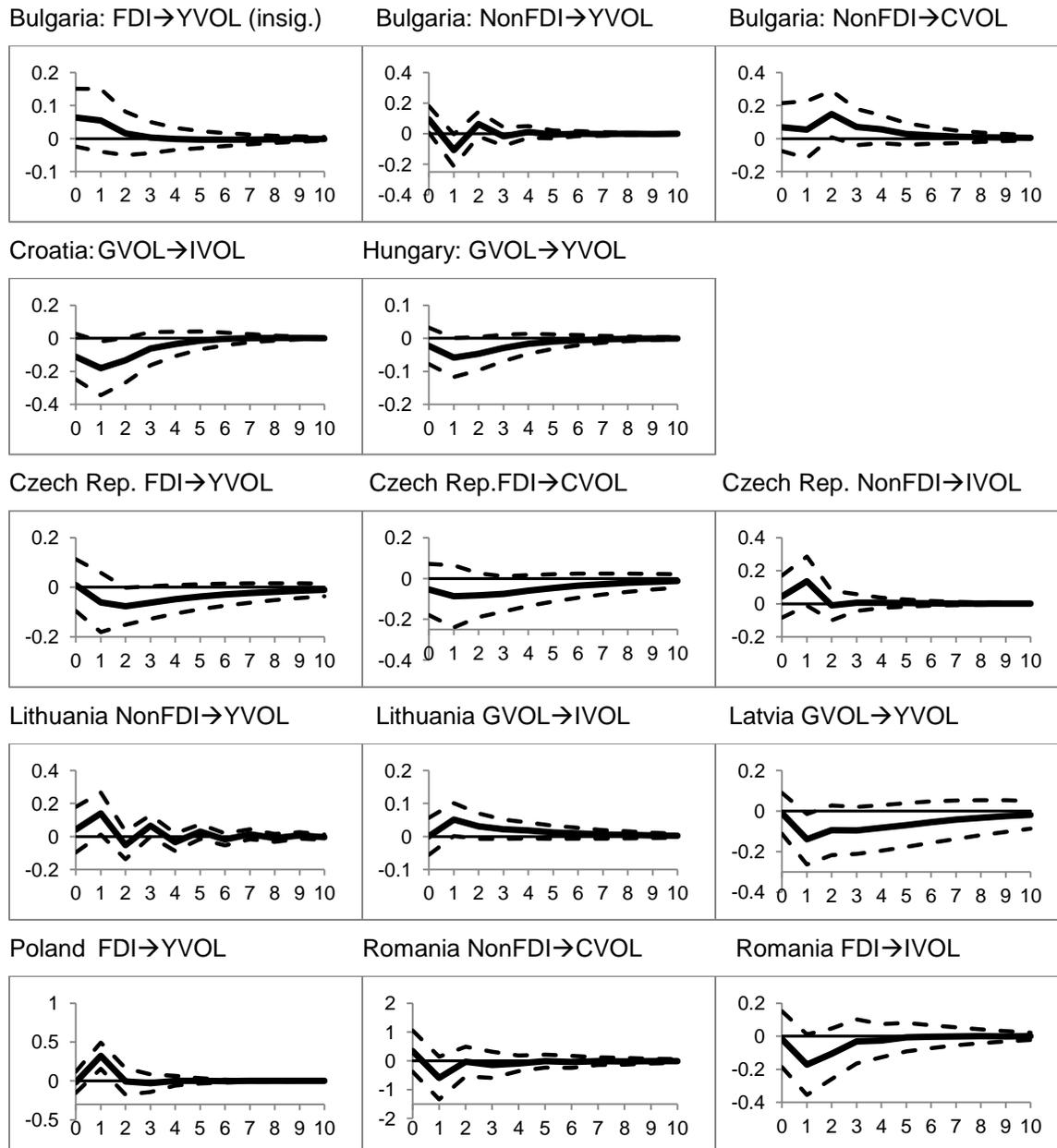
Table 1: Summary of Generalized Impulse-Response Functions

<i>Impulses to →</i>					<i>Impulses to →</i>				
<i>Bulgaria</i>	FDI	NONFDI	GVOL	REERVOL	<i>Lithuania</i>	FDI	NONFDI	GVOL	REERVOL
YVOL	Insig.	Neg.	Insig.	Insig.	YVOL	Insig.	Pos.	Insig.	Insig.
CVOL	Insig.	Pos.	Insig.	Insig.	CVOL	Insig.	Insig.	Insig.	Insig.
IVOL	Insig.	Insig.	Insig.	Insig.	IVOL	Insig.	Insig.	Pos.	Insig.
<i>Croatia</i>	FDI	NONFDI	GVOL	REERVOL	<i>Latvia</i>	FDI	NONFDI	GVOL	REERVOL
YVOL	Insig.	Insig.	Insig.	Insig.	YVOL	Insig.	Insig.	Neg.	Insig.
CVOL	Insig.	Insig.	Insig.	Insig.	CVOL	Insig.	Insig.	Insig.	Insig.
IVOL	Insig.	Insig.	Neg.	Insig.	IVOL	Insig.	Insig.	Insig.	Insig.
<i>Czech R.</i>	FDI	NONFDI	GVOL	REERVOL	<i>Poland</i>	FDI	NONFDI	GVOL	REERVOL
YVOL	Neg.	Insig.	Insig.	Insig.	YVOL	Pos.	Insig.	Insig.	Insig.
CVOL	Neg.	Insig.	Insig.	Insig.	CVOL	Insig.	Insig.	Insig.	Insig.
IVOL	Insig.	Pos.	Insig.	Insig.	IVOL	Insig.	Insig.	Insig.	Insig.
<i>Estonia</i>	FDI	NONFDI	GVOL	REERVOL	<i>Romania</i>	FDI	NONFDI	GVOL	REERVOL
YVOL	Insig.	Insig.	Insig.	Insig.	YVOL	Insig.	Insig.	Insig.	Insig.
CVOL	Insig.	Insig.	Insig.	Insig.	CVOL	Insig.	Neg.	Insig.	Insig.
IVOL	Insig.	Insig.	Insig.	Insig.	IVOL	Neg.	Insig.	Insig.	Insig.
<i>Hungary</i>	FDI	NONFDI	GVOL	REERVOL					
YVOL	Insig.	Insig.	Neg.	Insig.					
CVOL	Insig.	Insig.	Insig.	Insig.					
IVOL	Insig.	Insig.	Insig.	Insig.					

What is the role of the other types of volatility, both external and internal, in influencing this macroeconomic variability? Figure 1 shows the responses to variability in Government spending and in the real effective exchange rate. We find that *GVOL* increases investment volatility in Lithuania and Croatia, but reduces output variability in Latvia and Hungary. It is important to recall that Latvia maintained its Euro peg during the crisis by implementing an “internal devaluation,” whereby the government drastically contracted the economy so that prices would fall instead of the lats. Hungary, as well, received an IMF bailout, and while the forint is more flexible, we do see the two hardest-hit countries reducing output volatility through fiscal changes. When we look at the volatility of the real effective

exchange rate, we surprisingly find little evidence of any significant response to external shocks, in any country.

Figure 2: Selected Significant Generalized Impulse-Response Functions, with ± 2 Standard Error Bands



Our Forecast Error Variance Decompositions are available upon request, for all variables in each VAR, at one- and four-quarter horizons. These tend to confirm the results of our IRFs. In results that are similar to what was described above, shocks to non-FDI flows play an important role in determining

volatility in the region. In Bulgaria, this is true for consumption and output. In Latvia, Investment flows are influenced in a way that was not depicted in the IRFs. For the other countries, the outsize role of “hot” money is demonstrated in certain cases that match the IRF results. Examples include investment volatility in the Czech Republic and Poland and consumption volatility in Romania. Most importantly, while both types of investment make large contributions to the forecast error of Poland’s real GDP volatility, FDI inflows play the largest role of any flow or country studied here.

CONCLUSION

The transition economies of Central and Eastern Europe have, in the past decade, seen the potential downsides to both inflows and outflows of capital. Whereas before the 2008 financial crisis, excessive inflows had led to credit bubbles and high inflation, the “sudden stop” that followed helped usher in a severe economic contraction. This study looks at the role of capital inflows on macroeconomic volatility on nine economies in the region, from the mid-1990s to 2010. Modeling these types of volatility along with capital flows and other related variables in a VAR framework, we arrive at two main conclusions. First, “hot” non-FDI does seem to have more of an impact on output, consumption, and investment variability in the region than does Foreign Direct Investment. This confirms the idea that more volatile capital flows do indeed serve to destabilize these economies. At the same time, external volatility plays less of a role here than has been shown in certain previous studies. Secondly, each country responds to shocks in capital flows in its own way, which yields certain clear patterns. The Czech Republic, which has attracted a large amount of German manufacturing, has seen FDI reduce its level of output variability. The countries that have fixed their exchange rates to the Euro as a precursor to joining the common currency appear to have exposed themselves to macroeconomic instability as a result.

REFERENCES

- Ahmed, A.D. and S. Suardi. 2009. “Macroeconomic Volatility, Trade and Financial Liberalization in Africa.” *World Development*, 37(10): 1623-1636.
- Alper, C.E. 2002. “Business Cycles, Excess Volatility and Capital Flows: Evidence from Mexico and Turkey.” *Russian and East European Finance and Trade*, 38(4): 22-54.
- Chuhan, P, G. Perez-Quiros and H. Popper. 1996. “International Capital Flows: Do Short-Term Investment and Direct Investment Differ?” World Bank Policy Research Working Paper 1669.
- Claessens, S, M. P. Dooley and A. Warner. 1995. “Portfolio Capital Flows: Hot or Cold?” *The World Bank Economic Review*, 9(1): 153-174.
- Furceri, D. 2007. “Is Government Expenditure Volatility Harmful for Growth? A Cross-Country Analysis.” *Fiscal Studies* 28(1): 103-120.
- Hirata, H. S. H. Kim and M. A. Kose. 2007. “Sources of Fluctuations: The Case of MENA.” *Emerging Markets Finance and Trade*, 43(1): 5-34.

- Karras, G. 2006. "Trade Openness, Economic Size, and Macroeconomic Volatility: Theory and Empirical Evidence." *Journal of Economic Integration*, 21(2): 254-72.
- Kim, S. Y. 2007. "Openness, External Risk, and Volatility: Implications for the Compensation Hypothesis." *International Organization*, 61(1): 181-216.
- Kose, M. A.; Prasad, E.S.; K. Rogoff and S.-J. Wei. 2006. "Financial Globalization: A Reappraisal." *IMF Working Papers*: 06/189.
- Kose, M.A.; E.S. Prasad and M.E. Terrones. 2003. "Financial Integration and Macroeconomic Volatility." *IMF Staff Papers*, Special Issue, 50: 119-42.
- Newey, W. K. and K.D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55(3) 703-708.
- Pesaran, M. H., and Y. Shin. 1998. "Generalised Impulse Response Analysis in Linear Multivariate Models." *Economics Letters*, 58: 17-29.
- Phillips, P. and P. Perron. 1988. "Testing for a Unit Root in Time Series Regression." *Biometrika*, 75: 335-346.
- Ramey, G. and V. A. Ramey. 1995. "Cross-Country Evidence on the Link Between Volatility and Growth." *American Economic Review*, 85(5): 1138-1151.
- Razin, A. and A. K. Rose. 1994. "Business-Cycle Volatility and Openness: An Exploratory Cross-Sectional Analysis." In *Capital Mobility: The Impact on Consumption, Investment and Growth*, Cambridge University Press, 48-75.
- Sarno, L. and M. P. Taylor. 1999. "Hot Money, Accounting Labels and the Permanence of Capital Flows to Developing Countries: An Empirical Investigation." *Journal of Development Economics*, 59: 337-364.
- Sims, C. A. 1980. "Macroeconomics and Reality." *Econometrica* 48, 1-48.

Can Crowd Out Hamper Stimulus Programs In Recessions?

John J. Heim*

ABSTRACT

In well controlled statistical tests, crowd out was found related to government deficits financed by borrowing. Roughly equal effects were found for both recession and non-recession periods. Tax cut deficits were found more detrimental than spending deficits. Private borrowing was found to systematically decline with the growth of government deficits, and explained most variation in consumer and investment spending also related to deficits. Crowd out may be avoided if foreign borrowing increases loanable funds available, or if M2 money increases prior to the deficit. This study provides first time empirical evidence, using standard demand driven stimulus models, that crowd out occurs in recessions and non-recession periods in response to deficits, fully offsetting stimulus effects. Results are robust to differences in time period tested. They offer a plausible explanation for the lackluster results of recent U.S. government stimulus programs in offsetting the 2008 recession.

INTRODUCTION

When government finances deficits by borrowing from the savings pool, the reduction in loanable funds available to consumers and businesses is "crowd out". Borrowing by both consumers and businesses is extensively used even in recessions to supplement the purchasing power of income, for example, when buying a house or new machinery. Government borrowing to finance deficits may crowd out private borrowing, and hence, spending, offsetting some or all of the deficit's stimulus effects. Whether it actually does or not is an empirical issue, examined in this paper.

The paper tests whether consumer or business borrowing is negatively impacted by deficits, in either recessions or non-recession periods. It also tests if any declines in spending are systematically related to declines in borrowing, as theorized in the crowd out model. U.S. data for the period 1960-2000 are tested. Findings are tested for robustness by dropping the first or last ten year period from the sample, and then retesting.

THE NO - CROWD OUT MODEL

Standard simple models of the economy do not allow for borrowing - related crowd out. In such models the impact of taxes and government spending are derived using the GDP identity:

$$\text{GDP} = Y = C + I + G + (X-M) \quad (1)$$

A simple consumption function might be given as a linear function of disposable income (Y-T)

$$C = \beta(Y-T) \quad (2)$$

*Lecturer/Clinical Professor of Economics, Rensselaer Polytechnic Institute, Troy, NY 12180

substituting C into (1) gives

$$Y = \frac{1}{1-\beta} [-\beta T + I + G + X - M] \quad (3)$$

The clear expectation of standard model demand theory is that tax changes in are expected to be negatively related to the GDP, with a multiplier effect $-\beta/(1-\beta)$. Changes in government spending and net exports are related to GDP in the positive direction, with a multiplier effect $1/(1-\beta)$.

THE CROWD OUT MODEL

However, to test the hypothesis that savings otherwise used to finance consumer credit is diverted to finance government deficits, the consumption function must be modified to add the crowd out - causing factor, the deficit (T-G), where (T-G) = taxes minus government spending:

$$C = \beta (Y-T) + \lambda(T-G) \quad (4)$$

where lambda (λ) represents the marginal effect of deficit spending on consumer demand. With this function, the model becomes

$$\begin{aligned} \text{GDP} = Y &= \beta (Y-T) + \lambda(T-G) + G + I + (X-M) \\ &= \frac{1}{1-\beta} [(-\beta + \lambda) T + (1-\lambda) G + I + (X-M)] \end{aligned} \quad (5)$$

The impact of a change in T or G on the GDP depends on λ as well as β . The tax multiplier, is now $(-\beta + \lambda)/(1-\beta)$. The spending multiplier, is now $(1-\lambda)/(1-\beta)$. Both T and G marginal effects on the GDP will be smaller (in absolute terms) than they would have been without crowd out effects.

If crowd out has different effects in recession (Rec) and non-recession periods (NonRec), the formulation becomes

$$\begin{aligned} \text{GDP} = Y &= \beta (Y-T) + \lambda_{\text{Rec}}(T-G) + \lambda_{\text{NonRec}}(T-G) + G + I + (X-M) \\ &= \frac{1}{1-\beta} [(-\beta + \lambda_{\text{Rec}})T + (-\beta + \lambda_{\text{NonRec}})T + (1-\lambda_{\text{Rec}}) G + (1-\lambda_{\text{NonRec}}) G + I + (X-M)] \end{aligned} \quad (6)$$

We can see the impact of a change in T or G on the GDP depends on λ_{Rec} or λ_{NonRec} as well as β . The tax multiplier, is now $(-\beta + \lambda_{\text{Rec}})/(1-\beta)$ in recessions or $(-\beta + \lambda_{\text{NonRec}})/(1-\beta)$ in non-recessions. If crowd out is less in recessions, the tax cut multiplier effects will be larger than in non-recessions. The spending multiplier, is now $(1-\lambda_{\text{Rec}})/(1-\beta)$ or $(1-\lambda_{\text{NonRec}})/(1-\beta)$ and if crowd out is less in recessions, the spending multiplier will be larger in recessions than in non-recessions.

We can expand this model to include effects of crowd out on investment spending. Assume a simple investment model in which investment is determined by real interest rates (r) and access to credit, which varies with the government deficit (T-G).

$$I = \gamma(T-G) - \theta r \quad (7)$$

where gamma (γ) indicates the marginal effect of crowd out (the government deficit) on investment spending, and (θ) represents the marginal effect of real interest rates (r).

If we replace investment in the GDP identity with its hypothesized determinants, we obtain a typical “IS” equation:

$$GDP = Y = [1/1-\beta] [(-\beta + \lambda + \gamma) T + (1-\lambda-\gamma) G - \theta r + (X-M)] \quad (8)$$

In this IS equation, the normal stimulating impact of tax cuts on the GDP ($-\beta$) is offset by the effects of deficit – induced changes in credit available to consumers and investors ($\lambda + \gamma$). Tax stimulus effects may switch from negative to positive if the crowd out effects ($\lambda + \gamma$) are larger than the disposable income effect ($-\beta$). The normal stimulating effect of government spending is reduced from (1) to $(1-\lambda-\gamma)$, and stimulus effects are either reduced or become negative. The net exports multiplier effect stays the same, now becoming an even stronger stimulus relative to government spending or tax cuts. Results are shown in Table 1. Crowd out reduces the stimulus of spending deficits less than tax cut deficits, because the spending stimulus effect (1) to be offset by crowd out is larger than the tax stimulus effect ($-\beta$). Results are shown in Table 1.

Crowd out may or may not be a problem in recessions. It can be argued consumers and businesses borrow less in recessions, leaving savings available to finance new government deficits without causing crowd out. However, national savings may also drop in recessions due to falling incomes. If savings decline as much or more than private borrowing demand, new deficits will still cause new crowd out. Hence, arguments for and against crowd out in recessions can be made theoretically.

If crowd out effects are different in recessions than in non-recessions, the investment and IS functions change as follows:

$$I = -\theta r + Y_{Rec} (T-G) + Y_{NonRec} (T-G) \quad (9)$$

$$Y = [1/1-\beta] [(-\beta + \lambda_{Rec} + Y_{Rec}) T \text{ or } (-\beta + \lambda_{NonRec} + Y_{NonRec}) T + (1-\lambda_{Rec}-Y_{Rec}) G \text{ or } (1-\lambda_{NonRec}-Y_{NonRec}) G - \theta r + (X-M)] \quad (10)$$

Hence, the stimulus effect of tax cuts ($-\beta$) is offset by either the recession effect of crowd out ($\lambda_{Rec} + Y_{Rec}$) or the non-recession effect ($\lambda_{NonRec} + Y_{NonRec}$). The stimulus effect of government spending (+1) is offset by either the recession effect of crowd out ($-\lambda_{Rec} - Y_{Rec}$) or the non-recession effect ($-\lambda_{NonRec} - Y_{NonRec}$). These results are also summarized in Table 1.

The model we shall test later in this paper is a slightly different form of the model shown above. The model above was based on the GDP identity

$$Y = C + I + G + (X-M) \quad (11)$$

But we can just as accurately write

$$Y = C_D + I_D + G_D + X \quad (\text{where subscript D denotes domestically produced goods}) \quad (12)$$

This is an important distinction in calculating multipliers because only spending on domestically produced consumer or investment goods generates the multiplier effect on the GDP. Hence, the last formulation of the GDP identity may be the better form to use when calculating IS curve parameter estimates, since multiplier effects inherent in the coefficients are more correctly estimated. (We abstract from effects on exports of growth in import demand).

Because the data available does not allow separation of government purchases into domestic produced goods and imports, the form of the model we test is:

$$Y = C_D + I_D + G + X \quad (13)$$

Table 1: Effects of Crowd Out on Taxes and Government Spending Stimulus

	Without Crowd Out	With Crowd Out		Without Crowd Out	With Crowd Out
Tax coefficient	$(-\beta)$	$-\beta + (\lambda + \gamma)$	Spending Coefficient.	1	$1 - (\lambda + \gamma)$
	$(-\beta)$	$-\beta + (\lambda + \gamma)_{Rec}$		1	$1 - (\lambda + \gamma)_{Rec}$
	$(-\beta)$	$-\beta + (\lambda + \gamma)_{NonRec}$		1	$1 - (\lambda + \gamma)_{NonRec}$
Tax Multiplier (Average-All Per.)	$(-\beta)$ $(1 - \beta)$	$-\beta + (\lambda + \gamma)$ $(1 - \beta)$	Spending Multiplier	<u>1</u> $(1 - \beta)$	<u>$1 - (\lambda + \gamma)$</u> $(1 - \beta)$
Tax Multiplier (Recession Period)	$(-\beta)$ $(1 - \beta)$	$-\beta + (\lambda + \gamma)_{Rec}$ $(1 - \beta)$	Spending Multiplier	<u>1</u> $(1 - \beta)$	<u>$1 - (\lambda + \gamma)_{Rec}$</u> $(1 - \beta)$
Tax Multiplier (Non-Recession)	$(-\beta)$ $(1 - \beta)$	$-\beta + (\lambda + \gamma)_{NonRec}$ $(1 - \beta)$	Spending Multiplier	<u>1</u> $(1 - \beta)$	<u>$1 - (\lambda + \gamma)_{NonRec}$</u> $(1 - \beta)$

OTHER STUDIES

An extensive review of literature addressing the crowd out topic is available in Rensselaer Polytechnic Institute's Department of Economics Working paper #1103, available at the department's website.

METHODOLOGY

1960 - 2000 data on the determinants of consumption and investment spending were taken from the *Economic Report of the President 2002 and 2010. Flow of Funds Accounts* of the Federal Reserve were used to obtain data on consumer and business debt, changes in which were taken as measures of net consumer or business borrowing. The specific variables are identified in later sections.

Two -stage least squares regression was used since both consumption and investment are driven in part by income related variables (disposable income or the accelerator), and therefore 2SLS was needed to avoid issues of simultaneity. The remaining variables in the consumption and investment equations are lagged, or considered exogenous (e.g. the prime interest rate is rigidly set to reflect the federal funds rate, which is exogenously determined by the Fed; depreciation allowances are determined by prior year investment levels, etc.). They were used as 1st Stage regressors. Variables were tested in first differences instead of levels to address nonstationarity, serial correlation and multicollinearity issues commonly found

in time series data. Newey West corrections to standard errors were used to avoid heteroskedasticity problems. Durbin - Watson measures of serial correlation are shown with each regression. The 2SLS form of instrumental variables and the D.W. test for serial correlation are considered the most appropriate for use with time series data (Griffiths, Hill, Lim, 2011).

It is difficult to separate consumer imports out of total imports in the *Economic Report of the President*. The Bureau of Economic Analysis (BEA) has confirmed it does not categorize import and export data into same “C” and “I” and “G” categories used elsewhere in the national GDP accounts. Absent official determinations by BEA, economists must make their own evaluations of how to divide the data. For example, it is not clear from Table 104 in the *Economic Report of the President* how much of the value of motor vehicle imports or petroleum imports should be treated as inventory investment vs. sales to consumers. Data on imported services (Table B-106) does not distinguish between imports of services by businesses and consumers, though one might suspect the former dominate. Nor do the services data extend back beyond 1974. Hence, no deduction from total imports for business services imports could be made in calculating consumer imports.

Following Heim (2010), we take as our definition of consumer imports all imports except imports of capital goods and industrial supplies and materials. The theory behind this choice was that the best definition of “consumer” imports was the one whose variation was best explained (highest R^2) by the variables theoretically thought to drive demand for consumer imports. Other definitions of consumer imports, did not explain consumer behavior as well.

To obtain separate deficit, tax or government spending variables for recession periods, they are multiplied by a dummy variable taking the value (1) when there is a depression at some time during the data year, and (0) in non-recession years. For non-recession years, the dummy variable is reversed. National Bureau of Economic Research estimates (NBER 2009) were used to define recession years.

VARIABLES INCLUDED IN CONSUMPTION AND INVESTMENT MODELS

Theory suggests a wide range of variables are determinants of consumer demand. Individual studies have provided some empirical support for many variables, though not always controlling adequately for the rest. The consumption functions tested in this paper control for an extensive list of possible factors that might influence consumption, including crowd out, thereby helping ensure the correctness of the crowd out results. Lagged or average values are used with some of these variables, reflecting the findings of earlier studies. The variables used include:

C_T	= real consumer goods and services
$(Y-T)$	= real disposable income
$(T-G)$	= real government deficit, defined as one variable
PR	= real prime interest rate
DJ_{-2}	= a measure of wealth (Dow Jones Composite Index), lagged two years
XR_{AV}	= exchange rate average for current and past three years

- POP₁₆ = ratio of young (16 -24 year olds) to old (over 65) in population
 POP = population size
 ICC₋₁ = Index of Consumer Confidence (Conference Board), lagged one year
 M2_{AV} = real M2 money supply, average of past three years
 (C_{BOR}) = Net annual consumer borrowing = change in consumer debt (ΔC_{DEBT}) that period.

Theory also suggests many variables are determinants of investment demand. Here again, individual studies have provided some empirical support for many different variables, though results are not always obtained controlling adequately for other possible factors. The variables used here include:

- I_T = total spending on investment goods, both domestically produced and imported
 I_M = Spending on imported investment goods
 I_D = Spending on domestically produced investment goods
 ACC = a Samuelson accelerator variable measuring the economy's growth rate (Δ Real GDP)
 DEP = real business depreciation allowances
 CAP₋₁ = industrial capacity utilization, lagged one period
 r₋₂ = real Prime interest rate, lagged two periods
 DJ₋₂ = Dow Jones Composite Average, a proxy for Tobin's q
 PROF₋₂ = real corporate profits, lagged two periods
 B_{BOR} = Net annual business borrowing = change in business debt (ΔB_{DEBT}) that period.
 Other variables in the investment model are as defined in the consumption function.

TEST RESULTS, FINDINGS: "RECESSION/NO RECESSION" VERSUS "NO CROWD OUT" MODELS

Below, consumption and investment models are tested for separate recession and non - recession crowd out effects. The domestic consumption and investment models tested are the same as previously used. Regression results for the crowd out model are:

Domestically Produced Consumer Goods

$$\Delta C_D = .36\Delta(Y-T_G) + .20\Delta(T-G)_{Rec} + .35\Delta(T-G)_{NonRec} - 4.07\Delta PR + .22\Delta DJ_{-2} + 1.24\Delta XR_{AV} - 410.32\Delta POP_{16} + .01\Delta POP \quad (21)$$

(t =) (6.8) (2.2) (2.2) (-1.5) (1.1) (1.5) (-2.4) (2.3)

+ .13 Δ ICC₋₁ + 40.02 Δ M2_{AV} R²=86.4%
(0.4) (4.2) D.W.= 2.1

Domestically Produced Investment Goods:

$$\Delta I_D = +.50\Delta(T-G)_{Rec} + .52\Delta(T-G)_{NonRec} + .23\Delta ACC + .13\Delta DEP - .35\Delta CAP_{-1} - 8.15\Delta r_{-2} - .27\Delta DJ_{-2} + .44\Delta PROF_{-2} + 5.58\Delta XR_{AV} \quad (22)$$

(t =) (8.3) (3.5) (9.0) (0.5) (-0.3) (-7.0) (-1.4) (4.0) (5.7)

+ .008 Δ POP R²=89.8%
(3.7) DW =2.3

Using the parameter estimates from these domestic consumption (C_D) and investment (I_D) equations, the IS curve parameters are predicted to be:

Predicted IS Curve (With A Separate Crowd Out Variable (T-G) Included For Recessions/Non-recessions)

$$\Delta Y = +.53\Delta T_{Rec} + .80\Delta T_{NonRec} + .47\Delta G_{Rec} + .20\Delta G_{NonRec} + 1.56\Delta X - 6.35PR - .08\Delta DJ_{-2} + 10.64\Delta XR_{AV0123} - 640.10\Delta POP_{16} + .03\Delta POP + .20\Delta ICC_1 + 62.43\Delta M2 + .36\Delta ACC + .20\Delta DEP - .55\Delta CAP_{-1} - 12.71r_{-2} + .69\Delta PROF_{-2} \quad (23)$$

Actual IS Curve Test Results (With Separate (T) And (G) Variables For Recession/Non-Recessions)

$$\Delta Y = +.87\Delta T_{\text{Rec}} + .60\Delta T_{\text{NonRec}} - .65\Delta G_{\text{Rec}} - .23\Delta G_{\text{NonRec}} + .63\Delta X - 8.00\Delta PR + .24\Delta DJ_{-2} + 4.97XR_{\text{AV}} + 445.43\Delta POP_{16} \quad (24)$$

(t=) **(5.3)** **(2.3)** **(-1.1)** **(-0.6)** **(2.0)** (2.2) (0.6) (2.6) (1.3)

$$+ .05\Delta POP + 1.59\Delta ICC_{-1} + 44.51\Delta M2 + .59\Delta ACC + .68\Delta DEP + 8.36\Delta CAP_{-1} - .40r_{-2} + .22\Delta PROF_{-2} \quad R^2=97.8\%$$

(5.6) (3.5) (2.3) (10.4) (0.7) (2.3) (-0.1) (0.8) DW=2.5

The results indicate that

- Adding crowd out to the consumption model increases explained variance from 81.3% to 86.4%, slightly more than the average effect model (86.0%). For investment, adding crowd out increased explained variance about the same as the average crowd out model: from 75% to 89.8%, slightly less than the average model (90.0%). Though there are differences in specific recession/non-recession coefficients for taxes and spending, they are each within the confidence intervals of the other, suggesting they may be the same. This indicates dividing crowd out into recession and non-recession variables adds little information not already available in the average crowd out model. This actual IS curve results again suggest more than complete crowd out for tax cut deficits, and complete crowd out for spending deficits. The predicted IS curve results again indicate the same for tax cuts, but only partial crowd out for spending deficits. The lack of a significant difference in crowd out effects in the two periods is further explored in section 7.2.3. below.
- All four crowd out variables in the consumption and investment models are statistically significant at the 1% or 3% level. Statistical significance in both recession and non-recession periods strongly supports the hypothesis crowd out has negative effects in both periods.
- Deficits appears to have about the same marginal effect in recession and non-recession periods on investment (.50 vs.52). For consumption, point estimates of effects are larger for non-recession periods, but well within the other estimates confidence intervals, suggesting the differences may not be significant.
- The recession/non-recession crowd out model, predicted 9 of 17 IS curve coefficients more accurately than the no crowd out model. The no crowd out model predicted 7 better. The standard for judging results was an IS model tested with separate (T) and (G) variables for recession and non-recessions.
- When the standard for judging results was an IS model tested with only one set of (T) and (G) variables (average crowd out), the recession/non-recession model predicted 10 of 17 coefficients better than the no crowd out model.
- The regression coefficient estimates of IS curve crowd out effects, indicate more than complete crowd out of government spending deficit stimulus, but the coefficient is insignificantly different from zero, suggesting only full crowd out. Results do indicate tax cuts more than fully crowd out stimulus effects. Results are shown in Table 3 below. R/NR model predictions of actual IS coefficients were also found to be far more accurate than no crowd out model predictions.

Table 3: IS Curve Coefficients for Tax and Government Spending Variables

	Tax		Spending	
	Rec.	No Rec.	Rec.	No Rec.
No Crowd Out model Prediction	- 0.75	- 0.75	+ 1.75	+ 1.75
R/NR Crowd Out Model Prediction	+ 0.53	+ 0.80	+ 0.47	+ 0.20
Actual Regression Result	+ 0.87	+ 0.60	- 0.65	- 0.23
(t=)	(5.3)	(2.3)	(-1.1)	(-0.6)

- Actual IS curve regression coefficients suggest substantially different recession and non-recession period effects ($+.87\Delta T_{\text{Rec}} + .60\Delta T_{\text{NonRec}} - .65\Delta G_{\text{Rec}} - .23\Delta G_{\text{NonRec}}$), *with deficits causing a substantially worse crowd out problem in recession than non-recessions*. This may occur because savings fall faster than borrowing in recessions, necessitating a much larger cutback in credit based spending by both consumers and businesses than just that caused by the crowd out effects. This decline would be *coincidental with, but not part of, the crowd out effect*. Our models do not provide an easy way of disentangling the two effects. That said, these differences also cannot be considered different with any statistical certainty; 5% confidence intervals around each estimate contain the other as a possibility.

CONCLUSION

The findings indicate crowd out has a statistically significant negative effect on domestic consumer and investment spending in both recession and non-recession periods.

Though the IS curve coefficients indicate both tax cuts and government spending had larger crowd out effects in recessions than non-recessions, the confidence intervals around these estimates were large enough to indicate there may not be a real difference. The likely reason we find crowd out in recessions is because loanable funds availability (savings) drops as much or more than borrowing demand. *This is a key finding, for it is sometimes argued that crowd out is not a problem in recessions, when deficit stimulus is needed most, because private demand for loanable funds declines.*

Actual regression results for the IS curve in both recession and non-recession periods indicate crowd out *more than* fully offset all tax stimulus, but *only* fully offset government spending stimulus in both periods. Predicted results were the same for tax cuts, but only suggested partial crowd out for spending deficits. IS curve coefficients are predicted more accurately from crowd out models than from no crowd out models.

OTHER RESULTS

Additional tests were undertaken comparing the predictive ability of average and recession/no-recession models, and concluded little difference between them existed. Other tests indicated the same

models which explained the decline in consumption and investment fully explained the decline in business borrowing and most of the decline in consumer borrowing. A final section notes that crowd out can be avoided if foreign borrowing is used to finance stimulus programs rather than domestic borrowing, and increases in savings components of M2 in the three years immediately prior to stimulus borrowing, presumably because this offers an alternative source of loanable funds. These topics are examined in detail in Department of Economics Working Paper #1103, available from Rensselaer Polytechnic Institute's Department of Economics website.

BIBLIOGRAPHY

- _____. 2003, 2011. *Economic Report of the President, 2002 and 2010*. Washington, D.C. U.S. Government Printing Office.
- Heim, John J. October 2007. "How Much Does The Prime Interest Rate Affect U.S. Investment?" *Journal of the Academy of Business and Economics*. Vol. 7 (1).
- Heim, John J. October 2010. "Do Government Deficits Crowd Out Consumer and Investment Spending?" *Journal of the Academy of Business and Economics*, Vol. 10 (3).
- Hill, R.C., W. Griffiths and G. Lim. 2011. *Principles of Econometrics*. 4th ed. Hoboken, NJ: John Wiley & Sons.

A Method For Separating Income & Substitution Effects Of Exchange Rate Changes

John J. Heim*

ABSTRACT

Regression coefficient estimates of exchange rate total effects on aggregate demand are broken into separate income and substitution effects. Total effect (substitution and income) regression estimates can seem contrary to theory and common sense. Separating them into their two components shows this is not the case. The separation method also provides a simple test to determine if imports are normal or inferior goods. The paper finds consumer imports are normal goods, investment imports are inferior goods. The paper shows that if import total effects exceed domestic total effects, imports are a normal good. If smaller, they are inferior goods. (JEL: E00, F40, F43)

INTRODUCTION: CONSUMER DEMAND

A recent study indicated the U.S. exchange rate was systematically related to the level of consumer spending, particularly on imports (Heim 2008). In this study, demand for domestically produced or imported consumer goods was regressed on a range of variables commonly held to be determinants of consumer spending, including disposable income, interest rates, consumer wealth and the relative price of imports compared to domestic goods, as measured by the exchange rate. Higher exchange rate values indicate more foreign currency can be bought per dollar, which in turn can mean cheaper import prices. An additional determinant, measured by the government deficit, provides a measure of the extent to which consumers are crowded out of the credit market by government borrowing. The spending and tax variables are reported separately, rather than as a net figure, since preliminary testing indicated deficit increases due to increased government spending restrict consumer credit less than tax cuts. Key regression findings are summarized in the equations in table 1.

Subscripts of zero on variables, or no subscripts at all, mean the current period value of the variable is used. Subscripts with negative signs indicate the number of years the variable is lagged. The equations are estimated using first differences of the data e.g., $\Delta(C)_0$ to help reduce multicollinearity and autocorrelation problems, and sometimes non-stationarity problems in otherwise highly correlated data.

Adding the exchange rate variable seemed to have a major influence on demand for imported consumer goods, adding 6%-points to explained variance. However, explained variance did not increase in the other two models. These exchange rate coefficients show the total effect of an exchange rate change on consumer demand. Below this effect will be separated into its income and substitution effect components.

*Lecturer/Clinical Professor of Economics, Rensselaer Polytechnic Institute, Troy, NY 12180

Table 1: The Determinants of Demand for Total, Imported and Domestically Produced Consumer Goods (Nominal Exchange Rate Used)

$\Delta(C)_0$	$=.63\Delta(Y-T_G)_0$	$+ .47\Delta T_{G(0)}$	$+ .06\Delta G_0$	$- 6.22 \Delta PR_0$	$+ .60 \Delta DJ_{-2}$	$+ 2.69 \Delta XR_{AV0123}$	$R^2=91\%$
(t)	(18.2)	(4.3)	(0.5)	(-3.4)	(4.0)	(1.5)	D.W.= 1.7
$\Delta(M_{m-ksm})_0$	$=.06\Delta(Y-T_G)_0$	$+ .27\Delta T_{G(0)}$	$- .18 \Delta G_0$	$- 3.94 \Delta PR_0$	$+ .26 \Delta DJ_{-2}$	$+ 4.33 \Delta XR_{AV0123}$	$R^2=83\%$
(t)	(2.8)	(4.6)	(-1.9)	(-2.6)	(1.8)	(2.9)	D.W.= 1.5
$\Delta(C- M_{m-ksm})_0$	$=.57\Delta(Y-T_G)_0$	$+ .20\Delta T_{G(0)}$	$+ .24 \Delta G_0$	$- 2.28 \Delta PR_0$	$+ .34\Delta DJ_{-2}$	$- 1.64 \Delta XR_{AV0123}$	$R^2=74\%$
(t)	(16.0)	(1.6)	(1.3)	(-0.7)	(2.8)	(-1.0)	D.W.= 1.8

Where

C = Total Consumption

 M_{m-ksm} = Consumer Imports $C- M_{m-ksm}$ = Consumer Goods Domestically Produced $Y-T_G$ = Disposable Income T_G = Government Receipts G_0 = Government Spending on Goods & Services PR_0 = Real Prime Interest Rate DJ_{-2} = A Wealth Measure: the Dow Jones Composite Average XR_{AV0123} = The average nominal exchange rate (trade weighted) for the current and past three Years

The Federal Reserve's trade weighted nominal Broad exchange rate was used above; a related study (Heim 2009) used the Federal Reserve's real Broad exchange rate in the same models, and yielded similar results except the exchange rate variable, whose results varied somewhat, as expected.

Table 2: The Determinants of Demand for Total, Imported and Domestically Produced Consumer Goods (Real Exchange Rate Used)

$\Delta(C)_0$	$=.66\Delta(Y-T_G)_0$	$+ .49\Delta T_{G(0)}$	$+ .04\Delta G_0$	$- 6.92 \Delta PR_0$	$+ .62 \Delta DJ_{-2}$	$+ 2.83 \Delta XR_{AV0123}$	$R^2=92\%$
(t)	(29.2)	(5.7)	(0.3)	(-3.2)	(4.9)	(3.2)	D.W.= 2.0
$\Delta(M_{m-ksm})_0$	$=.11\Delta(Y-T_G)_0$	$+ .30\Delta T_{G(0)}$	$- .20 \Delta G_0$	$- 5.00 \Delta PR_0$	$+ .34 \Delta DJ_{-2}$	$+ 3.03 \Delta XR_{AV0123}$	$R^2=85\%$
(t)	(6.3)	(5.0)	(-2.0)	(-3.5)	(4.5)	(5.6)	D.W.= 1.8
$\Delta(C- M_{m-ksm})_0$	$=.55\Delta(Y-T_G)_0$	$+ .19\Delta T_{G(0)}$	$+ .24 \Delta G_0$	$- 1.92 \Delta PR_0$	$+ .28\Delta DJ_{-2}$	$- .20 \Delta XR_{AV0123}$	$R^2=74\%$
(t)	(16.2)	(1.5)	(1.3)	(-0.6)	(1.9)	(-0.2)	D.W.= 1.8

Adding the exchange rate variable to the total, imports, and domestically produced consumer goods models above increases explained variance by 2%, 8% and 0% respectively. Notice the estimated total effect of exchange rates on consumer demand for imports is larger than the estimated total effect on

domestic goods, and that the estimated total effect on domestic goods is negative. Later we will show that this implies consumer imports are normal goods and that the substitution effect outweighs the income effect.

INTRODUCTION: INVESTMENT DEMAND

Similarly, the (2008a) study indicated the exchange rate played the following role in determining the level of spending on domestic and imported investment goods:

Table 3: The Determinants of Demand for Total, Imported and Domestically Produced Investment Goods (Nominal Exchange Rate Used)

ΔI	$=.28\Delta ACC + .95\Delta DEP + 1.48\Delta CAP_{-1} + .52 \Delta T_G - .63\Delta G - 6.40\Delta PR_{-2} - .20 \Delta DJ_{-2} + .16 \Delta PROF_{-2} + 6.92 \Delta XR_{AV0123}$	$R^2=.89$
t=	(7.9) (3.2) (1.0) (5.6) (-2.9) (-3.8) (-0.8) (0.9) (3.8)	DW =1.9
$\Delta(M_{ksm})$	$=.04\Delta ACC + .38\Delta DEP + 1.52\Delta CAP_{-1} + .07 \Delta T_G - .22\Delta G + 1.54\Delta PR_{-2} + .19\Delta DJ_{-2} - .10 \Delta PROF_{-2} + 2.52 \Delta XR_{AV0123}$	$R^2=.70$
t=	(1.8) (4.1) (2.2) (2.5) (-2.1) (1.3) (2.6) (-1.1) (-2.1)	DW =2.4
$\Delta(I-M_{ksm})$	$=.24\Delta ACC + .57\Delta DEP - .04\Delta CAP_{-1} + .44 \Delta T_G - .41\Delta G - 8.00\Delta PR_{-2} - .38 \Delta DJ_{-2} + .26 \Delta PROF_{-2} + 4.39 \Delta XR_{AV0123}$	$R^2=.85$
t=	(8.7) (1.9) (-0.0) (5.4) (-2.0) (-4.9) (-1.5) (1.5) (1.8)	DW =1.6

Where

I = Total Investment Demand

M_{ksm} = Demand for Imported Investment Goods

$I-M_{ksm}$ = Demand for Domestically Produced Investment Goods

ACC = The Accelerator, a measure of the growth rate of the GDP each year

DEP = Depreciation Levels of Capital Equipment

CAP_{-1} = % of manufacturing capacity currently being utilized, lagged one year

$\Delta PROF_{-2}$ = Corporate profits, lagged two years

Other variables used are defined in the consumption equations. Subscripts have the same meanings as before and first differences of the data are again used. Notice that, unlike the consumption equations, in the investment equations the larger estimated total effect of a change in exchange rates on demand is in the domestic demand equation, not import demand. We will show later that this is a sign that investment imports are inferior goods.

These coefficients estimates of the exchange rate total effects will be separated below into income and substitution effects. The exchange rate does add significantly to the explanatory power of some equations. The exchange rate appears to have a major influence on demand for imported investment goods, adding 6%-points to explanatory power, but also adds 4% to explained variance when added to the total investment demand and 2% to domestically produced investment goods demand models.

We note that the regression results indicate that for every single - point (~ 0.8%) decline in nominal Broad exchange rate from 2000 levels, making imports more expensive, there appears to be a \$4.39

billion decrease in demand for domestically produced investment goods as well as a 2.52 billion decrease in demand for imported investment goods. We will show later this finding for domestic goods is not nearly as irrational as it appears to be at first blush. It is totally consistent with our estimates of the sum of income and substitution effects for investment goods, particularly our estimates showing investment goods, as a group, are inferior goods and that this trait is transmitted through the substitution effect.

Results for the same investment demand model using the real Broad exchange rate (Heim 2009), are presented in Table 4 below. They were very similar; with the total effect estimates above not changing much for any of the variables, except the exchange rate, which was expected:

Table 4: The Determinants of Demand for Total, Imported and Domestically Produced Investment Goods (Real Exchange Rate Used)

ΔI	$=.28\Delta ACC + 1.37\Delta DEP + .69\Delta CAP_{-1} + .52 \Delta T_G - .61\Delta G - 8.46\Delta r_{-2} - .10\Delta DJ_{-2} + .35 \Delta PROF_{-2} + 4.97 \Delta XR_{AV0123}$	$R^2=.89$
t=	(6.9) (4.7) (0.4) (5.3) (-3.4) (-3.5) (-0.4) (2.0) (4.2)	DW =2.3
$\Delta(M_{ksm})$	$=.05\Delta ACC + .46\Delta DEP + 1.25\Delta CAP_{-1} + .07 \Delta T_G - .14\Delta G + 1.12\Delta r_{-2} + .30\Delta DJ_{-2} - .11 \Delta PROF_{-2} - .40 \Delta XR_{AV0123}$	$R^2=.64$
t=	(1.9) (4.5) (1.4) (2.0) (-1.7) (0.7) (3.4) (-1.09) (-0.7)	DW =2.1
$\Delta(I-M_{ksm})$	$=.24\Delta ACC + .91\Delta DEP - .15\Delta CAP_{-1} + .45 \Delta T_G - .47\Delta G - 9.59\Delta r_{-2} - .40 \Delta DJ_{-2} + .47 \Delta PROF_{-2} + 5.37 \Delta XR_{AV0123}$	$R^2=.88$
t=	(7.8) (3.0) (-0.4) (6.0) (-2.9) (-7.3) (-1.9) (4.1) (4.1)	DW =2.1

Adding the exchange rate variable increases explained variance in the total, imports and domestically produced goods models by 4, 0 and 5% respectively.

Ultimately, the sign of the total effect of an exchange rate change on spending is the sum of its two parts: the “pure” income effect and the substitution effect. If the substitution effect is negative, and large, it may “swamp” the positive income effect, and leave a negative sign. If not, the sign may be positive. For example, by separating the total effects of a change in exchange rates in tables 1-4 above into their separate income and substitution effects, we can explain results which otherwise seem illogical, or at least puzzling. We (again) note that the regression results indicate that for every single - point decline in the nominal exchange rate (imports more expensive) there appears to be a \$4.39 billion decrease in demand for domestically produced investment goods and a 2.52 billion decrease in demand for imported investment goods. The decrease in import demand is understandable since if import prices are increasing, the real income effect is negative. Further, if investment goods are normal goods, the increase in relative import prices should cause further movement out of imports. However, we do not find demand for domestically produced investment goods increases. This may be because the negative income effect swamps the positive substitution effect, or if imports are an inferior good, the overall effect may be out of domestic goods because both effects are negative. To know for certain, we must parse out the separate income and substitution effects.

There is a simple method for breaking down the total effect of an exchange rate change into its income effect and substitution effect components, so that we can resolve such questions. The method will

use information we already have on the estimated total effect of exchange rate changes on and the information inherent in the following identities for domestic (D) and imported (M) goods:

$$\text{Total \$ Effect } (T_D) = \$ \text{ Income Effect } (I_D) + \$ \text{ Substitution Effect } (S_D)$$

$$\text{Total \$ Effect } (T_M) = \$ \text{ Income Effect } (I_M) + \$ \text{ Substitution Effect } (S_M)$$

Since economic theory holds that *the real value* of (pure) substitution effects are symmetric except for sign, this means that in money terms

$$S_D = - (S_M)$$

i.e., $T_D - I_D = - (T_M - I_M)$

Since T_D and T_M are known from regression analysis, this leaves us with one equation in two unknowns: I_D and I_M , *the dollar value* of the pure income effect. However, we will show that these two pure income effects must be the same. Therefore, we have but one equation in one unknown to solve, which is a simple task. This will be done further below.

OTHER EMPIRICAL ESTIMATES OF INCOME AND SUBSTITUTION EFFECTS: A LITERATURE REVIEW

Elmendorf (1996) has excellent estimates the total effect of changes in interest rates on consumption, but does not break them down into income and substitution effects. Baker, Gruber and Milligan (2003) examined the impact of Canada's government retirement programs on work incentives, but, again, did not attempt to separate the total effect of the onset of retirement income into income from substitution effects. Others have attempted to separate these two effects, but have used "hypotheticals", such as "what if" survey responses, instead of data to estimate one of the effects. The other effect is then inferred from indirect evidence. For example, Kimball and Shapiro (2008) estimated income effects of a income increase by asking survey respondents "what if" they won a sweepstakes. How would their work habits change if they won an independent income for life? Using these responses as "income effects", restrictions from labor theory, and known total effects of approximately zero, they inferred substitution effects of a magnitude similar to income effects. Ward and Worach (2005) examined whether the cheaper price of land line phone service available to low income people ("Lifeline" service) affected their demand for cell phone services compared to others. They found a 17.8% difference before controlling for income and demographic characteristics. After controlling for these factors, only 3.1% remained. Since they had attempted to control for income, they were inclined to view the remaining 3.1% as a substitution effect. To ensure the 3.1% was not some contaminated by some residual income effect mixed in, they examined usage of other communication services (cable TV and internet) and income level. When income and demographic characteristic were controlled for, they could not find any significant remaining relationship between usage of cable TV or internet, and usage of the "lifeline" service by poor people. They concluded the 3.1% effect was the substitution effect.

INCOME AND SUBSTITUTION EFFECTS IN THEORY

In standard economic theory, utility is derived from consumption. Utility varies as the combination of goods consumed changes. The combinations considered here are domestically produced goods and imported goods. (D) represents the bundle of domestically produced goods consumed by businesses (investment goods) or consumers (consumer goods); (M) represents the bundle of investment or consumer imports. The utility relationship is given as

$$\text{Utility (U)} = f(D, M)$$

One example of this relationship might be

$$U = D * M$$

$$\rightarrow D = U/M$$

i.e., utility grows in both D, M subject to diminishing returns. This function provides us with an example of a standard – shaped hyperbolic indifference curve in which U is increasing in D, M.. For example, we might find (were utility cardinally countable),

$$U = 100 = D_{10} * M_{10}$$

$$U = 400 = D_{20} * M_{20}$$

Where the subscripts on D, M represent the real quantities consumed. Consumers (and businesses) choose utility maximizing combinations of D and M, given their budget constraints

$$P_D * D = \text{Budget (B)} - P_M * M$$

If the budget is 20 and prices are $P_D = P_M = \$1$, the feasible combinations of goods the consumer can buy with a the budget is given by $(D = 20 - M)$, where “20” might be interpreted as \$20 billion and $U = 100$ at utility maximization. The combination of goods that fully expends the budget and provides the highest utility level is $D=10$ and $M=10$. Other purchasable combinations provide lower utility, for examples

$$D_{=18} = 20 - M_{=2} \Rightarrow U = 15 * 5 = 36$$

$$D_{=15} = 20 - M_{=5} \Rightarrow U = 15 * 5 = 75$$

$$\underline{D_{=10} = 20 - M_{=10} \Rightarrow U = 10 * 10 = 100}$$

$$D_{=5} = 20 - M_{=15} \Rightarrow U = 15 * 5 = 75$$

$$D_{=2} = 20 - M_{=18} \Rightarrow U = 15 * 5 = 36$$

Etc....

We can see the *dollar amounts of each good which maximize utility are the same*: ($\$1 * D = \$1 * M = 10$). In this specific case, the quantities are also the same. In the more general case dollar equivalence will remain, but quantity equivalence will not. Theory suggests the ratio of the goods selected is inversely related to the ratio of their prices, which implies that for budget (income level) changes, prices remaining constant, utility maximization requires allocation of equal money amounts to both products when income changes. This is true for any income change, from zero income on up. We can measure this “pure income effect” by simply increasing the consumer’s budget without changing the relative prices of (D,M). If income doubles, the vertical and horizontal intercepts on the budget constraint double, but the slope of

the budget line (P_M/P_D) remains the same. It now touches the new (and higher) indifference curve where the new curve has the same marginal rate of substitution (MRS) as before. Therefore, both before and after the change

$$\text{MRS} = \partial D / \partial M = P_M / P_D \quad (\text{Prager, 1993})$$

Which implies

$$\partial D * P_D = \partial M * P_M$$

Or in discrete terms

$$\Delta D * P_D = \Delta M * P_M$$

Clearly this indicates that (except for sign) the money value of substitution effects must be identical. The dollar amount substituted out of one good must equal the amount substituted into the other.

This formulation also clearly indicates that if incomes change *not due to price changes, i.e., prices remaining constant*, we have a “pure income effect”. In order for the above condition above to be met if income is increasing, *spending on both goods must change by the same amount*, no matter what the initial income level. This must hold for all budget levels, e.g.,

$$\Delta 1D * P_D = \Delta 1M * P_M$$

$$\Delta 2D * P_D = \Delta 2M * P_M$$

$$\Delta 3D * P_D = \Delta 3M * P_M$$

Etc....

Hence we conclude this standard theoretical formulation shows that the money value of the “pure” income effect must be the same for two goods when income changes, prices constant. The money value of pure substitution effects are also the same (except for sign). These two findings are of key importance in inferring income and substitution effects from data on total effects.

INVESTMENT DEMAND: DERIVING INCOME AND SUBSTITUTION EFFECTS FROM ESTIMATED TOTAL EFFECTS:

Six cases are evaluated to determine income and substitution effects of exchange rate changes on investment goods. An additional six cases test consumer goods in the same way.

- Cases 1 and 3 test, whether imported investment goods, as a group, are normal goods. Nominal and real exchange rate changes are tested separately).
- Cases 2 and 4 test whether imports should be considered an inferior good, with the inferiority trait passing through the substitution effect (again, nominal and real exchange rates changes are tested separately).
- Finally, a fifth and sixth cases are tested. They test the hypothesis that imports are inferior goods, but that the inferiority trait passes through the income effect. Again, separate tests are run using the nominal and real exchange rates.

Using the method described, these six tests, applied separately to investment and consumption, lead this study to conclude

- consumer imports are normal goods (as a macroeconomic grouping).

- investment imports are inferior goods (as a macroeconomic grouping), with the inferiority trait passing through a negative substitution effect.

Details of these tests, as well as a more complete exposition of the theory, are published in the full version of this paper under the same title in the *Journal of International Finance and Economics*, Vol. 9 (3), Oct. 2009, pp. 69-82.

REFERENCES

- Baker, M, J. Gruber and K. Mulligan. May 2003. "The retirement Incentive Effects Of Canada's Income Security Programs." *The Canadian Journal of Economics*. Vol. 36 (2).
- Elmendorf, D. 1996. "The Effect of Interest Rate Changes on Household Saving and Consumption: A Survey." *Finance and Economics Discussion Series 96-44*. Board of Governors of the Federal Reserve System.
- Heim, J. J. 2008. "How Falling (Nominal) Exchange Rates Have Affected The U. S. Economy And Trade Deficit." *Journal of International Business and Economics*. Vol. VIII, (1).
- Heim, J. J. 2009. "How Falling (Real) Exchange Rates Have Affected The U. S. Economy And Trade Deficit." *Proceeding of the 61st Annual Conference of the New York State Economics Association*. Ithaca, NY. (In Press)
- Kimball, M. and M. Shapiro. 2008. "Labor Supply: Are the Income and Substitution Effects Both Large Or Both Small?" Working Paper 14208. National Bureau of Economic Research. July.
- Prager, J. 1993. *Applied Microeconomics*. Boston: R. J. Irwin, Inc.
- Ward, M.R. and G.A. Woroch. "Fixed-Mobile Telephone Subscription in the U.S." Department of Economics Working Paper #0501. University of Texas at Arlington. March 2005.
- Wold, H. and L. Jureen. 1953. *Demand Analysis*. Wiley Publications in Statistics. New York: John Wiley & Sons;

Degrowth and the Social Structure of Accumulation

Kent A. Klitgaard*

ABSTRACT

The world economy is now at a point of systemic crisis. The lack of economic growth is producing high levels of unemployment and poverty. At the same time, cumulative growth of the material economy is depleting our resource base threatening humanity with potentially catastrophic climate change. A new set of economic theories must be developed in order to understand the process of degrowth that will be imposed by the planet's biophysical limits. The approach of the Social Structure of Accumulation is a potentially fruitful starting point for understanding change in institutional structure.

INTRODUCTION

The economies of the developed world face a fundamental dilemma as the second decade of the 21st century approaches. Unemployment rates hover near the double-digit level, the government's capacity to stimulate spending is limited by debt and political processes, and poverty levels are on the rise. From the mainstream point of view the situation calls for growth, and there is no shortage of opinion as to how to best stimulate growth. While economists and politicians may disagree about the best strategy by which to achieve growth, few question the efficacy of growth itself.

Yet, at the same time, there is ample reason to think seriously about the benefits of growth. The economy exists within the parameters of a finite and non-growing planet. Nearly all scientific measures show that we are perilously close to the planet's limits if not, in many instances, in overshoot. However an economy in overshoot cannot grow its way into sustainability. The relation between the limits of the planet and the limits of the economy are treated theoretically as either non-related events or a tradeoff. At the present moment, jobs are at the top of the political agenda. What we need, in order to attain a sustainable economy, is a set of institutional arrangements by which human needs can be met while simultaneously living meaningful lives within the limits posed by the dynamics of the economy and by nature.

Economic activity takes place within social and biophysical contexts. By a social context I mean the set of institutional arrangements that govern how individuals interact with each other and with collective organizations. The biophysical context is set of the limits that are imposed by the laws of nature such as the laws of thermodynamics, energy density of fuels, nutrient availability, and the properties of electromagnetic radiation. We are currently in a systemic crisis of a failed growth economy. The social institutions are capable of providing neither "the good life" for the majority of the Earth's seven billion people, nor sustained economic growth on a biophysically limited planet. Therefore the problems cannot somehow be "fixed" by means of marginal adjustments. The institutions themselves must be reconfigured

*Professor of Economics and Patti McGill Peterson Professor of Social Sciences. Wells College. Aurora, NY

to accommodate the new realities of an economy approaching rapidly its limits. The fundamental purpose of this paper is to introduce contemporary readers to heterodox and institutional theories for the purposes of developing a theory that synthesizes more complete and systemic economic theory. The particular analysis that is developed in this paper is the Social Structure of Accumulation (SSA). While the SSA approach holds particular promise for understanding the energy-short and climate-compromised world to come it, in turn, must be adapted. The SSA approach lies well within growth-oriented economics. The social structure of accumulation was originally defined as “the specific institutional environment in which the capitalist accumulation process is organized” (Gordon, *et al.* 1982: 9). Furthermore, the theory asserts that steady capital accumulation requires a well-functioning and supportive set of institutions, and once these institutions are constructed, rapid capital accumulation will follow (Kotz 2006). However, we need to find a set of social institutions that provide for human needs, employment, and opportunity in the absence of the ability to grow. We must first find a set of institutions that will guide the process of degrowth as the economy must back away from overshoot.

SOME EVIDENCE OF BIOPHYSICAL LIMITS

There are a wide variety of biophysical limits to human activity. This paper will focus on three: the global peak of petroleum production; climate change induced by human activity; and the ecological footprint of humanity. Petroleum is close to peaking on a global scale. Oil alone accounts for 40% of U.S. energy use, with coal and gas accounting for about 25% each. All the “new” alternatives (wind, solar) account for less than 1% combined (Hall and Klitgaard 2011). We are truly dependent upon fossil fuels to power our economy. Energy analyst Cutler Cleveland and colleagues attribute 90% of gains in productivity to the increased application of fossil fuels (Cleveland, *et al.* 1984). Once the global oil supply peaks, indicating we have extracted half of the petroleum created in geological time, energy will become less available and more expensive as demand continues to grow, at the current rate 2.5% per annum. The U.S Energy Information Agency expects energy consumption to continue this growth trajectory, more than doubling by 2035 (USEIA, 2011). Once global production peaks, the best oil will have already been acquired, leaving only that which is more expensive, harder to extract, and lower quality. Consequently the cost, in terms of energy, for acquiring new energy (better known as the Energy Return on Investment or EROI) will continue to increase. When the Spindletop field of East Texas was initially tapped in the early 1900s about 100 units (joules or kilocalories) of energy were returned for every one put into the process. By 1974 the EROI for domestic oil had fallen to 9:1 and by 1980 to 3:1 (Hall and Klitgaard 2011). Domestic oil production peaked in 1970, while estimates for the global production peak range between seven years ago and 2050 (Campbell and Lahererre 1998). The age of cheap oil is clearly coming to an end. Since cheap and available energy has been such a primary driver of economic growth, then economic growth will be compromised, forced by the decline of a flexible, energy dense source of fuel for which few, if any, substitutes, are on the horizon. Historically, economic progress has been driven by the substitution of high EROI energy sources for low EROI sources. The medieval economy was based on

wood (biomass) wind, and the muscle power of humans and animals. The initial stage of the industrial revolution was powered by coal, and later by oil and natural gas (Li 2006). If no viable replacement can be found, and the scientific evidence for such a prospect is scant then “business as usual” will not proceed as it has in the recent past. Limited and expensive petroleum will limit the rate of economic growth, not by choice, but by necessity.

At the same time a case can be built that humans need to restrict the growth of the material economy as regards carbon emissions. Preindustrial atmospheric carbon concentrations stood at approximately 275 parts per million (ppm). Historically CO₂ concentrations have varied between 200 and 300 ppm over the past million years Atmospheric CO₂ concentrations of 200 ppm or less have correlated with ice ages, while levels in the 275-300 ppm range have produced the geologically unprecedented climate stability of the Holocene period. Since about 1750 humans have added about 2 ppm to the atmosphere annually. As of now carbon dioxide equal concentrations approximately 390 ppm. Climatologist James Hansen, of NASA Goddard Space Centers, argues that an atmospheric carbon concentration of 350 ppm represents a theoretical “tipping point,” beyond which damage to climate stability will be irreversible. The controversy surrounding climate change has long been posed as a problem that will be faced by our grandchildren. However the mounting evidence on droughts, floods, sea surface warming, sea-level rise, glacial melt, the cost of storm damage, and insect infestation indicate it is our generation that must deal with the consequences of a warmer planet now. Bill McKibben asserts that our planet will be little like the planet on which human civilization evolved if carbon emissions are not limited immediately (2010).

While the climate science is settled, the political debate rages, essentially around the issue of economic growth. Any attempt to reduce carbon emissions will entail not just increasing energy efficiency, but will also necessitate reduced consumption. The nations of the world have been unable to agree on a framework for reducing carbon emissions, despite numerous recent attempts (Copenhagen, Cancun, etc.) precisely because reducing carbon would entail the reduction of growth in the material economy in a time of overall world crisis. A similar analysis of the need for reduction in the size of the material economy is found in the ecological footprint literature.

Ecological footprints measure the amount of food, fiber, and non-fossil energy that is required to support the existing lifestyle of a particular area. The amount of land needed per capita to produce water, food and fiber, housing and waste management is then normalized to a unit called the global hectare (gha) which is simply a hectare ($1 \times 10^4 \text{ m}^2$) of average productivity. In 2006, according to the Global Footprint Network the world’s ecological footprint for the world was 2.2 gha, yet the biocapacity was but 1.8 gha. In other words, humanity is using about 120% of the planets annual flow of primary productivity. Moreover, the distribution of resource use is far from equally distributed. The ecological footprint for the United States in the same year was 9.6 gha. If everyone in the world were to consume as does an “average American” the world would need nearly five planets to support this level of consumption. However we only have one! Dietz, *et al.*(2007) estimate that the human ecological footprint is growing at 2.12% per year, and that by 2015 we will need 1.6 planets to provide food, fiber and energy to a growing

human population. Clearly this is unsustainable, especially when the driving force of the short-term ability to exceed carrying capacity, cheap oil, is at its global peak. Based on these data Trauger *et al.* (2006) presented ordinary least squares results indicating that the U.S. economy would have to shrink significantly in order to achieve a steady-state equilibrium within biophysical limits. Trauger estimates that a steady-state sustainable balance in the United States would entail a population of less than two-thirds of the present level (200 million) and a per capital annual income of only \$18,500.

This is the dilemma faced at the beginning of the 21st century. If we continue the unsustainable growth path of the past century humanity will most probably face ever-increasing resource shortages, and a more variable climate than the one we are currently facing. If carbon emissions grow at the current level, atmospheric carbon concentrations may reach 1200-1400 ppm. Indeed, if all the Copenhagen proposals were implemented, carbon concentrations would still exceed 725 parts per million by 2100, more than double the levels that have produced the Holocene climate stability. Yet a cessation of growth, under our present institutional arrangement is commensurate with high levels of unemployment, rising poverty, world tensions and, for many, increased hopelessness and cynicism. Economists need to develop a theory that can understand how the economy can meet human needs while remaining within the planet's carrying capacity. Unfortunately, at this time, we do not have such a complete theory. It is towards providing such a theoretical framework that the present essay now turns, tracing the origins and the implications of the Social Structure of Accumulation School.

A BRIEF HISTORY OF THE SOCIAL STRUCTURE OF ACCUMULATION

The SSA framework began in the mid-1970s with empirical studies on structured labor markets. Labor Market Segmentation was a historical process by which labor markets were divided into non-competing sub-markets with different labor market characteristics and work rules. Secondary labor markets, were characterized by low wages and productivity, high turnover, and little skill development. Primary labor markets exhibited high wages, stability and written work rules. The theory also separated the upper tier primary labor markets into two separate segments. Subordinate primary jobs were those in the mass production industries. They were routine, but replete with job ladders, negotiated wages and work rules, and often unions. Independent primary labor markets included professional labor markets and allowed for, if not expected, some degree of creativity and independence (Reich *et al.* 1973).

In 1974 Harry Braverman published *Labor and Monopoly Capital* where he traced the degradation of work in the twentieth century, driven by the needs of large corporations to reduce unit labor cost by increasing productivity and control of the point of production. Braverman chronicles the historical separation of conception from execution, especially as regards organizational changes such as scientific management, and the impact of technological change, in producing work that was devoid of meaning. He also predicted that the standardization and routine that characterized manufacturing would be imposed upon "white collar" and service employment as well. Braverman saw little hope for the resurrection of meaningful work as long as the institutions of large-scale capitalism remained in place.

Contested Terrain by Richard C. Edwards appeared in 1979, arguing that labor market segments evolved historically in response to crises of control at the point of production. Different systems of control characterized different segments of the labor force. The competitive era of capitalism employed simple and direct methods of labor control. The coming of the industrial revolution brought new forms of control. Mass production industries were subject to technical control, or machine pacing, and the nearly unlimited power of the foreman. Edwards contends that further development of capitalism brought forth not greater homogenization, but segmentation. In the years following the Great Depression and Second World War, many companies embarked on strategies to overcome the worst excesses of technical control and instead link the interests of the worker to those of the company. Under a regime of bureaucratic control an employee will advance by means of obeying the all-important company rules. A greater degree of control was offered those who labored under bureaucratic control than those working under technical or simple control strategies. In conclusion different forms of control evolved to fit different economic times, and each segmentation period was dominated by a form of control. Competitive capitalism utilized primarily simple control, while the homogenization period relied on technical control. The segmentation phase was characterized by bureaucratic control. Most importantly, the driving force behind changes in economic structures and institutions were changes in control at the point of production.

The social structure of accumulation approach made its formal appearance with the 1982 publication of *Segmented Work, Divided Workers* by. This book linked changes in the organization of labor and systems to control to broader oscillations, or long swings, in macroeconomic activity and identified three separate long swings in economic activity. The authors explained expansion and contraction by means of changes in the labor process and in the institutional structure. Long periods of accumulations were driven by a set of social institutions that were compatible with profitability and growth. When those social institutions decayed, long periods of economic stagnation followed.

Each SSA went through three phases. In the decline of the long swing traditional methods of profit making become ineffective and the more advanced and innovative capitalists explored new methods of organizing work. As these become generalized they help drive the next long period of economic prosperity in which the social structure is consolidated. But as conditions change the existing social structure is no longer conducive to profit making and accumulation. Consequently the SSA decays with the declining phase of the long swing. In the depth of the depression that signals the decay of the existing SSA the more innovative capitalists explore once again new technologies and modes of organization. These are consolidated in the expansion phase of the next long swing. Gordon, Edwards and Reich identified three complete SSAs. The initial proletarianization of the labor force lasted from the 1820s until the 1870s. The second SSA, homogenization, erected on the foundation of oligopoly, mass production, operative labor, and a management style they called "the drive system," was explored in the long downward swing of the 1870s-90s, consolidated in the early years of the 20th century, and decayed in the interwar years. The final SSA, *labor market segmentation*, was explored in the Great Depression, consolidated fully during the long post-war boom period and began to decay in the early 1970s. But a

question remained unanswered. Could a global phenomenon such as long swings in economic activity be explained fully by changes in the labor process in the United States?

The SSA theorists responded in 1990 by expanding their conception from a labor process-based SSA of segmentation to a broader post-war social structure of accumulation. In *After the Wasteland* Bowles, Gordon, and Weisskopf specified four pillars of prosperity upon which the postwar SSA was consolidated. These included *Pax Americana*, or the postwar peace established upon the basis of US dominance of manufacturing, money, and military might. A second core institution was a limited capital-labor accord whereby manufacturing workers negotiated the sharing of productivity gains with capital. A capital-citizen accord produced a mild Keynesian welfare state, as the gains of the New Deal were expanded and enabled by a commitment to economic growth. Finally the application of anti-trust law and regulations to accept large-scale corporations where they existed, plus the co-respective behavior of oligopoly firms limited the kind of cutthroat competition that proved so ruinous in the long decline of the 1870s-1890s.

This was sufficient to produce growing productivity, rising profits, strong capital accumulation and growth until the early 1970s, when internal and external forces combined to cause the decay of the postwar SSA. The costs of maintaining a complex empire began to mount and add to the country's balance of payments deficit. Moreover the domestic supply of oil peaked in 1970 and in 1973 the country was subjected to significant run-ups in the price of oil and gasoline. The Bretton Woods Accords collapsed when volume of claims exceeded the gold stock. Facing the loss of profits linked to the decline of US hegemony, U.S. corporations could no longer "afford" a limited capital-labor accord. The costs of military spending in conjunction with civilian spending necessary to maintain the capital-citizen accord at a time of high unemployment produced demand pull-inflation. This combined with the cost-push inflation driven by resource price increases, union contracts with cost of living adjustments, and the administered pricing of oligopolies produced simultaneous inflation and recession. Moreover Europe and Japan began to compete effectively with the United States on world markets, increasing the pressure upon profits. The postwar SSA simply could not withstand the myriad pressures of the stagflation era and began to decay. Productivity growth declined from 2.7% per year in the 1950s to 0.3% per year in the 1970s, while GNP growth fell accordingly, from 4% in the 1950s and 1960s to 2.6% per year in the 1980s (Bowles, 1990: 6). Yet a new SSA was not immediately constructed as the groups and classes that benefited from the old institutional arrangement still possessed sufficient power to block a new, and more conservative, agenda based upon more business friendly principles. A period of impasse ensured for nearly a decade.

Out of this decay of the old order a new SSA was explored. Keynesian economic theory and policy was rejected in favor of "supply side economics." The new approach consisted of a commitment to five major policy initiatives designed to reverse the declines in productivity growth, corporate profitability, and international power. The first entailed the use of contractionary monetary policy to plunge the economy into the "cold bath" of recession, thereby stemming the tide of inflation and increasing unemployment sufficiently to restrict union power and increasing wages. Secondly workers' organizations were attacked directly and the National Labor Relations Board was staffed with administrators hostile to the very notion

of collective bargaining. The third component was the deregulation of business in general and finance in particular. Tax policy was employed to shift the distribution of income towards those at the top. Finally the last policy aspect was a remilitarization and more aggressive use of the military to achieve international goals of re-establishing the hegemony of the immediate postwar period.

Because contradictory forces exist in the institutional structure the acceptance of the program did not lead to rapid capital accumulation. The tight money policy that served to reduce inflation and break the power of labor also increased interest rates enough to choke off real investment. High interest rates draw foreign portfolio capital and increase the value of the dollar on international markets. This is beneficial as it makes access to raw materials, especially dollar-denominated oil, cheaper. However the same high interest rates make exports more expensive and exacerbate the trade deficit that emerged in the 1980s. Although the profit share increased with the increase in inequality, the output-capital ratio fell sufficiently, due to the expansion of excess capacity, that the growth in profit rates remained sluggish. In the end the 1980s produced economic growth, but that growth was based primarily on the expansion of debt.

By the early years of the 21st century SSA theorists debated whether the conservative explorations of the 1980s had been consolidated into a new SSA based on neoliberalism. Kotz and Wolfson (2010) argue that a new neoliberal SSA was consolidated over the last decades of the 1900s. Kotz (2009) argues that a core set of neoliberal institutions have been sufficiently implemented to establish a new SSA. These include 1) a belief in small government, 2) deregulation of domestic finance, 3) deregulation of labor markets and a more market-based approach to industrial relations, 4) the free movement of international money capital and finance, 5) international relations to establish US military power, and 6) the advancement of the belief that the neoliberal agenda was inevitable.

However they also believe that the internal dynamics and contradictions of the neoliberal SSA have not manifest themselves as rapid and sustained capital accumulation. The rate of GDP growth in the neoliberal SSA (1990-2007) was 2.99% per year, a rate not significantly higher than the growth rate of 2.93% during the decay of the postwar SSA. The deregulation of finance has led to an expansion of debt and the rise of asset bubbles, and the traditional role of the SSA in balancing the rivalry between industry and finance has not been forthcoming. The balance has been tipped towards finance. The increase in inequality has made it difficult to resolve the conflict between productivity increases and effective demand. The postwar SSA capital labor accord allowed the sharing of productivity gains, leading to increased consumption. The neoliberal SSA is based on the idea that productivity gains belong to capital. In the neoliberal SSA increases in productivity justify reduction of employment not increased compensation. This exacerbates the problem of effective demand and pushes consumers to contract increasing amounts of debt. The trade and payments deficits plus the globalization of finance removes the independence of domestic authorities. The United States has reached the limits of economic expansion by means of debt-financed fiscal policy and is at or near the limits of monetary policy. Kotz, in fact, advocated the separation of the SSA from the need to produce sustained growth, and suggested replacing term social structure of accumulation with that of Institutional Structure. He shows that Liberal

Institutional Structures produce lower rates of growth than do Regulated Institutional Structures (Kotz 2006). Harkening back to *Segmented Work, Divided Workers*, Wallace and Brady contend that we have already entered a new SSA which they term Spatialization. This encompasses the global mobility of capital with new computer technologies. The method of control is called technocratic control, and is built upon information technology. While computerization gives a small minority tremendously creative work, the computer routinizes the work of most (Wallace and Brady 2010).

This is the dilemma that economists must face directly. If an economy is at its biophysical limits it cannot continue to grow without risking potentially severe ecological crises in the near future. But if a growth-dependent economy fails to grow it produces unemployment, loss of wealth and social dislocation in the immediate period.

A SUSTAINABLE INSTITUTIONAL STRUCTURE

The neoliberal SSA is now teetering on the brink of collapse, as the world seems to have reached a period of “peak debt.” Despite the conservative belief that further deregulation will bring about another period of vibrant growth, the economies of the world’s developed economies are entering a period of austerity. Yet most economists believe that the market forces will again balance the world economy if given sufficient time. But it is likely that by the time the internal imbalances of the world economy are rectified, the biophysical limits will have begun to assert themselves more seriously. In this case growth is neither possible without further ecological overshoot, nor wise. Instead we must build an institutional structure built upon sustainability. But first, the world must enter a period of degrowth, in order to step back from the impending biophysical limits.

The most difficult task will be to accommodate employment and opportunity in the absence of economic growth. In this case, something other than mass consumption of increasingly energy-intensive and sophisticated gadgets must replace the current vehicle of well-being. Secondly, a more equal distribution of income must replace the emphasis on the growth of income, as the biophysical system can no longer accommodate income growth without the increasing probability of ecological catastrophe.

A Sustainable Institutional Structure (SIS) must be based on meaningful work, with less emphasis on productivity and quantity. Here, a return to the original focus of SSA theory, changes in the labor process, will be crucial for understanding how work has been stripped of its meaning. If ecological economists do not treat seriously the need for work, then the insights about the limits of nature will not be heard. If growth-oriented economists neglect biophysical limits, then economics will be of little use in a climate compromised and resource scarce world. For vast majority of human history, until the age of fossil hydrocarbons, human society was characterized by two phenomena: production for use and utilization of the solar flow as an energy source. The brief era of fossil hydrocarbons allowed us to not only transcend the diffuse and hard to capture energy source of the sun, but also allowed a vast increase in the scale of production and exchange. I can only hope that we have learned enough from the expansion of science that accompanied the fossil fuel era that we can do a better job of capturing and storing the solar flow

than did our medieval ancestors. When fossil hydrocarbons are sufficiently depleted and climate stability compromised, we will, out of necessity, return to a smaller-scale, more community-based system of production for use. The hard work is envisioning a new and sustainable institutional structure, and managing the transition to degrowth needed to step back from our current overshoot of nature's limits, with a minimum of social dislocation. This will entail a far greater amount of work on my part.

REFERENCES

- Bowles, S, D. Gordon, and T. Weisskopf. 1990. *After the Wasteland*. Armonk, NY: M.E. Sharpe, Inc.
- Braverman, H. *Labor and Monopoly Capital*. New York: Monthly Review Press.
- Campbell, C., & J. Laherrere. 1998. "The End of Cheap Oil." *Scientific American*. March: 78-83.
- Cleveland, C, R. Costanza, C.A.S. Hall and Kaufmann. 1984. "Energy and the US Economy: A Biophysical Perspective." *Science* 225: 890-897.
- Deitz, T, E. A. Rosa and R. York. 2007. "Driving the Human Ecological Footprint." *Front. Ecol. Environ.* 5(1): 13-18.
- Edwards, R. 1979. *Contested Terrain*. New York: Basic Books.
- Gordon, D, R. Edwards, M. Reich. 1982. *Segmented Work, Divided Workers*. Cambridge: Cambridge University Press.
- Hall, C.A.S. and K. A. Klitgaard. 2011. *Energy and the Wealth of Nations*. New York: Springer
- Kotz, D. 2009. "The Financial and Economic Crisis of 2008: A Systemic Crisis of Neoliberal Capitalism." *Review of Radical Political Economics*.(41): 305-317.
- Kotz, D. and M. Wolfson. 2010. "Global Neoliberalism and the Contemporary Social Structure of Accumulation." In *Contemporary Capitalism and Its Crises*. McDonough, T., Reich, M., & Kotz, D. Eds. Cambridge: Cambridge University Press.
- Minqi, Li. 2006. "Capitalism with a Zero Profit Rate?: Limits to Growth and the Tendency for the Rate of Profit to Fall." In *Growth and Crisis, Social Structure of Accumulation Theory and Analysis*. McDonough, T., Reich, M., Kotz, D., & M-A. Gonzalez-Perez. Eds. Galway, Ireland: Centre for Innovation and Structural Change.
- McKibbern, B. 2010. *Eaarth*. New York: Times Books.
- Reich, M, D. M. Gordon and R. S. Edwards. 1973. "A Theory of Labor Market Segmentation." *American Economic Review*. 63(2): 359-365.
- Trauger, D and G. Mansi. 2007. "Ecological Footprints: Tracking Economic Steps Towards Sustainability and Biodiversity Conservation." The Wildlife Society. Tucson, AZ. September 26.
- United States Energy Information Agency. 2011. *World Energy Outlook*
- Wallace, M. and D. Brady. 2010. "Globalization or Spatialization? The Worldwide Spatial Restructuring of Labor Process." In *Contemporary Capitalism and Its Crises*. McDonough, T., Reich, M., & Kotz, D. Eds. Cambridge: Cambridge University Press.

Factors Determining Consumer Sentiment - Evidence from Household Survey Data

Kajal Lahiri^{*}, Yongchen Zhao[†]

ABSTRACT

Consumer sentiment is essential in policy-making and consumption forecasting due to the wealth of information contained in such measure. Unfortunately, such information is not clearly identified in the literature. We study the five components of University of Michigan's consumer sentiment index using household survey data from January 1978 to March 2009 in nonparametric ordered choice model. We find that the most important factors affecting consumer sentiment are their own perceptions and expectations on the economic conditions, the effectiveness of government policies, and news reports. After controlling for these factors, demographic characteristics, macroeconomic conditions and forecasts account for little in addition.

INTRODUCTION

Looking back at the most recent recession, one may find it surprising that the Index of Consumer Sentiment started to fall before April 2007, while the Stock and Watson's monthly GDP estimates started to fall notably from about May 2008. However, this is merely another time when changes in the Index of Consumer Sentiment lead similar changes in the real economy. Such important relationship between consumer sentiment and the real economy, proved repeatedly by the history of our economy, justifies the important role consumer sentiment measures play in business and economic research. However, as is pointed out in Ludvigson (2004), whether consumer confidence surveys contain meaningful independent information about the economy is still largely undetermined, even though work has been done to determine the determinants of consumer sentiment.

The issue of what is captured by consumer confidence measures should be illuminated by an investigation of the factors, which determine sentiment. However, the literature provides a limited discussion of these factors, and offers conflicting evidence. For example, Praet and Vuchelen (1989) and Jennings and McGrath (1994) show that increases in the value of the U.S. dollar affects consumer confidence in the U.S., Germany and France, but not in the U.K. and Italy. They find changes in the U.S. stock market index to affect positively German confidence, but negatively sentiment in the U.K. and to have no effect in France and Italy. Other studies attempt to explain sentiment with general macroeconomic indicators (Lovell, 1975 and Garner 1981), financial assets (Mishkin, 1978) and liabilities or political events (Vuchelen, 2004 and Blood and Phillips, 1995). Throop (1992) shows that a structural

^{*} Department of Economics, University at Albany, SUNY, klahiri@albany.edu

[†] Department of Economics, University at Albany, SUNY, hzhao@albany.edu

model for the US. ICS including inflation rate, unemployment rate and short-term interest rates as explanatory variables explains sentiment fairly well in times of usual political and economic activity but collapses around points of extraordinary events such as the Persian Gulf War. Estelami, et al. (2001) focus on consumer price knowledge and find that macroeconomic factors explain considerable proportion of variations. In a more recent study, Ludvigson (2004) investigates the relationship between consumer sentiment measures and consumer spending and suggests that “it is possible that there are more complex, possibly nonlinear, interactions between consumer confidence and economic variables ...”. Vuchelen (2004) shows that “a few well-chosen variables” explain sentiment of Belgian consumers quite well.

In light of these inclusive findings, this paper begins by examining the evidence on the explanatory power of sentiment in models for consumption using monthly data on the University of Michigan’s ICS and on personal consumption expenditures, dis-aggregated into durable, nondurables, and services for the period 1978:01 to 2009:03. We base our nonparametric estimation on a household level survey data and look very closely at the overall explanatory power of our model as well as the incremental explanatory power of each category of independent variables. Unlike previous studies of this issue, which generally uses quarterly data, we use high-frequency monthly data. Since the original ICS is reported monthly, quarterly data will lose much of the variation in sentiment through averaging.

The focus of this paper is the issue of what determines consumer sentiment. All studies on this subject use time series analysis to explain an aggregate measure of consumer sentiment. It should be noted, however, that measures of sentiment are derived from individual survey data, which contain information about consumer expectations and personal characteristics unavailable at the aggregate level. These data can help fully answer the question of what groups of factors affect sentiment, and how much of it is explainable at all. In Section 2, we examine the data used in this study. Section 3 discusses the models and variables, which are grouped into four broad categories. Section 4 focuses on different aspects of the explanatory power of our models. Section 5 shifts attentions to individual factors and identifies the ones that are important and interesting. Section 6 briefly discusses the specification test for our nonparametric model and Section 7 concludes.

SURVEY OF CONSUMERS

The index of consumer sentiment (ICS) was developed by George Katona (1951) as a measure of consumer’s changing attitudes about the business conditions and job prospects, and the Survey Research Center at the University of Michigan has continued to produce it on a regular basis to measure consumer confidence in the United States. The ICS has generally led the U.S. business cycle and the 1990-1991 recession has been widely attributed to a drop in the ICS in the wake of Iraq’s invasion of Kuwait. The SRC’s survey of Consumer Attitudes and Behavior on which the ICS is based, started in 1952, and has been conducted quarterly since 1960 and monthly since 1978. The five questions on which the ICS is based are given below.

PerFin_Current: “We are interested in how people are getting along financially these days. Would you say that you (and your family living there) are better off or worse off financially than you were a year ago?” Possible responses are: better off; same; worse off. *PerFin_Expected*: “Now looking ahead – do you think that a year from now you (and your family living there) will be better off financially or worse off, or just about the same as now?” Possible responses are: better off; same; worse off. *BusCond_12m*: “Now turning to business conditions in the country as a whole do you think that during the next 12 months we’ll have good times financially, or bad time, or what?” Possible responses are: good times; good with qualifications; pro-con; bad with qualifications; bad times¹. *BusCond_5y*: “looking ahead, which would you say is more likely – that in the country as a whole we’ll have continuous good times during the next five years or so, or that we’ll have periods of widespread unemployment or depression, or what?” Possible responses are: good times; good with qualifications; pro-con; bad with qualifications; bad times. *BuyCond*: “About the big things people buy for their homes such as furniture, refrigerator, stove, television, and things like that. Generally speaking do you thin know is a good or a bad time to buy major household items?” Possible responses are: good; pro-con; bad.

Based on the responses to the above five questions, we can compute the balance statistic as follows
 $(\% \text{ of respondents who answer better or good} - \% \text{ of respondents who answer worse or bad}) + 100$

Then simple average of the five balance statistics is computed to obtain index of consumer sentiment. The ICS is reported relative the base month – February 1960 = 100. For our time series analysis we use monthly data on the ICS dating from January 1978 to March 2009 (375 months). This is the most recent and complete data available to the public at the time this study began.

The design of the survey allows us to extract additional information about individual’s expectations and perceptions of general and personal economic conditions, as well as personal characteristics, which is not contained in the five index questions alone. Specifically, the survey asks questions regarding the individuals’ perception of how the economy has done over the past one year, how the government is doing its job at present, and whether the individual has recently heard any good or bad news about the general economic and political situation in the country. There are also questions about the individuals’ expectations about general economic conditions over the next year, about prices, real income, unemployment, and interest rate on borrowing, all with answers in the three-response framework better/same/worse or up/same/down. There is also information about the respondents’ employment status, education, marital status, gender, race, age, region of residence, household size, and annual income. In addition to the response to the five index questions, we use information on more than 50 other questions to explain response to each of the component questions at the individual level. We used raw data sets from monthly survey, which amounts to a total of 375 months. We have a final sample contains 156,070 observations.

MODEL AND VARIABLES

As outlined previously, we utilize the information from over 50 questions from the survey of consumers asked consistently throughout the sample period 1978:01 – 2009:03 to analyze the factors, which determine individual responses to each of the five ICS component questions in an ordered response model. In this study we are specifically interested in what are the main determinants of the responses to each question; are they the same for all of the five components; and what effect do they have on the components of sentiment measure. Answering these questions will enable us to better discriminate among the existing hypotheses about the informational content of the ICS.

As stated before, we combine the response, when necessary, so that for all of the five components, we have three different kinds of responses: good / same / bad or up / same / down. This enables us to consider the questions raised before in a unified latent-variable framework. Let $y_{it}^* = X_{it}\beta + \varepsilon_{it}$, where the index it corresponds to individual i interviewed at time t , with lower values of y^* denoting higher optimism, i.e., better/up = 2, same = 1, worse/down = 0, and, $X_{it} = (E_{it}, I_{it}, M_t, F_t)$ is a vector of variables affecting an individual's level of optimism, in which E is individual's expectation and perception, I is a set of variables explaining individual idiosyncrasies, M is a set of macroeconomic variables and F is a set of macroeconomic forecasts. The data for F comes from monthly mean of Blue Chip forecasts. Note that a set of monthly dummy variables are included in the I category and a dummy variable for whether the economy is in a recession is included in the M category.

More specifically, for each category *Expectations/Perceptions (E)* category contains individual's perception on overall news, news specifically related to inflation, news specifically related to unemployment, news specifically related to war, current business condition compared with that from one year ago and the performance of government economic policies, as well as individual's expectation of future business conditions, future interest rate, future price level, future real income and future unemployment. *Personal characteristics/idiosyncrasies (I)* category contains individual's age, real per capita family income, racial origin, marital status, education level, gender and location of residence. *Macroeconomics conditions (M)* category contains the levels and percentage changes from previous month of the index of coincident indicators, index of leading indicators, industrial production, S&P 500 index, personal income and purchasing managers' index, as well as the levels of inflation rate, 3-month treasury bill rate, 5-year government bond rate, unemployment rate, along with a dummy variable indicating NBER recession and the standard deviation of the S&P 500 index in the previous month. *Macroeconomic forecasts (F)* category contains monthly mean of Blue Chip forecasts, with one year horizon, of inflation rate, real GDP and unemployment rate.

Given the latent variable y^* , the observation rule is $y_{it} = \begin{cases} 0 & \text{if } y_{it}^* < \mu_1 \\ 1 & \text{if } \mu_1 \leq y_{it}^* < \mu_2 \\ 2 & \text{if } \mu_2 \leq y_{it}^* \end{cases}$, where ε_{it} is an

independently distributed error term with distribution function $F(\cdot)$. The model can be estimated by maximum likelihood. The likelihood function is given by $\ln(L) = \sum_{i=1}^N \sum_{j=1}^J (Pr[y_{it} = j])$. For the error

distribution F , we consider two types of specification here. Firstly we consider a usual Probit² specification, where $F = \Phi$ is a normal distribution function. Secondly, we consider a nonparametric specification (Stewart, 2003) where the distribution function of the error term is thus given by $F_K(x) = \frac{\int_{-\infty}^x (\sum_{k=0}^K \gamma_k \varepsilon^k)^2 \phi(\varepsilon) d\varepsilon}{\int_{-\infty}^{\infty} (\sum_{k=0}^K \gamma_k \varepsilon^k)^2 \phi(\varepsilon) d\varepsilon}$, which is a family of distributions given parameter K . In this study, K is chosen for each of the five components separately according to the BIC statistic.

EMPIRICAL RESULTS

We have discussed the explanatory power of the model and its variation across the five components of the sentiment index and variation over the entire sample period. In this section we focus on the individual factors in each category that has interesting and important impact on the components of sentiment. Table 1 give the regression result of the nonparametric model for each of the sentiment components where marginal effects are reported.

Table 1: Marginal Effects of Selected Variables

Variable	PerFin_Current		PerFin_Expected		BusCond_12m		BusCond_5y		BuyCond	
	y=1	y=3	y=1	y=3	y=1	y=3	y=1	y=3	y=1	y=3
GoodNewsOnly			-0.0027	-0.0034	-0.0122	0.0559	-0.0032	0.0100	0.0030	0.0051
BadNewsOnly			-0.0088	-0.0114	0.0117	-0.0543	0.0020	-0.0063	-0.0049	-0.0084
GoodGovt	-0.0059	0.0765	0.0105	0.0130	-0.0001	0.1044	-0.0221	0.1327	0.0164	0.0248
PoorGovt	0.0373	-0.1026	-0.0249	-0.0332	0.0407	-0.1359	0.0495	-0.1164	-0.0222	-0.0430
Grade_9_11	0.0308	-0.0895	-0.0284	-0.0384	0.0073	-0.0372	0.0482	-0.1140	-0.0240	-0.0473
High_School	0.0181	-0.0603	-0.0041	-0.0052	0.0003	-0.0019	0.0215	-0.0585	-0.0001	-0.0002
Some_College	0.0111	-0.0411	0.0123	0.0152	-0.0008	0.0058	0.0044	-0.0135	0.0038	0.0062
White	-0.0087	0.0337	0.0188	0.0248	-0.0060	0.0315	-0.0059	0.0179	0.0177	0.0330
Black	-0.0049	0.0315	0.0858	0.0989	-0.0001	0.0004	0.0244	-0.0653	0.0130	0.0200
Hispanic	-0.0014	0.0070	0.0190	0.0233	0.0031	-0.0178	0.0187	-0.0519	-0.0029	-0.0048
North_Central	0.0014	-0.0066	-0.0223	-0.0297	0.0013	-0.0080			-0.0001	-0.0001
Northeast	0.0068	-0.0272	-0.0285	-0.0385	0.0043	-0.0236			0.0014	0.0023
South	-0.0025	0.0137	-0.0044	-0.0056	-0.0005	0.0035			-0.0014	-0.0024
Male	-0.0031	0.0176			-0.0042	0.0612	-0.0200	0.0977	0.0209	0.0308

LHF variable (y) is shown in the header row. A blank means the variable is not in the regression model. Full results including explanation of variable names, coefficients, significance and other statistics available upon request from the author.

Expectations and perceptions: We find that perception of government economic policy has an important role in explaining PerFin_Current, BusCond_12m, and BusCond_5y. On average, the marginal effect of perception of government performance is about 5-7%; meaning that other things being equal, the likelihood of people have a good sentiment is 5-7% higher if they think the government is doing a good job. Interestingly, the effect of this on PerFin_Expected and BuyCond is much smaller. A possible explanation is that while consumers do think that government economic policy is important to the overall economy and to the change in their own financial situation, they tend not to think of it as an important factor in forming expectation about their own future and making their own purchasing decisions. Meanwhile, we find that the effect on the sentiment is bigger if consumers think the government is doing a poor job than the effect if consumers think the government is doing a good job. We also find that whether consumers have heard any good or bad news mainly affects BusCond_12m but not the rest of the components, in particular, BusCond_5y. This may be because of the time-sensitiveness nature of news stories – they mostly focus on what's happening “now” instead of on what's going to happen in a relatively long term. In addition, our result shows that good news has a bigger effect than bad news, somehow contradicting to traditional wisdom that consumers are more sensitive to bad news. The marginal effect of good news on BusCond_12m is about 6% while that of bad news is around 1%. In the rest of the variables in this category, consumers' perception and expectation of overall economic condition, price level, and real income are the most important factors in terms of magnitude of marginal effects.

Personal idiosyncrasies: In terms of racial origin, we find that white consumers generally have higher sentiment, in all five sentiment components, then black consumers, then Hispanics. For example, white consumers are about 2-3% more likely to think that their own financial situations as well as the general economic conditions have been improving and will continue to improve in the future. This number is generally less than 1% (or not statistically significant) for black consumers and Hispanics. In terms of gender, male consumers tend to have higher sentiment, especially for long term expectations – BusCond_5y, where male consumers are 10% more likely to have higher sentiment or unlikely to have lower sentiment. For other components of sentiment, the marginal effect of being male is generally 2-6%. The level of education also has a significant effect on consumer sentiment, though the effect is not so clear for some components of sentiment. Our result shows that consumers with higher education level have significantly higher sentiment in PerFin_Current and BusCond_5y – consumers with a college degree or higher are about 1.5% more likely to have higher sentiment than consumers without high school diploma. For the rest of the sentiment components, the effect of education is not clear. As for residence location, we find that for current personal financial situation PerFin_Current, consumers residing in the south have higher sentiment then consumers residing in the west and the consumers residing in the northeast have lowest sentiment. For expected personal financial situation PerFin_Expected and 12-month ahead expectation of business conditions BusCond_12m, consumers from the west have higher sentiment than consumers from northeast. For the rest of the sentiment components, the effect of location of residence is not clear.

Macroeconomic variables and forecasts: Macroeconomic variables and macroeconomic forecasts are generally not so important in determining personal financial situations, both past and expected, PerFin_Current and PerFin_Expected, though an increase in unemployment increases consumers sentiment in PerFin_Current. The rationality behind this seemingly unreasonable result may be that the worse the condition for their peers, the better the consumers feel for themselves, provided that most of them are still employed. As for business conditions, changes in industrial production, forecasted inflation, 5-year government bond rate and unemployment rate has strong effects on BusCond_12m, while the index of coincident indicators and forecasted unemployment rate strongly affect BusCond_5y. For example, a 1% increase in industrial production makes consumers 1.13% more likely to think that the overall business condition will be better in one year; a 1% increase in the coincident indicator makes consumers 1.27% more likely to think that business condition will be better in five years. Unemployment also has a strong effect on BuyCond.

Information content of sentiment: While the current setting of the model does not permit a statistical test on the information content of consumer sentiment, the above analysis and findings do provide some solid clues. In general, we find that sentiment captures consumers' perception and expectation of economic fluctuations, based on their day-to-day experience (hypothesis 3 and 4). Firstly, variation in the sentiment components can be largely explained by variation in consumer's perception and expectations of their own situation, the information they get from news media as well as the condition of the economy, at least the part of the economy they experience. If sentiment was an unmediated experience, and reflects nothing more than official statistical reports, professional forecasts and noise, as in hypothesis 5, we would expect actual lagged macroeconomic variables and macroeconomic forecasts to be able to explain a large portion of the variation in sentiment on the individual level. However, we find contrary. Moreover, macroeconomic variables and macroeconomic forecasts, by themselves, do not provide sufficiently large explanatory power. Secondly, sentiment is clearly not a direct reflection of political and business news, otherwise, our news variables would be expected to capture a lot more variations than what they do now. Not to mention that even they do capture a lot more variation, they are still the perception of the consumers, i.e., different consumers may interpret the same piece of news differently, especially when such interpretation is used to form expectation of the future. Our results show that new variables generally have less than 2-3% marginal effect, about the same as all other important variables. In particular, news of war, unemployment and inflation do not have any important effects on sentiment components, which make us believe that consumers tend not to overreact to particular word(s) that appear in news report. So it is highly unlikely that one news story containing a particular bad word, e.g., the "r" word – recession would trigger a sudden fall in consumer sentiment. This is supported by the result that good news has larger effect on sentiment compared with bad news. The weak performance of news variables as determinants of sentiment directly contradicts the results of Blood and Phillips (1995), who find that recession related headline news have a significant negative prior influence on ICS, even after accounting for the actual state of the economy. Our findings support the view of Linden (1982) and

Blendon, et al. (1997) that consumers' expectations are formed in "conversations between neighbors over the backyard fence" and are not a direct reflection of media coverage or published statistics (Garner, 1991). Finally, we find a clear difference between consumers' sentiment regarding their own personal situations and the conditions of the overall economy. Their personal situations are more likely to be affected by their demographic characteristics and other idiosyncrasies rather than macroeconomic variables or their forecasts. But macroeconomic variable do greatly affect consumers sentiment on business conditions. Expected business conditions have marginal effects about 4-6% for BusCond_12m and BusCond_5y. Expected prices, expected income and expected unemployment situation generally have marginal effects about 1-2% for BusCond_12m and BusCond_5y. Expected business condition and expected employment are also important in BuyCond, with marginal effects more than 1%.

We can see that in terms of overall explanatory power, the model explains from 17% to 56% of the total variation. BusCond_12m – the one-year ahead expectation of the overall business condition is best explained with a pseudo- \bar{R}^2 of 55.86% in probit model and 54.50% in nonparametric model. The least well explained variables are PerFin_Current – personal financial condition compared to a year ago and BuyCond – buying condition for large household items, with a pseudo- \bar{R}^2 of around 17-18% for both the probit model and the nonparametric model. Considering that BusCond_5y – expectation of five year ahead business condition is the second best explained variable with a pseudo- \bar{R}^2 about 40% and PerFin_Expected – expectation of one year ahead personal financial situation is the third best explained variable with a pseudo- \bar{R}^2 of about 30-40%, we come to the conclusion that the model explains expectations about the overall economic condition better that it explains expectations about personal situation. This observation is further confirmed by that fact that macroeconomic variables accounts for a larger proportion of variation in explaining overall business conditions. Another finding we get is that in terms of overall explanatory power, the probit model provides a result that is highly consistent with what the nonparametric model provides. The differences in pseudo- \bar{R}^2 for four of the five variables are within 1 to 2%, except for PerFin_Expected. While the probit model explains about 30% of the total variation of this variable, nonparametric model explains about 50%. This obvious increase in explanatory power may be due to the non-normality of the error distribution for this variable.

In terms of the explanatory power of each category of variables, the two models again give highly consistent results. Generally, expectations and perceptions category has the highest incremental explanatory power, ranging from about 10% for BuyCond and PerFin_Current to 52% for BusCond_12m. In the rest of the categories, personal idiosyncrasies category plays an important role in explaining PerFin_Current and PerFin_Expected, providing pseudo- \bar{R}^2 of 8-10% for the former and 5-8% for the later. Macroeconomic variables and macroeconomic forecasts play equally important role in explaining business conditions and buying condition. And not surprisingly, these two categories are more important in explaining BusCond_12m than the rest two variables. Note that while overall explanatory power for BusCond_5y is much higher than that for BuyCond, macroeconomic variables and macroeconomic forecasts are much less important in BusCond_5y. This shows that uncertainty of long-term expectation of

business condition is so large that the variation caused by idiosyncratic factors is much more than variations caused by the variation in the common information set – current and professionally forecasted economic conditions.

SUMMARY

In this study, we focus on the informational content of the Index of Consumer Sentiment, in particular, its five components. Models based on household level survey data are estimated assuming normal error distribution as well as assuming nonparametric distribution. Our results provide some solid clue on how much we can explain variations in the components of consumer sentiment, what are the important factors affecting components of consumer sentiment, and what information consumer sentiment index reflects.

We find that up to 48% of variations in personal sentiment, and up to 56% of variations in sentiment on business conditions can be explained by variables including consumers' perception and expectation of social economic fundamentals, macroeconomic variables and their professional forecasts, and consumers' demographic characteristics such as gender, racial origin, education level, age, and region of residence. Among them, consumers' perception and expectation are the most important category of variables, accounting for 11% to 53% of the total variation in components of sentiment. Macroeconomic variables and forecasts generally provide much lower explanatory power around 5% to 20%, especially for personal sentiment components where they only provide about 3% to 5% explanatory power.

The most important and influential factors affecting components of sentiment at the individual and household level are consumers' expectation of future business conditions and unemployment level, as well as consumers' perception of the performance of government economic policies and news stories. Consumer's response to general questions regarding, for example, the overall business condition contains complex information from all sources and balance statistics constructed for this kind of questions are not easy to interpret. Our results thus suggest the incorporation of the statistics constructed from direct questions about expected unemployment or family income in the sentiment index.

The above finding suggests that the informational content of consumer sentiment is complex yet tractable. The information contained in sentiment measure is neither purely macroeconomic fluctuation nor the tone of news reports delivered by mass media. Rather, different information takes up the role of the principle driving force for different component of consumer sentiment. It is more informative if components of the sentiment index are treated separately in economic and econometric models. When the components are combined, information of the driving force of changes in sentiment may be obscured.

ENDNOTES

1. Responses good time / good with qualifications and bad with qualifications / bad times are grouped together respectively for index construction. Same for the next question.

2. A logit specification is also considered and proved to produce very similar results as that produced by a probit specification. Thus, the result from logit model is omitted from this paper but is available from the authors upon request.

REFERENCES

- Blendon, R. J, J. M. Benson, M. Brodie *et al.* 1997. "Bridging the Gap between the Public's and Economists' Views of the Economy." *The Journal of Economic Perspectives*.11(3):105–118.
- Blood, D. J. and P.C.B. Phillips. 1995. "Recession Headline News, Consumer Sentiment, The State of the Economy and Presidential Popularity: A Time Series Analysis 1989-1993." *International Journal of Public Opinion Research*, 7(1):2.
- Estelami, H, D. R. Lehmann and A. C. Holden. 2001. "Macro-Economic Determinants of Consumer Price Knowledge: A Meta-Analysis of Four Decades of Research." *International Journal of Research in Marketing*. 2001;18(4):341–355.
- Garner, C. A. 1991. "Forecasting Consumer Spending: Should Economists Pay Attention to Consumer Confidence Surveys?" *Economic*, 57.
- Jennings, E. J and P. McGrath. 1994. "Consumer Sentiment and International Conditions." *The Journal of Economics*. XX:17-21.
- Katona, G. 1951. "Psychological Analysis of Economic Behavior." McGraw-Hill: New York.
- Linden, F. 1982. "The Consumer as Forecaster." *Public Opinion Quarterly*, 46(3):353–360.
- Lovell, M. C. 1975. "Why Was the Consumer Feeling So Sad?" *Brookings Papers on Economic Activity*, 473–479.
- Ludvigson. S. C. 2004. "Consumer Confidence and Consumer Spending." *Journal of Economic Perspectives*. 18(2):29–50.
- Mishkin, F. S, R. Hall, J. Shoven, T. Juster and M. Lovell. 1978. "Consumer Sentiment and Spending on Durable Goods." *Brookings Papers on Economic Activity*, 217–232.
- Praet, P. and J. Vuchelen. 1989. "The Contribution of Consumer Confidence Indexes in Forecasting the Effects of Oil Prices on Private Consumption." *International Journal of Forecasting*. 5(3):393–397.
- Stewart, M. B. 2003. "Semi-Nonparametric Estimation of Extended Ordered Probit Models." *The Stata Journal*. 4(1):27-39.
- Throop, A. W. 1992. "Consumer Sentiment: Its Causes and Effects." *Federal Reserve Bank of San Francisco Economic Review*, 1:35–59.
- Vuchelen, J. 2004. "Consumer Sentiment and Macroeconomic Forecasts." *Journal of Economic Psychology*, 25(4):493–506.

Students' Economic Ideology and Deficit Reduction Choices

Michael McAvoy and William P. O'Dea*

ABSTRACT

We measure the students' economic ideologies in a capstone economics course defining economic ideology as how one views the world. Students are measured to be slightly liberal in ideology. Students submit deficit reduction choices from "The New York Times Deficit Project," to reduce the projected long-term deficit to 3 percent of expected economic output. We compare the ideology measurement to students' deficit reduction choices. While students on average prefer spending cuts, their choices to reduce the deficit have popular liberal leanings: reducing the size of the military and increasing taxes on very high income earners. Medicare and Social Security reform are commonly selected as options for deficit reduction as well.

In this paper, *The New York Times* deficit reduction project (Leonhardt 2010b, Leonhardt and Marsh) is paired with an ideology measurement instrument (Hadsell, McAvoy and McGovern). An undergraduate student often thinks of herself as an adherent of one particular ideology or another. Often, too, undergraduate economics courses focus on perfectly competitive markets and prices to determine the best outcomes without discussion or incorporation of alternative ideologies. The survey instrument employed uses simple economic statements to measure a student's ideology along a continuum. The *Times* deficit reduction project presents choices which may be favored specifically by partisans of one ideology or another. Nevertheless, however a student may think her ideology is classified can measure it using this survey instrument. When combined with the deficit reduction project, she gains a deeper understanding why she prefers one deficit reduction option as opposed to another. In this paper, the instrument ideology scores are compared to choices that students make to reduce the long term deficit. We show that ideology scores tend to correlate with the deficit reduction choices.

In *The New York Times* Economic Scene column, published on November 13, 2010, David Leonhardt (2010b) analyzes the federal budget and asks his readers to make choices to maintain the long term federal budget deficit at 3 percent of output. Readers must choose specific tax increases and spending cuts summing to \$1.345 trillion, to reduce the projected deficit to 3 percent of output by 2030. These choices are an exercise to illustrate the difficulty of reducing the deficit no matter one's political ideology. Leonhardt focuses on the long-term deficit since Baby Boomers will have an increasing impact on the deficit as they begin to retire in large numbers. In Leonhardt's judgment, focusing on deficit reduction from the long view causes the reader to realize the enormity of the task. Why 3 percent rather than a balanced

* Department of Economics, Finance and Accounting, SUNY College at Oneonta, Oneonta, NY, 13820.

◇ We thank Lester Hadsell for his comments on an earlier draft. We remain responsible for all errors and omissions.

budget? Leonhardt (2010a) utilizes budget analysis by Auerbach-Gale, experts on the budget. A deficit of 3 percent is sustainable: one year's deficit may be paid for by economic growth the following year. In addition, interest costs on the accumulated debt are manageable given ongoing deficits of that proportion.

As Leonhardt (2010b) notes, spending cuts alone or revenue increases alone will not balance the long term budget. The numbers Leonhardt presents show that the long term deficit reduction target is 5.5 percent of 2030's gross domestic product. However, domestic discretionary spending is projected by the Congressional Budget Office to be 3 percent of gross domestic product, and military spending about 4 percent. In other words, deep cuts to social programs alone or the military alone will not be enough to eliminate the long term deficit. Therefore, both spending cuts and tax increases are necessary. Spending cuts in social programs are difficult for Democratic liberals to accept, while tax increases are similarly unappealing to Republican conservatives. Leonhardt observes that the bipartisan commission agreed to make substantial cuts to social programs while allowing the Bush era tax cuts to expire on incomes above \$250,000. Factoring in other spending cuts and tax increases, the commission's recommendations reduced the long term deficit by one-third, an amount too small to reduce the deficit to 3 percent of GDP.

Readers of the Leonhardt's column (2010b) are presented with a graphic showing a menu of spending and revenue options. Readers are challenged to reduce the projected 2030 deficit by \$1.345 trillion and post their choices online (Leonhardt and Marsh 2010). Readers are presented with instructions for a selection of options within categories summarized in Table 1. While the spending options total \$1.81 trillion and the revenue options total \$1.955 trillion, due to instructions within categories, the maximum spending cut a reader could choose is \$1,590 billion, and the maximum tax increase is \$1,625 billion.

Table 1: Spending and Tax Options in *The New York Times* Deficit Reduction Project

Spending Options	Billions	Revenue Options	Billions
Domestic Programs and Foreign Aid	\$160	Estate Tax	\$170
Military	505	Investment Taxes	70
Health Care	680	Income and Employer Taxes	715
Social Security	465	Tax Reform	545
		Other New Taxes	455
Total	\$1,810	Total	\$1,955

Among the responses posted to reduce the long term deficit, Leonhardt (2010c) observes that the least popular choice selected (fewer than 10 percent of respondents) allowed the expiration of the Bush tax cuts for incomes less than \$250,000 per year. The most popular spending cut chosen by respondents (more than 80 percent) reduced the military to below its size before the Iraq War. Leonhardt (2010c) notes that respondents' long term choices to reduce the deficit likely reveal their ideological preferences:

“But when it came to tax cuts for incomes above \$250,000, people’s opinions appeared to diverge according to their political views. Those who preferred spending cuts – a conservative group in all likelihood – generally wanted this tax cut to remain in place. Among those who closed the deficit mostly with tax increases – probably a liberal group – the expiration was the single most selected policy.”

The deficit reduction puzzle is an interesting exercise for college students as well. Just as ideology is observed in responses by Times readers, we may observe how students’ ideology affects their choices. Similar to Hadsell, economic ideology is, “a view of the world that influences what one considers important,” and this view affects behavioral modeling and policy positions. To measure ideology, they develop a survey instrument that measures economic ideology along a continuum. A student records her strength of agreement (towards score 6) or disagreement (towards score 0) for twelve short statements in a seven point Likert type scale. The questions are provided in Table 2. The score for a person is the sum of her answers to the statements (see note to Table 2). The larger (smaller) her score, the more liberal (conservative) she is. A score of 36 is the midpoint. Hadsell finds the survey score measurement to be consistent with the person’s stated ideology.

Table 2: Survey Questions and Students Average Score to Each

Number	Question	Score
n1*	If people are poor it is mostly because of their own actions.	2.94
n2	The price of pharmaceuticals (drugs) should be regulated by the government so that the drugs are more affordable to the average person.	2.29
n3	Government is needed to help the less fortunate.	3.71
n4*	At birth, everyone has an equal opportunity to become rich.	3.65
n5	The government should provide unemployment compensation and retraining to people who lose their jobs as a result of a weak economy.	2.24
n6	People with very high incomes should be heavily taxed.	3.29
n7	Consumers need active government protection from unethical business practices.	2.47
n8	Women and minorities face significant discrimination in the labor market.	3.53
n9	Sick persons should always be provided healthcare whether or not they have ability to pay.	3.71
n10	The government should provide basic housing to those unable to pay market price rent.	2.76
n11	The government should use tax money to subsidize the development of environmentally sustainable (green) technologies.	2.71
n12	Every employed person should be guaranteed a fair wage.	3.71

Note: Answers are scored 0 (strongly disagree) to 6 (strongly agree), except the numbers with asterisks which are scored (6 - answer). Once summed, the larger (smaller) the score, the more liberal (conservative) is the person. See Hadsell, McAvoy and McGovern (2010, p14).

SUNY Oneonta offers the Senior Seminar in Economics, a capstone course for the economics major as well as the finance minor. The course is offered during the spring semester and seniors commonly enroll. These students' views are of interest as they have completed several upper division economics and/or finance courses and are relatively knowledgeable about current events. Students are required to read *The New York Times* daily, prepare summary logs, and participate in discussions of current events.

Nineteen students completed the ideology survey. Of those students that reported their gender, 26 percent are female, and 74 percent are male. The average ideology score is 37.02, meaning the average student respondent scores slightly above the midpoint of 36. The survey requests respondents to provide six additional data observations, one of which is a self-rating. Four students identified themselves as conservative, while eleven students self-identified as liberal, the other four reporting were neither. The survey indicates that six students are conservative or slightly conservative, four students are indifferent (neither slightly liberal nor conservative), and the other nine students are slightly liberal to liberal. Therefore, while this sample is small, the survey average slightly understates liberal self-identification. Hadsell notes that their survey measurement is not intended to predict a respondent's ideology; instead, the instrument is intended to spark interest in economic principles and understand better the role values play in modeling the behavior observed in the world and the resultant policies favored. Nevertheless, given Leonhardt's predictions, our seniors should prefer tax increases to spending cuts, on average.

Seventeen students submitted *Times* deficit reduction choices. These choices are presented in Tables 3 and 4. While the students had to submit at least \$1,345 billion as tax increases and spending cuts, on average, they reduced the deficit by a larger average amount, \$1,463 billion, of which \$862 billion (58.9 percent of selected options) represent spending reductions, and \$601 billion tax revenue increases (similarly 41.1 percent). At first blush, our typical liberal leaning seminar student generally prefers spending cuts. Spending cuts represented 47.6 percent of the total cuts possible, while tax increases 30.7 percent of the total possible. However, the options include instructions restricting some choices. When these choices are applied given the instructions, the outcome strengthens the conservative leaning choices our self-identified average liberal students make. Under the adjusted proportions, students select about 54 percent of the spending cuts and 37 percent of the tax increases. On the surface, at least, our ideologically measured and self-identified liberal leaning students were more willing to reduce spending than increase taxes.

Table 3: Students' Spending Reduction Choices, \$ Billions

SPENDING OPTIONS (1)	Student Average Spending Cut (n=17) (2)	Percent Students Choosing Option (n = 17) (3)	Students Scoring Conservative Ideology, Averages, (n = 6) (4)	Students Scoring Liberal Ideology, Averages, (n = 9) (5)
DOMESTIC PROGRAMS				
Farm Subsidies	\$4.24	29.41%	\$7.33	\$2.80
Foreign Aid	10.76	64.71%	7.83	11.90
Earmarks	9.24	64.71%	9.83	8.40
Federal Workforce	2.65	17.65%	5.00	1.50
Government contractors	5.65	35.29%	10.33	3.40
Cut Pay 5 %	5.00	29.41%	5.67	3.40
Other Cuts	10.59	35.29%	5.00	9.00
Aid to States	14.59	35.29%	27.33	8.40
MILITARY				
Cancel Weapons	8.59	47.06%	6.33	7.20
Reduce Fleets	7.12	29.41%	4.17	9.60
Nuclear Arsenal	20.24	52.94%	6.67**	26.60**
Reduce Active Personnel	37.53	76.47%	24.67	39.20
Reduce Noncombat Comp	17.94	35.29%	16.83	20.40
Withdraw IR/AF by 2015, or (a)	61.47	41.18%	99.67	29.80
Withdraw IR/AF by 2013 (a)	69.65	41.18%	28.33*	101.40*
HEALTH				
Malpractice Reform	4.94	35.29%	7.17	2.80
Medicare eligibility age 68	23.00	41.18%	18.50	22.40
Medicare eligibility age 70	30.65	29.41%	34.83	31.20
Cut Medicare Growth	330.24	58.82%	373.67	281.00
SOCIAL SECURITY				
Disability Insurance	6.76	41.18%	10.67	3.40
Reduce benefits for upper income	25.41	47.06%	18.00	32.40
Alternate inflation measure	33.65	41.18%	27.00	32.80
Raise retirement age to 68, or (a)	20.82	29.41%	23.50	21.30
Raise retirement age to 70 (a)	101.35	41.18%	163.67	49.40
TOTAL SPENDING CUTS	\$862.06		\$942.00	\$759.70

a: The choice is one of the adjacent options or neither, * the difference is statistically significant at 10%, ** 5%

Table 4: Students' Tax Increase Choices, \$ Billions

REVENUE OPTIONS (1)	Student Average Tax Increase (n = 17) (2)	Percent Students Choosing Option (n = 17) (3)	Students Scoring Conservative Ideology, Averages (n = 6) (4)	Students Scoring Liberal Ideology, Averages (n = 9) (5)
ESTATE TAX (Pick one)				
Exempt 1st \$5 million 36% Rate	\$1.18	5.88%	\$0.00	\$2.00
Exempt 1st \$3.5 million 45% Rate	13.24	29.41%	15.00	9.00
Exempt 1st \$1million	42.88	41.18%	34.83	41.60
INVESTMENT TAXES (Pick one)				
Raise capital gains tax on high income	7.24	29.41%	12.50	4.80
Raise capital gains tax on all	13.53	29.41%	7.67	9.20
INCOME and EMPLOYER TAXES				
Surtax on Income > \$1million	50.29	52.94%	31.67	57.00
SS Tax applied to income>\$106,800	23.53	23.53%	16.67	30.00
Reduce health insurance tax break	36.82	23.53%	52.00	15.70
Expiration of BUSH cuts >\$250,000 (a)	47.35	41.18%	76.67	34.50
Expiration of BUSH cuts <\$250,000 (a)	14.82	5.88%	0.00	0.00
TAX REFORM (Pick one)				
Reduce mort. deduct. for high income	19.12	35.29%	18.17	21.60
Eliminate breaks cut rates (SIMP)	82.35	47.06%	116.67	52.50
Eliminate breaks cut rates less	74.12	23.53%	0.00	126.00
OTHER NEW TAXES				
Carbon Tax	41.65	58.82%	35.17	35.50
Tax on banks	66.41	64.71%	68.00	61.80
National Sales Tax	66.06	23.53%	46.67	56.20
TOTAL TAX INCREASES	\$600.59		\$531.67	\$557.40

a: The choice is pick one of the adjacent options or neither.

Similar to *Times* readers, Table 3 shows 76 percent of our students prefer to reduce the size of the military, whether they prefer spending cuts or tax increases. In addition, about 64 percent want to cut foreign aid and reduce earmarks. A majority of the students would reform both Medicare and Social Security, primarily by increasing the age of eligibility to at least 68 years. In terms of revenue increases, Table 4 shows at least one-half of the students wants a surtax on incomes above \$1 million. The students' least popular tax increase, selected by one student, is to allow the Bush tax cuts to expire on incomes below \$250,000, matching the smaller than 10 percent selection of this option by *Times* readers

(Leonhardt 2010d). Our students on average reveal themselves to be slightly conservative, given their preferences for spending cuts. While increasing the age of eligibility for Medicare and Social Security may not be a liberal position, very good financial actuary reasons can support the selection of these options, independent of ideology. Furthermore, in cutting spending, our students very much prefer to target the military for cuts, a common liberal preference.

Cox presents the submissions by 6,989 Twitter users. Leonhardt (2010d) observes the online submitters via Twitter are likely younger. At least two-thirds of all Twitter users favor reducing the size of the military, ending the Bush tax cuts for \$250,000 plus earners, reducing Social Security benefits for high income earners, reducing nuclear arsenal and space spending, canceling or delaying some weapon programs, and imposing the millionaire's income tax. Cox further compares the deficit reduction choices by segmenting the entries into either those that reduced the deficit by more than via 75 percent tax increases (604 submissions), or by more than 75 percent via spending cuts (509 submissions). For those that prefer tax increases, at least two-thirds eliminate the Bush tax cuts on incomes of less than \$250,000, increase taxes on incomes in excess of \$1 million, tax carbon emissions and banks, increase the estate tax, increase the income subject to the Social Security payroll tax. For those Twitter users preferring spending cuts, at least two-thirds reduce the size of the military, reduce Social Security benefits, reduce the nuclear arsenal and space spending and eliminate or delay some weapons programs, reduce troops in Iraq and Afghanistan, eliminate farm subsidies, increase the age of eligibility and cap Medicare growth, enact medical malpractice reform, eliminate earmarks, and cut contractors, cut military compensation, both cut the pay and size of the civilian federal workforce, cut foreign aid and aid to states, and other cuts such as to national parks.

This study collects more precise information on ideology than Leonhardt (2010d) and Cox. Student ideology is directly measured with the survey tool developed by Hadsell. While Hadsell measures ideology along a continuum, they intend not to label specifically respondent's ideology, and prefer for their respondents to observe that one who scores low views behavior, models, and policies very differently than one who scores high. In this study, the midpoint of 36 is selected as a demarcation between liberal and conservative ideology measures. Ideology is labeled upon a specific, carefully considered, and ultimately arbitrary threshold. Students' answers for the ideology survey are linked to his or her choices for various budget reduction options. All seventeen students submitting reduction choices also self-identified on the submitted ideology surveys. Given the caveat of small sample sizes, as a group, they are liberal leaning: they self-identify an average 2.65 (neither to slightly liberal), and they score an average 37.00 (liberal). Five students self-identified as neither liberal nor conservative, and scored higher than 36 (liberal). Four students self-identified as slightly liberal or liberal, and scored 36 or below (neither or conservative). Two students self-identified as slightly conservative or conservative, and scored 36 or above (neither or liberal). Of the remaining six students, four students self-identified as slightly liberal or liberal and scored above 36, and two students self-identified as slightly conservative or conservative and

scored below 36. In sum, six students score below 36 (this study's separation score for "conservative" classification), and nine students score above 36 (this study's threshold score for "liberal" classification).

Among students scoring different than 36, their choices are shown in Tables 3 and 4, in columns 4 and 5. Students scoring conservative clearly favor reducing spending relative to students scoring liberal, and students scoring conservative increase taxes slightly lower than students scoring liberal. Keeping in mind the small sample sizes, students scoring conservative cut domestic programs more, reduce military spending less, reduce Medicare growth more, and are more willing to raise the age for Social Security benefits. In testing for differences in the means, two differences were statistically significant. Students scoring liberal are more willing to reduce the nuclear arsenal (5 percent level of significance) and more quickly draw down military forces in our Iraq and Afghanistan conflicts (10 percent level of significance). In Table 4, none of the differences in means is statistically significant. Amongst the tax increases, both sub-samples of students scoring conservative or liberal are shown to increase taxes in similar amounts, although the students scoring liberal increase taxes slightly more. The students scoring conservative increase taxes on wealthier and higher income households, but also are more likely to close loopholes and to lower rates. These students can be interpreted as broadening the tax base and treating all forms of income similarly. Students scoring liberal increase taxes by greater amounts on wealthier and higher income earners, lower rates less, and maintain popular deductions for middle income earners.

Some of the survey questions are directly related to the options in the deficit reduction project. For instance, the survey questions listed in Table 2 show numbers 2 and 9 are related to health care, number 3 promotes social programs in support of low income and other marginal groups, number 6 increases income taxes, 7 favors regulation, and 11 promotes the environment. On a continuum from 0 (strongly disagree) to 6 (strongly agree), the sub-sample average ideology scores for n2 is 2.29, number 3 is 3.71, number 6 is 3.29, number 7 is 2.47, number 9 is 3.71, and number 11 is 2.71. Tables 3 and 4 compare the reduction choices for fifteen students scoring other than 36. As previously stated, students with conservative scores favor spending cuts, while students with liberal scores slightly prefer more tax increases.

Comparing average ideology scores for health care to Medicare choices in the subsample, our students on average are moderately liberal (sub-sample average score is 3.71 on n9); however, they choose to reform Medicare eligibility by increasing the age and reducing spending growth. Their average choices for Medicare match closely their ideology for preferring not to regulate drug prices (average score is 2.29 on n2). Liberal scoring students favor a smaller cut in Medicare and are less in favor of raising the retirement age to 70. While students slightly agree that the less-fortunate must be assisted by the government (average score on n3 is 3.71), Tables 3 and 4 show they choose larger spending cuts than tax increases to reduce the deficit, no matter their ideology score. Consistent with popular views of partisan ideology, our liberal scoring students prefer smaller reductions to social programs relative to our conservative scoring students.

Comparing the scores in Table 4 to the tax increases selected in Table 3, the students slightly agree that higher incomes should be taxed more (n6 score is 3.29). Indeed, the students scoring liberal choose an average 63.52 percent of the \$315 billion available for tax reform, while students scoring conservative select 42.81 percent. Liberal scoring students choose a smaller national sales tax, conservative scores selecting an average 16.61 percent, and liberal scores selecting an average 20 percent, of the available \$315 billion tax increase. While the students slightly disagree that consumers need government protection (score on n7 is 2.47), their tax increase on banks is relatively high, conservative scores selecting an average 66.02 percent, and liberal scores selecting an average 60 percent of the \$103 billion available. The students slightly disagree that the government should direct resources to green technologies (score on n11 is 2.71); and, they select a moderate carbon tax increase, conservative scores selecting an average 49.52 percent and liberal scores 50 percent, from \$71 billion available. Consistent with popular partisan interpretation of current issues, liberal scoring students favor tax increases on wealthy or high income earners more than conservative scoring students, and they favor smaller tax increases on low income earners.

The ideology survey and *The New York Times* deficit reduction project continue to be relevant to current events. In combination, they offer the instructor an opportunity to increase student participation in class and well as their interest in important current issues. When matching the survey results to the options selected by those students, the instructor has an opportunity to discuss ideology, partisan views, and economic models with their assumptions. While more statistical results and conclusions would be ideal for a study of this type, the tentative results presented are promising, and a larger sample size or replicating this exercise frequently should present more conclusive evidence that the survey scores predict the economic choices students make. At the least, the ideology survey and *The New York Times* deficit reduction project offers students a chance to gain a greater understanding of the world around them.

REFERENCES

- Cox, Amanda, Matthew Ericson, David Leonhardt, Bill Marsh and Kevin Quealy. November 20, 2010. "How Twitter Users Balanced the Budget." *The New York Times*. <http://www.nytimes.com/interactive/2010/11/21/weekinreview/21leonhardt-graphic.html?ref=weekinreview>.
- Hadsell, Lester, Michael McAvoy and Jaime McGovern. September 9, 2010. "A Survey to Measure Economic Ideology in Principles Students." Working paper, .
- Leonhardt, David. 2010a. "Behind The Times's Deficit Project." *The New York Times*, Economix blog. November 13, 2010. <http://economix.blogs.nytimes.com/2010/11/13/behind-the-scenes-deficit-project.html>.
- Leonhardt, David. 2010b. "O.K., You Fix the Budget." November 13, 2010. *The New York Times*. <http://www.nytimes.com/2010/11/14/weekinreview.14leonhardt.html>

Leonhardt, David. 2010c. "How Readers Chose to Fix the Deficit." November 20, 2010. *The New York Times*. <http://www.nytimes.com/2010/11/21/weekinreview/21leonhardt.html>.

Leonhardt, David. 2010d. "7,000 Ways to Fix the Budget." November 21, 2010. *The New York Times*. <http://graphics8.nytimes.com/packages/pdf/weekinreview/20101121-deficit-graphic.pdf>.

Leonhardt, David and Bill Marsh. November 13, 2010. "Get a Pencil. You're Tackling the Deficit." *The New York Times*. <http://graphics8.nytimes.com/packages/pdf/weekinreview/20101114-deficit-graphic.pdf>.

A Merger Effect on Different Airline Groups : Empirical Study on the Delta-Northwest Merger in 2008.

Sora Park*

ABSTRACT

The airline industry has performed inconsistently over the past decades since deregulation in 1978. Primarily, there has been a significant profit discrepancy between legacy airlines and low-cost carriers. Starting from the early 2000s, four of the largest legacy airlines chose to consolidate with one another with an effort to increase efficiency and profitability. In particular, Delta Air Lines and Northwest Airlines have consolidated in 2008, becoming the largest commercial airline in the world. This paper examined the motivation behind this merger as well as the anticompetitiveness concerns arising from this merger. It also analyzed the merger effect on the airfares on top 1,000 U.S. domestic city-pair routes in relation to the number of passengers, the distance, and the market share. All airlines were categorized into three groups – Legacy Airlines, Delta-Northwest, and Low-Cost Carriers – and the interactions between the airfares and the variables were examined within each airline group and across different airline groups. It was found that the airfares of legacy airlines and Delta-Northwest decreased at a faster rate than the low-cost carriers while their market shares increased simultaneously irrespective of whether they offered the lowest prices or had the largest market share on each route.

INTRODUCTION

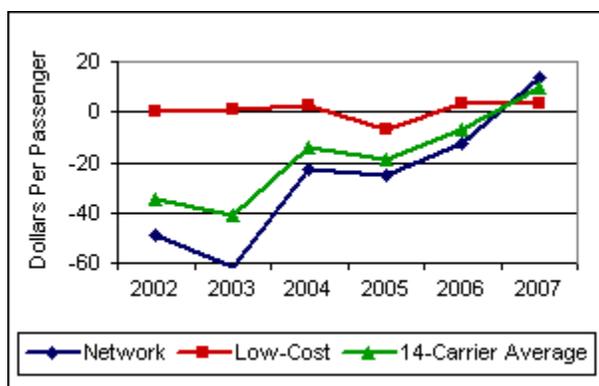
The U.S. domestic airline industry can be largely divided into two different groups; one is legacy airlines and the other one is Low-Cost Carriers (LCCs). In general, legacy airlines refer to those carriers who existed before the Airline Deregulation Act of 1978. As they operate on almost any domestic and international routes, they provide first and business class, airport lounges, and other high-level of services with various types of aircrafts. On the other hand, LCCs entered the market after deregulation and operate under the low cost structure that maximizes their profit by primarily serving direct flights on the denser routes at lower prices with fewer services.

The distinct difference in those two groups of airlines can also be found in the profit gap between them. Since deregulation in 1978, the airline industry has been volatile in terms of its financial stability, while the demand for air travel has steadily increased. There have been over 160 airline bankruptcies since 1978, yet particularly, the largest airline bankruptcies took place in the early 2000's; US Airways and United Airlines entered Chapter 11 bankruptcy in 2002, Delta Airlines and Northwest Airlines in 2005. Besides bankruptcies, the legacy airlines alone incurred nearly \$25 billion in operating losses from 2000 to 2003, while LCCs gained \$1.3 billion in profit during the same time period (US GAO, 2004). Figure 1

* University at Buffalo- the State University of New York, Buffalo, NY 14260.

shows the operating loss per originating passenger was more than \$60 for legacy airline in 2003, whereas the operating profit per passenger was positive for LCCs at the same time.¹ In the meantime, mergers and acquisitions in the airline industry have increased considerably. Starting with the American Airline's acquisition of TWA in 2001, America West acquired US Airways in 2005, Delta Air Lines acquired Northwest Airline in 2008, and United Airline consolidated with Continental Airline as a way to increase their profitability through consolidation. Excluding the merger between America West and US Airways in which US airways announced to become a low-cost carrier from a legacy airline, the rest of mergers were between legacy airlines. Then the question arises as to whether legacy airlines could be able to turn around its financial difficulty by recent merger trends. In fact, Figure 1 shows that the profit gap between the legacy airlines and the LCCs became narrower since 2005.² The operating loss eventually turned to a positive profit for legacy airlines by the beginning of 2007 and it surpassed the operating profit of the LCCs, confirming the positive effects from the mergers.

Figure 1: Performance Measure in the Airline Industry: System Operating Profit or Loss per Originating Passenger



The merger and acquisition was considered inevitable to minimize asset devaluation in order prevent any domino effect in the airline industry as most airlines at issue were “too big to fail” and because the industry itself was widely vertically-integrated and interdependent to each other. However, the flood of horizontal merger could impose a threat to the incumbents and consumers due to the likelihood that the integration of major airlines with long-established networks may dominate the market and less the competitiveness of the airline industry. Nonetheless, the United States Department of Justice granted immunity from antitrust issues for most airline mergers in the 2000's, and four of the largest legacy airlines chose to consolidate with one another while the efficiency and competitive effects of mergers remained controversial.³

As the airline industry has undergone several transformations by eliminating inefficient operators and has consolidated its excess capacity by mergers and acquisitions, it was imperative to evaluate the merger effect on the merged airlines and the non-merger airlines. Although there have been extensive studies about price effects after mergers in the 1980s, there have not been enough studies about the

recent mergers and the corresponding changes in the performance of merging firms in relations to those of other legacy airlines and low-cost carriers. Thus, the empirical results on the current merger will provide an opportunity to find the changes within each airline and across different airlines as well as the general outcomes of the consolidation in the airline industry.

This paper examined the empirical evidence on the merger effect of Delta-Northwest. The Delta-Northwest merger in 2008 was an unprecedented event where two of the largest legacy airlines consolidated with antitrust immunity, although the probability of potential anticompetitive effects remained high. The significance of this event arose not only from the fact that the merging firm could harm the other airlines and consumers significantly through its combined market power, but also from the speculations on merger benefits. By categorizing the U.S. domestic airlines into three different groups and comparing their performances two quarters before and after merger in the course of four years, this paper analyzed the merger effect on airfares in terms of the changes in the number of passenger, flight distances, and market shares on top 1,000 U.S. domestic city-pair routes. In addition, this paper investigated the antitrust issue associated with the potential anticompetitiveness of the proposed merger between Delta and Northwest.

LITERATURE REVIEW

There have been numerous studies devoted to analyze the effects of mergers by using panel data. Lichtenberg and Kim (1989) compared the performance of merged airlines by analyzing load factor, flight duration, and other input prices from five different mergers that occurred in the 1970's and the 1980's. Morrison (1996) specifically looked at the two different mergers between Northwest Airlines and Republic Airlines in 1986 and between US Air and Piedmont Aviation in 1987 and found that fares on those routes affected by the mergers jumped as high as 30 percent. Peters (2003) predicted substantial increase in post-merger prices on overlap routes of merged airlines and compared them to actual post-merger prices from six major airline mergers in the 1980's. Berry and Jia (2009) found that the air travel demand became more price sensitive in 2006 than that of the 1990's. Also, they found that travelers put more emphasis on direct flights than hub airports, which explained the legacy airline's consistent losses over the past decade. Overall, there has been extensive analysis supporting the price increase after the merger.

On the other hand, Ito and Lee (2003) identified the market characteristics of low cost carriers. LCCs typically competed in the markets where passenger density was high enough for them to exploit their competitive advantage by providing the quick-return point-to-point service. In particular, when LCC's networks were not large enough to generate more capacity in the late 1990's and the early 2000's, they had no choice but to choose denser markets in which they serve non-stop routes for local passengers. Looking at Ito and Lee's data on the number of the largest 1,000 domestic city-pair markets in 1990, low-cost carriers only served 78 city-pairs whereas they now serve 613 city-pairs on the latest data in the third quarter of 2010. As LCCs' networks expanded throughout the U.S. domestic market, Tan (2011) found that legacy airlines lowered their fares before and after LCC's entry while LCCs did not change their

fares. Although LCCs could be disadvantageous upon entering the market due to entry barriers such as legacy airlines' dominance at airport gates and landing/take-off slot controls, LCC's low cost structure enabled them to charge lower fares. Once highly concentrated markets by legacy airlines faced the threat of price competition, and the market competitiveness has been improved where low-cost carriers successfully competed against legacy airlines.

In summary, there were two different evidences at work; mergers made airfares to increase and LCCs' presence made legacy airlines to lower their fares. If those two factors are present, how will the merger between legacy airlines affect LCCs? On the other hand, will legacy airlines have to alter their pricing strategy when they face competition from LCCs?

BACKGROUND

(I) WHY DO AIRLINE GO INTO THE M&A PROCESS?

The primary purpose of mergers and acquisition is to enable airlines to eliminate duplicative operating costs, therefore reducing their total costs on labor, service, and operations. It also enables the merging firms to integrate their financial assets, facility, and technology. As a result, airlines look for mergers and acquisitions in pursuit of both reducing operating costs and maximizing revenue. For example, if two airlines with the most market overlap merge to one firm, the cost saving from eliminating the duplicate operations would generate a significant cost saving effect. Especially when the market demand is at the low, the need for mergers and acquisitions tends to increase. Experts say that the motivation behind industry mergers and acquisition is largely based on the airlines' need for the combined market share at the integrating airlines' major hubs, because it can maximize revenue (US GAO, 2010). While the presence of low-cost airlines may bring down airfares in the market, a dominance of merging firms through integration could lock in their profit against cost competition. Furthermore, airlines also seek to merge with one another as a means to expand their networks. As a way to compete with legacy airlines in the wider markets, LCCs now expand to the international market through strategic alliance, code sharing, or even through a merger with other LCCs.

In the U.S. domestic market after deregulation in 1978, price dispersions and excessive competition over capacity intensified, and it led to many airlines' financial failures; the largest five legacy airlines filed Chapter 11 bankruptcy and only 9 of the 23 legacy airlines served the U.S. domestic market by 2007. (Borenstein and Rose, 2007) Besides, legacy airlines lost nearly \$33 billion from 2001 to 2005, whereas LCCs remained profitable regardless of increasing cost of fuel. In accordance with the changes in the market structure followed by financial reorganization and exit, mergers and acquisitions increased ever more since the early 1980's until today. Specifically from the beginning of 2001, American Airline acquired Trans World Airlines, US Airways acquired America West in 2005, Delta merged with Northwest in 2008, and Continental merged with United in 2010.

(ii) Delta-Northwest Merger

On April 14, 2008, Delta and Northwest Airlines announced a \$17.6 billion merger, which would make them to be the largest commercial airline in the world. Both of them were legacy carriers which existed prior to the 1978 Deregulation Act. Also, they operated under the hub-and-spoke system in which passengers have to go through a hub airport in order to transfer to other connecting flights. In 2006, Delta was the second largest airline in the top 5,000 markets in terms of the number of markets served and the number of passengers. Northwest was the fourth largest, while Southwest being in the first place and American Airline in the third. On September 26, 2008, two airlines' shareholders approved the merger and Antitrust Division at the Justice Department approved the merger on October 29, 2008. As of December 2009, Northwest's aircrafts have operated under Delta, and Northwest's hubs have been fully consolidated with Delta's brand.

(iii) Antitrust and Anticompetitiveness Issues

As discussed above, airlines merge or acquire other airlines to reduce costs on duplicate operations and consolidate their financial stability, but the proposed merger and acquisition needs to be approved by the US Department of Justice (DOJ) as required by the Hart-Scott-Rodino Act. This act requires an acquisition of assets be reported to the DOJ and its Antitrust Division examines and determines whether the merger is likely to have an adverse effect on consumers by following their analytical framework called *Horizontal Merger Guidelines*. Under these guidelines, the US DOJ defines the relevant product market and geographic market where the merging firms operate and determines the likelihood of increased concentration ratio in these relevant markets.

Defining a relevant geographic market could be ambiguous depending on the industry, products, and geographic areas. Generally, a relevant market indicates products and their geographic area where they are produced and sold by a monopolist. The monopolist in this market is likely to have the power to increase his price above the competitive price due to the market conditions such as the inelastic demand or insubstitutability of the product. In the airline industry, each city-pair market is recognized as a separate geographic market and the nonstop flights can be considered as separate products because they are an equivalent substitute for connecting flights (Goetz, 2002). For antitrust purposes, a relevant market is defined as nonstop service in a city-pair market (Klein and et al, 1999). Thus, a relevant market in reviewing a potential merger constitutes an aggregate of each city-pair market in which the merging firms operate. The next stage of evaluating the potential merger is to quantify the market power and the concentration ratio of merging firms and their likelihood of potential anticompetitive effects in the market. Market power is measured by market share, and airlines with the larger market share in the relevant market is likely to extract fare premiums because they have established dominance over airport gate assignment, inelastic demand in their city-pair routes, and such. The higher the market share a firm has, the more the market is concentrated. Thus the relevant market is likely to be shared by only few firms, and it gives them an incentive to engage in anticompetitive conduct and fully exploit their monopoly

power. According to the US GAO, a concentrated market is where a single airline accounts for at least 60% of the passenger enplanement or two airlines with 85% (US GAO, 1993). They found that average fares tend to be higher in the concentrated markets. The US DOT also identified a concentrated market with a single firm accounts for 70% of the market (US DOT, 1990). The Herfindahl Hirschman Index (HHI) is considered as a good measure for concentration ratio as well for evaluating horizontal mergers, which will be discussed later in this paper.

Following the *Horizontal Merger Guidelines*, the US DOJ analyzes other anticompetitive effects on competition in the corresponding industry and the likelihood of failure of merger proposing firms in the absence of a merger, which is weighed against the potential consumer benefits from expanded service of the merged firm. The recent cases of the Delta-Northwest merger in 2008 and the United-Continental merger in 2010, the US DOJ has granted full approval, deciding that the potential benefits for consumers will outweigh the potential anticompetitive effects in the airline industry. According to the public announcement from DOJ on October 28, 2010 in closing their investigation in the proposed merger of Delta and Northwest, they stated that "...the merger likely will result in efficiencies such as cost savings in airport operations, information technology, supply chain economics...consumers are also likely to benefit from improved service...by combining under single ownership..."⁴

The financial viability of the airline industry has been a significant issue for the industry itself as much as its potential negative impacts on consumers and the government in the forms of increased airfares, unemployment, and federal financial assistance. While the airline industry went through numerous bankruptcies and mergers since deregulation, new airlines and existing airlines continued to increase their capacity, only to exacerbate reduce their revenues. Past mergers and acquisition were intended to reduce this excess capacity in effort to improve their profitability, yet the reduction in industry capacity was temporary due to the expansion of existing and new airlines. Thus the rationale behind the merger approval is to enable the larger airlines to combine their duplicative operations on overlapping routes so as to reduce excess capacity, achieve cost savings, and ultimately generate more revenues. Besides the fact that the integration of computer systems, airline fleets, facilities, and human capital of two different airlines is not an easy task, the integration process may face challenges from the consumer side. Capacity reduction poses a potential price increase issue in certain markets after a merger.

The major difference between low-cost carrier's operations and legacy airlines' operations is the network structure. Legacy airlines use hub-and-spoke networks where passengers fly through a hub in order to get to their final destination, while LCCs use a point-to-point system where they serve non-stop flights between denser airports. LCCs enjoy their profitability by focusing on economies of scale from traffic density and they have less motivation to enter the market with the lower passenger traffic. On the other hand, legacy airlines expand their networks due to their cost advantage on economics of scope and they are likely to have more incentives to serve the markets where they do not face a competition from LCCs. In fact, some studies showed that airfares increased on some markets after mergers and acquisition in the 1980s (Morrison, 1996). A few other studies showed that dominant airline at an airport

charged higher fares due to the barriers to entry (Borenstein, 2005). In addition, a recent study found that Legacy Airlines reduced their fares upon the entry of LCCS in the market, whereas the LCCs did not alter their pricing. *Overall, cost reduction could provide the merging Legacy Airlines an increased revenue flow at the expense of consumers.

The U.S. airline industry has undergone many changes since the deregulation. While the airlines took the freedom of adopting their most efficient pricing without the previous restrictions on their operations on specific routes or airports, it essentially led to them to engage in predatory pricing in a way to expand their travelers, some airlines charged higher fares for business passengers while this profit was subsidized for the loss from heavily discounted fares. Many studies acknowledged this fare dispersion and found that smaller routes with less competitive alternatives tend to have higher rates compared to larger routes or airports (US GAO, 1996a).

Along with the pricing dynamics, complaints of predatory pricing indicated aggressive competition that has been persistent in the U.S. airline industry. There have been several allegations that legacy airlines were involved in predatory pricing behavior in which they sharply reduce their fares and increase their capacity in an effort to eliminate the new entrants with an expectation of recouping the losses by charging consumers with supra-competitive prices once the entrants were driven out of the market. In fact, the U.S. DOT received 32 complaints alleging anticompetitive conduct such as unfair pricing, capacity responses, and barriers to entry from 1993 to 1999 (Transportation Research Board, 1999). Nonetheless, U.S. District Court ruled in favor of American Airlines and Northwest Airlines against Vanguard Airlines and Reno Airlines respectively, reasoning that the actions taken by those legacy airlines were proper alternative response with sufficient justification (US GAO, 2008).

The US DOJ in 2008 published a report on the potential effects of mergers and acquisition in the US domestic airline industry, illuminating on the inevitability of financial consolidation amongst the airlines and evaluating the concerns with the Delta-Northwest merger (US GAO, 2008). They examines the extent of efficiencies arising from the route expansion and cost saving from the merger, which was weighed against the financial standing of merging firms as the potential failure of one of the merging firms would substantially depreciate its assets and harm the industry. They also evaluated the probability of merger partners facilitating its enhanced market power after the proposed merger. Additionally, they accentuated the changes in the airline market structure due to the entry of LCCs in recent years, and proposed that competition would not subside simply by consolidation of legacy airlines as opposed to the entry of a new carrier. They looked at several studies on how the presence of LCCs on routes mitigated the dominance of legacy airlines and spurred the pricing competition. They found that the hub dominance was not correlated to a gain in market power (Borenstein, 2005). To determine the competitiveness of the U.S. airline industry and its resilience to the proposed merger, US DOJ also examined the possible changes in passenger traffic, market share, and the number of competitors on routes where merger partners operate as well as financial conditions of the merger partners.

(IV) MERGER ANALYSIS

As previously discussed, the U.S. domestic airline industry had a history of new entry and consolidation. The four-firm ratio was 46% from 1978 and 69% from 1985 to 1991, while it was 56% in 2000 (Goetz, 2002). Using the data from the Bureau of Transportation Statistics (BTS), the four-firm ratio was 41% in 2006 with a slight increase at 46% in 2010.⁵ While the concentration ratio has decreased consistently from the late 1980's, it has shown an increase during the Delta-Northwest merger period although a subsequent study is needed in the future to see whether this is a temporary effect from the merger or the permanent integration in the U.S. market.

In evaluating horizontal mergers, the HHI index of market concentration was calculated by adding the squares of the market shares of each firm in the relevant market. These merger partners' hub airports were chosen for the HHI calculation in order to analyze any changes in market share and concentration ratio during the merger period from 2008 to 2010. If the HHI is below 1000, the market was unconcentrated, while HHI above 1800 indicated a high level of concentration. As shown in Table 1, both airlines' hub airports had the HHI greater than 1800 with the market share as high as 66%, indicating the already established high concentration. Atlanta airports showed an increase in both the market share and the HHI, implying the merger enhanced the merger partners' market power. On the other hand, Detroit and Minneapolis showed a significant reduction in the market share from approximately 61% to 49% and 66% to 53% respectively. The HHI at those airports dropped simultaneously, while the numbers are still well above the threshold for the high level of concentration. To the contrary, New York J.F.K airport seemed to have an ambiguous effect from the merger; the market share increased while the HHI decreased substantially. This could imply that the merger had a positive effect on the market share of the merger partner, whereas it subsequently lowered the other airlines' relatively large market share at the same airport.

Table 1

	2006		2007		2009		2010	
Delta-Northwest Hub Airport	market shares	HHI Index	market shares	HHI Index	market shares	HHI Index	market shares	HHI Index
Atlanta (ATL)	52.76%	3825	58.06%	3681	53.42%	3641	56.85%	3945
Detroit (DTW)	60.98%	4128	59.32%	3974	49.94%	3219	48.90%	3157
New York (JFK)	14.91%	3090	15.33%	2990	19.81%	2572	21.65%	2601
Minneapolis- St.Paul (MSP)	66.17%	4711	65.89%	4684	52.93%	3562	52.89%	3680
Salt Lake City (SLT)	40.48%	2690	39.79%	2884	38.85%	2895	40.70%	2940

In comparison to individual airport markets, city-pair markets weigh more importance, because the city-pair routes are defined as relevant markets with regard to the antitrust issues. Before analyzing the relationship between the air fare and the market share in the regression analysis later in this paper, the changes in Delta-Northwest's market share on several top 1,000 U.S. domestic city-pair routes were summarized in Table 2. The sample city-pair routes where Delta-Northwest had the largest market share were selected and compared between the pre-merger and the post-merger time period. The numbers in parentheses presented the airlines and their market which provided the lowest fare on the corresponding routes where Delta-Northwest had the largest market share. Interestingly, Delta-Northwest's market share had not changed from 2006 to 2010 in the New York-Washington D.C city pair market, yet there was a quite change in airfare. Before merger, Delta had the largest market share on this route while it also provided the lowest fare. However, United Airline replaced Delta's place after the merger, providing the lowest fare on the route. On the other hand, the Atlanta-Salt Lake City city-pair market showed an increase in Delta-Northwest's market share from 67 to 76 percent while the market share of the lowest fare provider, Continental Airlines, decreased substantially from 3 to 1.5 percent. This could imply that the merger enabled Delta and Northwest to expand their market share on this route and probably engage in price competition against Continental Airlines.

Several studies showed a significant increase in the HHI on the city-pair routes. For the routes between the cities with large populations, merger partners tended to dominate on these routes after consolidation. Combining the existent market share alone gave much more market power for the merger firms, and the actual data in the late 1990's showed that the combined concentration was increased to 70 percent of nonstop traffic on the corresponding routes (Goetz, 2002). Nevertheless, antitrust scrutiny involves a multilateral analysis that assesses the effects of the proposed merger at the national level, the individual airport level, and the city-pair routes although this may give contrasting results depending on the definition of the relevant market.

There are several studies on how airlines have influenced the fares of the other airlines by exercising increased market power through mergers and acquisitions. Some found that merger-affected routes showed a significant price increase while this increase was correlated with a gain in concentration (Kim and Singal, 1993). The enhanced market power could appear in the form of predatory pricing in which the merging firm enjoys economies of scale and drove out its competitors by reducing their airfares below the competitor's prices. On the contrary, the merging firm could utilize its enhanced market share against the non-merger competitors and increase its price, taking advantage of its network expansion and the corresponding consumers' preference. Particularly if the merging firms had solid market share at smaller airports and on unpopulated city-pair routes even before the merger, the integration process between the merging partners was likely to give them dominance in these markets. Furthermore, as the merging firms had a control over the airfares on these markets, the merger may led to a tacit collusion between the merging firms and non-merger firms in which they implicitly agreed or followed the pricing behavior of the leading firm, primarily in the form of higher prices.

DATA AND METHODOLOGY

The purpose of this paper was to analyze the merging airline's pricing behavior compared to other airlines during pre- and post-merger periods, especially on whether this merger had altered the market structure and to what extent the merger affected different airline. To examine the presence of these merger impacts, the changes in operations of legacy airlines and LCCs were compared to that of Delta-Northwest along with the changes in the consumer demand which was measured by the daily non-stop passenger traffic. A potential change may include an aggressive pricing strategy of the merging firm which would coincide with a gain in its market power in the corresponding city-pair routes. While legacy airlines in general were unprofitable in contrast to the LCC's lower operating cost structure, a merger between the large legacy airlines could have enabled them to sufficiently reduce their fares through economies of scale from the merger. On the other hand, the merger between these two largest legacy airlines could have disincentivized other airlines' competitive pricing due to the reduced number of competitors in the same market, thus it could have led to higher fares after the merger. While many studies have found the significant correlation between a price increase and an airline merger, specifically on the routes or in the airport markets where the merging airline directly faced competition from other airlines, this paper examined the comprehensive merger effect and its trend in the market regardless of whether or not the merging firm operated in the corresponding market. In other words, this paper investigated the general influence of the Delta-Northwest merger in the U.S. domestic airline market, rather than comparing the Delta-Northwest operating routes only.

The top 1,000 largest U.S. domestic city-pair markets have been collected to analyze the Delta-Northwest merger effect on the major, dense domestic routes. The data was available from the Office of Aviation Analysis under the Department of Transportation (U.S. DOT). The Department of Transportation releases an Air Travel Consumer Report each month to inform consumers about service quality of each airline. As part of this effort, the Office of Aviation Analysis publishes a quarterly report to provide average fare information in the 1,000 largest domestic city-pair markets, which accounts for 75 percent of total passengers in the 48 contiguous states and 70 percent of total domestic passengers. This report includes the origin city, the destination city, the number of one-way passenger per day on each city-pair route, the non-stop distance, the average market fare per mile, the airlines with the largest market share on each route and their fares, and the airlines with the lowest airfare and their market share.⁶ The extended version of top 1,000 city-pair routes was available only on the U.S. DOT website, which listed all city-pair routes with average at least 10 passengers each day where either the origin city or destination city is a hub and has more than one airport. To minimize the different characteristics of each city-pair routes, only duplicate routes were selected across 2006, 2007, 2009, and 2010. Thus, the same routes across different years were compared, and this eliminates the variations across the selected routes and enables a direct comparison between other variables such as airfares and the consumer demand. In order to provide a direct test on merger effects without the deterrence from the trending and seasonality issues, the third quarter of each year is compared for the merger period. Since the Delta merger took place in

April 2008 and it has completed in October 2008, two quarters before and after the third quarter of 2008 have been gathered for data analysis. The third quarter of 2006 and 2007 are recognized as a “before-merger” period and 3rd quarter of 2009 and 2010 as “after-merger” period.

The effect of merger could lead to different results owing to the different operating structure of each airline. To compare the Delta-Northwest merger to non-merger legacy airlines and LCCs, airlines were categorized into three different groups.⁷ The first group was low-cost Carriers, the second was a group of three legacy airlines which did not merge or acquire during the Delta-Northwest merger.⁸ Last group consists of Delta and Northwest. If Northwest was identified on quarterly reports prior to the merger, it is grouped under Delta Airline. Price dispersion amongst different airlines was extensive depending on the extent of competition, market power, and the size of the market. Therefore, the average fares might severely distort the true information on fares. Instead, the lowest fares were compared to see whether they had a significant relationship with each airline group in the post-merger period. In addition, the fares of the largest carrier on each route were weighted against other market variables so as to analyze the connection between market power and the changes in fares. Table 3 descriptive statistics summarized the data. The total of 4002 city-pair routes was collected during the 2006-2007 and 2009-2010. The mean fare of the airlines with the largest market share was \$180.93, implying that these airlines with market power charged a higher price than the average at \$178.56 whereas the mean of the lowest fare was \$156.24. On the other hand, the mean of the largest market share was 60.14 percent while the mean market share of the lowest fare is 37.14. In the market with the lowest fare, LCCs accounts for 74 percent, while Delta contributed only 11 percent. In the market with the largest market share, LCCs still accounted the majority at 57 percent, while legacy served 24 percent of the market and Delta-Northwest served 18 percent of the market.

Table 3 descriptive statistics: Descriptive Statistics on the airlines with the largest market share on top 5,000 U.S. domestic markets (n=4002)

	Minimum	Maximum	Mean	Standard Deviation
Distance	129	2724	1081.80	653.292
Daily passenger	199	9910	771.79	854.279
Fare of airlines with largest market share	70.96	452.58	180.93	63
LCC dummy	0	1	.57	.496
Legacy dummy	0	1	.24	.427
Delta dummy	0	1	.18	.381
Merger dummy	0	1	.50	.500
Average fare	76.04	426.11	178.56	55.540
Largest market share on the route	16.64	100	60.14	19.70

Table 4: Descriptive Statistics on the airlines with the lowest fare on top 5,000 U.S. domestic markets (n=4002)

	Minimum	Maximum	Mean	Standard Deviation
Distance	129	2724	1081.80	653.30
Daily passenger	199	9910	771.79	854.28
Average fare	76.04	426.11	178.56	55.54
Market share of lowest-fare airlines	1.02	100	37.14	27.98
LCC dummy	0	11	.74	.470
Legacy dummy	0	1	.15	.356
Delta dummy	0	1	.11	.318
Merger dummy	0	1	.50	.50
Lowest fare	66.48	380.27	156.24	47

The dataset available at U.S. DOT was organized in two different categories⁹; one with the airlines with the lowest fares and the corresponding number of passenger, the distance, and the market share; another one was the same information on the airlines with the highest market share on each city-pair routes. As discussed before, the average fares were disregarded due to the high level of price dispersion in the airline industry. To identify changes on the airfares in relation to the changes in other available variables, the following expression was used for each group.

$$(1) \text{Fare}_{\text{low}} = \alpha_0 + \beta_1 \text{distance} + \gamma_1 \text{pax} + \delta_1 \text{mksh} + D_m + \varepsilon$$

$$(2) \text{Fare}_{\text{high}} = \alpha_1 + \beta_2 \text{distance} + \gamma_2 \text{pax} + \delta_2 \text{mksh} + D_m + \varepsilon$$

The variable Fare_{low} was the lowest fare in the first dataset, while $\text{Fare}_{\text{high}}$ was the fare of the airlines with the highest market share in the second dataset. The regressor *distance* was the distance between the two different cities of each city-pair routes, *pax* was the number of average daily passengers on each route, *mksh* was the market share of each corresponding airlines on each route. The last variable D_m was a dummy variable that takes the value 1 for the post-merger period from 2009 to 2010, while it was 0 for the pre-merger period from 2006 to 2007. The expressions above examines whether or not the Delta-Northwest merger led to a change in airfares of non-merging airline groups relative to that of Delta-Northwest.

While the above expressions examined the merger effect within each airline group, another term was added to analyze the aggregate merger effect among different airline groups.

$$(1) \text{Fare}_{\text{low}} = \alpha_0 + \beta_1 \text{distance} + \beta_2 \text{pax} + \beta_3 \text{mksh} + \delta_{\text{delm}} D_{\text{delm}} + \delta_{\text{ccm}} D_{\text{ccm}} + \delta_{\text{legm}} D_{\text{legm}} + \gamma_1 (\text{mksh} * D_{\text{delm}}) + \gamma_2 (\text{mksh} * D_{\text{ccm}}) + \gamma_3 (\text{mksh} * D_{\text{legm}}) + \varepsilon$$

$$(2) \text{Fare}_{\text{high}} = \alpha_1 + \beta_4 \text{distance} + \beta_5 \text{pax} + \beta_6 \text{mksh} + \delta_{\text{delm}} D_{\text{delm}} + \delta_{\text{ccm}} D_{\text{ccm}} + \delta_{\text{legm}} D_{\text{legm}} + \gamma_4 (\text{mksh} * D_{\text{delm}}) + \gamma_5 (\text{mksh} * D_{\text{ccm}}) + \gamma_6 (\text{mksh} * D_{\text{legm}}) + \varepsilon$$

Again, the same term was applied to two different sets of data; one organized by the lowest fare and another one organized by the largest market share. Here the total population of three airline groups was classified into six groups: before merger low-cost carrier, after merger low-cost carrier, before merger legacy airlines, after merger legacy airline, before merger Delta-Northwest, after merger Delta-Northwest. The corresponding dummy variables were taken for each groups: D_{delm} took the value of 1 if the data is for Delta-Northwest after merger or 0 for before merger, D_{legm} took 1 for post-merger legacy airlines or 0 for pre-merger legacy airlines, D_{ccm} took 1 for post-merger low-cost carriers or 0 for pre-merger low-cost carriers. Consequently, α_0 was the intercept for the base case for Delta-Northwest in the markets with the lowest fare. In addition, the interaction terms were added to test the change in each airline group's market share conditional on the existence of merger and whether it could be related to the change in the airfare. An increase in the fare with the effect of low-cost carriers after merger was likely to imply that the airfares of non-merging firms are positively affected by the Delta-Northwest merger. Likewise, if a decrease in the fare corresponded with a decrease in the market share of legacy airlines after merger, it will indicate that the airfares and the market share of the non-merging firms are negatively affected by the merger. Thus, this test will examine Delta-Northwest's own price change after its merger as well as whether its merger influenced the airfares of non-merging firms in terms of the distance, the number of passengers, and the market share.

RESULTS

Table 5 shows the changes in airfare for each airline group in relation to the changes of the distance, the market share, the daily passengers, and the merger dummy when the airlines provided the lowest fares on each city-pair routes and when the fares on the routes where each airline group had the largest market share respectively. For the lowest fare, low-cost carriers' fare increased by 3.3 percent after merger while their market share and the number of passengers were negatively related with the fare. That is, the market share and the passenger volume for low-cost carriers were negatively related with the airfare at the statistically significant level over the period of four years regardless of whether or not the merger took place, while the price of flight ticket increased by 3.3 percent after the Delta-Northwest merger. However, the same regression led to different results for legacy airlines; the merger did not affect the fare of the low-cost carriers. Similarly, the fare of Delta-Northwest did not change after its merger while the distance and the market share were positively related with its fare. An increase in the lowest fare was positively related with an increase in distance for all airline groups, except for LCCs. On the other hand, the airfare was negatively related with the number of passengers in general while only the change in passenger volume for LCCs was not statistically significant. The airfare was negatively related with LCC's passenger volume at the 90% significance, yet it showed more significance to that of Delta-Northwest. The change in market share did not seem to affect the airfare, except for that of LCCs which

showed a statistically significant negative relationship between its market share and the lowest fares. In summary for all the city-pair markets where airlines offered the lowest fares, only low-cost carriers' fare increased after merger, while the merger did not affect the fare of the other two groups.

Table 5

Variables	Coefficient	Standard Error	t-value	Prob> t	R ² _{adj}	Sample Size
Models:	(1) $\text{Fare}_{\text{low}} = \alpha_0 + \beta_1 \text{distance} + \gamma_1 \text{pax} + \delta_1 \text{mksh} + D_m + \varepsilon$					
	(2) $\text{Fare}_{\text{high}} = \alpha_1 + \beta_2 \text{distance} + \gamma_2 \text{pax} + \delta_2 \text{mksh} + D_m + \varepsilon$					
<i>Low-Cost Carriers:</i>						
Model (1)						
Constant	3.34886	0.06583	50.87	<.0001	0.4612	2938
Distance	0.29677	0.00671	44.22	<.0001	0.4612	2938
Passenger	-0.05696	0.00547	-10.41	<.0001	0.4612	2938
Market Share	-0.01321	0.00445	-2.97	0.0030	0.4612	2938
Merger Dummy	0.03347	0.00800	4.18	<.0001	0.4612	2938
Model (2)						
Constant	3.96690	0.12768	31.07	<.0001	0.4540	2288
Distance	0.27043	0.00846	31.95	<.0001	0.4540	2288
Passenger	-0.06782	0.00725	-9.35	<.0001	0.4540	2288
Market Share	-0.09248	0.01629	-5.68	0.0030	0.4540	2288
Merger Dummy	0.04885	0.00973	5.02	<.0001	0.4540	2288
<i>Legacy Airlines:</i>						
Model (1)						
Constant	3.83013	0.13823	27.71	<.0001	0.2820	602
Distance	0.19483	0.01349	14.45	<.0001	0.2820	602
Passenger	-0.02236	0.01369	-1.63	0.1030	0.2820	602
Market Share	0.05292	0.00926	5.72	<.0001	0.2820	602
Merger Dummy	-0.02919	0.01866	-1.56	0.1182	0.2820	602
Model (2)						
Constant	3.19570	0.14749	21.67	<.0001	0.3695	969
Distance	0.28869	0.01277	22.60	<.0001	0.3695	969
Passenger	-0.00229	0.00790	-0.29	<.0001	0.3695	969
Market Share	0.03659	0.01984	1.84	0.0654	0.3695	969
Merger Dummy	0.03822	0.01334	2.87	0.0043	0.3695	969
<i>Delta-Northwest:</i>						
Model (1)						
Constant	3.76352	0.12605	29.86	<.0001	0.3015	483
Distance	0.18534	0.01402	13.22	<.0001	0.3015	483
Passenger	-0.01556	0.01149	-1.35	0.1761	0.3015	483
Market Share	0.05147	0.00879	5.86	<.0001	0.3015	483
Merger Dummy	0.03106	0.01761	1.76	0.0784	0.3015	483
Model (2)						
Constant	3.43002	0.20279	16.91	<.0001	0.3028	772
Distance	0.21968	0.01507	14.58	<.0001	0.3028	772
Passenger	-0.07812	0.01181	-6.62	<.0001	0.3028	772
Market Share	0.22565	0.02822	8.00	0.0030	0.3028	772
Merger Dummy	-0.08454	0.01746	-4.84	<.0001	0.3028	772

For the markets where each airline group had the most market power, the fare of low-cost carriers increased by 4.8 percent after merger while the fare of legacy airlines increased by 3.8 percent. On the contrary, the fare of Delta-Northwest indicated a significant reduction of its fare by 8.4 percent after merger. In fact, the positive relationship between the market share and its fare demonstrated that Delta-Northwest's fare has been reduced after its merger while its market share has increased in the markets where the merging firm was already dominant relative to the other two groups. This is also confirmed by the fact that the market shares of two other airline groups did not have a strong relationship with their fares; low-cost carriers' market share was actually negatively related with fare while legacy airlines' market share did not have any significance to the fare.

The merger effect on the lowest fare was analyzed in Table 6. Legacy airlines fare showed a largest reduction in this market while their market share rose by the greatest among three airline groups. Likewise, the merger dummy for Delta-Northwest reflected its reduced airfare during the post-merger period and placed the merged firm right behind legacy airlines. The difference of the fare reduction between legacy airlines and Delta-Northwest was only 0.7 percent. While the change in the market share of legacy airlines accounted for a price increase by 9.5 percent and 7.6 percent for Delta-Northwest, the market share of LCCs only represented a price increase by 1.5 percent. Since the greater number of LCCs already offered the lowest fare in this market, the least price drop seemed reasonable. However, it was found that legacy and Delta-Northwest were quickly catching up with the low-cost carriers; both of their fares decreased almost two folds than LCCs whereas their market shares increased significantly during the post-merger period.

In comparison, the interaction terms for the market with the largest market share indicated that each airline's market share positively affected their airfare after merger, with Delta-Northwest's gain being the greatest. Delta-Northwest's market share after merger increased its fare by 29 percent while that of low-cost carriers' only increased their fare by 6.6 percent. On the other hand, the merger dummy suggested that the reduction in airfare was the greatest for Delta-Northwest. While both merger dummy variables for legacy and LCCs are negatively related with their fares, the intensity of fare reduction was almost three times greater for Delta. In summary, Delta-Northwest showed the greatest gain in its market share on the city-pair routes where they already had the largest market share while it had the biggest reduction in airfare after merger compared to its counterparts. LCCs, to the contrary showed the least reduction in their fare. Overall, the distance and the number of daily passenger were found to be consistent throughout all the regression results; the distance was positively related with fares while the passenger volume was negatively related. These were expected results, because airfares are generally higher for the long-haul flights while higher load factor decreases operating costs per passenger.

Table 6

Model: (1) $\text{Fare}_{\text{low}} = \alpha_0 + \beta_1 \text{distance} + \beta_2 \text{pax} + \beta_3 \text{mksh} + \delta_{\text{delm}} D_{\text{delm}} + \delta_{\text{ccm}} D_{\text{ccm}} + \delta_{\text{legm}} D_{\text{legm}} + \gamma_1 (\text{mksh} * D_{\text{delm}}) + \gamma_2 (\text{mksh} * D_{\text{ccm}}) + \gamma_3 (\text{mksh} * D_{\text{legm}}) + \varepsilon$

N=4002; R ² =0.4395				
Variables	Coefficient	Standard Error	t-value	Prob> t
Constant	3.5865	0.0571	62.8000	< 0.0001
Distance	0.2694	0.0057	46.9055	< 0.0001
Passenger	-0.0556	0.0048	-11.4990	< 0.0001
Market Share	-0.0202	0.0051	-3.9568	0.0001
After Merger LCC	-0.0580	0.0272	-2.1313	0.0331
After Merger Legacy	-0.1747	0.0420	-4.1560	< 0.0001
After Merger Delta	-0.1488	0.0468	-3.1786	0.0015
Mksh* After Merger LCC	0.0152	0.0079	1.9381	0.0527
Mksh* After Merger Legacy	0.0954	0.0138	6.9175	< 0.0001
Mksh* After Merger Delta	0.0765	0.0156	4.9206	< 0.0001

In comparison, the interaction terms for the market with the largest market share indicated that each airline's market share positively affected their airfare after merger, with Delta-Northwest's gain being the greatest. Delta-Northwest's market share after merger increased its fare by 29 percent while that of low-cost carriers' only increased their fare by 6.6 percent. On the other hand, the merger dummy suggested that the reduction in airfare was the greatest for Delta-Northwest. While both merger dummy variables for legacy and LCCs are negatively related with their fares, the intensity of fare reduction was almost three times greater for Delta. In summary, Delta-Northwest showed the greatest gain in its market share on the city-pair routes where they already had the largest market share while it had the biggest reduction in airfare after merger compared to its counterparts. LCCs, to the contrary showed the least reduction in their fare. Overall, the distance and the number of daily passenger were found to be consistent throughout all the regression results; the distance was positively related with fares while the passenger volume was negatively related. These were expected results, because airfares are generally higher for the long-haul flights while higher load factor decreases operating costs per passenger.

Model:

$$(2) \text{Fare}_{\text{high}} = \alpha_1 + \beta_4 \text{distance} + \beta_5 \text{pax} + \beta_6 \text{mksh} + \delta_{\text{delm}} D_{\text{delm}} + \delta_{\text{ccm}} D_{\text{ccm}} + \delta_{\text{legm}} D_{\text{legm}} + \gamma_4 (\text{mksh} * D_{\text{delm}}) + \gamma_5 (\text{mksh} * D_{\text{ccm}}) + \gamma_6 (\text{mksh} * D_{\text{legm}}) + \varepsilon$$

$$N=4002; R^2=0.4617$$

Variables	Coefficient	Standard Error	t-value	Prob> t
<i>Constant</i>	3.7773	0.1084	34.8569	< 0.0001
<i>Distance</i>	0.2944	0.0071	41.7444	< 0.0001
<i>Passenger</i>	-0.0444	0.0055	-8.0446	< 0.0001
<i>Market Share</i>	-0.0888	0.0171	-531971	< 0.0001
<i>After Merger LCC</i>	-0.3528	0.1090	-3.2352	0.0012
<i>After Merger Legacy</i>	-0.3652	0.1477	-2.4730	0.0134
<i>After Merger Delta</i>	-1.0589	0.1835	-5.7699	< 0.0001
<i>Mksh* After Merger LCC</i>	0.0656	0.0267	2.4584	0.0140
<i>Mksh* After Merger Legacy</i>	0.1373	0.0376	3.6484	0.0003
<i>Mksh* After Merger Delta</i>	0.2904	0.0454	6.3942	< 0.0001

CONCLUSION

This paper examined the Delta-Northwest merger in 2008 and its merger effect on different airline groups in the U.S. domestic market. Two different markets were introduced where the first market consisted of information on the airlines that offered the lowest fare on top 1,000 city-pair routes while the second market was comprised of the airlines with the largest market share on the same routes. These two different markets were compared within each group with the merger dummy showing the changes in the airfare in the post-merger period. In the markets where low-cost carriers offered the lowest price, the Delta-Northwest merger positively affected the LCC's fare. Yet, in the markets where each airline group held the largest market share, only Delta-Northwest's fare fell significantly while the other two airline groups' fares increased. In another regression results with the interaction term between the market share and the merger dummy, it was found that the airfares of legacy airlines and Delta-Northwest decreased at a faster rate than the low-cost carriers. One noticeable fact is that fares of legacy and Delta-Northwest decreased while their market shares increased simultaneously, irrespective of whether the airlines offered the lowest prices or had the largest market share on each route.

As the empirical results showed, the merged firm altered its pricing behavior through its merger while this affected the performance of non-merger airlines as well. As Delta-Northwest aggressively lowered its fares through consolidation, non-merging airlines may face an increased level of competition directly and indirectly. The US DOJ approved the merger between Delta and Northwest Airlines based on the reasoning that potential anticompetitive effects from the proposed merger would not outweigh the potential benefits for consumers. While there were speculations over potential anticompetitiveness effects with regard to the Delta-Northwest merger due to both airlines' pre-existing market power in the U.S. domestic market, the merged firm was actually found to facilitate price competition. Indeed, the regression results in this paper confirmed the reduction in airfares on 1,000 most popular domestic routes in the U.S. after the Delta-Northwest merger in 2008. Legacy airlines and Delta-Northwest reduced their

fares while increasing their market share at a faster rate than low-cost carriers, which suggested more heightened competition across those airline groups at the benefits for consumers, although more studies need to follow in subsequent years for the long-term effects.

ENDNOTES

1. *Performance Measures in the Airline Industry* available at http://www.bts.gov/programs/airline_information/performance_measures_in_the_airline_industry/
2. US DOT Research and Innovative Technology Administration (RITA) Website: http://www.bts.gov/programs/airline_information/performance_measures_in_the_airline_industry/
3. Among the mergers by legacy airlines in the 2000s, United Airlines made an announcement to acquire US Airways, but it soon withdrew the offer just before the DOJ barred the potential merger based on antitrust issues. The rest of the mergers were approved by the DOJ; American Airlines acquired Trans World Airlines (TWA) in 2001, US Airways bought out by America West Airlines, Delta Air lines merged with Northwest Airlines in 2008, United Airlines merged with Continental Airlines in 2010.
4. Statement of the Department of Justice's Antitrust Division released on October 29, 2008, available at http://www.justice.gov/atr/public/press_releases/2008/238849.htm
5. The four-firm ratio was measured by the top four airlines' total number of domestic passengers proportional to the total number of domestic passenger in each given year.
6. This is a slightly different approach than using the DB1B data or T100 data that are widely used in studies on the airline industry. DB1B is a 10 percent random sampling of tickets on the directional airport-pair markets, yet smaller LCCs with less frequency is often omitted from the sample.
7. Since the regional airlines normally operate either by code sharing or alliance with larger airlines, the performance of regional airlines have been disregarded for simplicity.
8. The group of Low-Cost Carriers includes Allegiant Air, Alaska Airlines, AirTran Airways, America West Airlines, Frontier Airlines, JetBlue Airways, Midwest Express Airlines, Spirit Air Lines, USA 3000 Airlines, US Airways, and Southwestern Airlines. The group of legacy airline consists of American Airlines, Continental, and United Airlines. Although Continental and United announced a merger agreement in 2010, their final agreement was not announced until 2011 and the integration process will be finalized in 2012. Since their integration has not started until 2011, they are considered as legacy airlines without merger and are sorted out as a control group.
9. Data is available at http://ostpxweb.dot.gov/aviation/x-50%20Role_files/consumerairfarereport.htm. See table 1A for 3rd quarter of each year from 2006 to 2010.

REFERENCES

- Berry, S. and J. Panle. 2010. "Tracing the Woes: An Empirical Analysis of the Airline Industry." *American Economic Journal*. Vol. 2(3): 1-43.
- Borenstein, S. 2005. "U.S. Domestic Airline Pricing, 1995-2004." Working Paper No. CPC05-48, EconWPA.
- Borenstein, S. and N.L. Rose. 2007. "How Airline Markets Work...Or Do They? Regulatory Reform in the Airline Industry." Working Paper No 13452, National Bureau of Economic Research.
- Bruce, J. K. and Lee, D., and E. Singer. 2010. "City-Pairs vs. Airport-Pairs: A Market-Definition Methodology for the Airline Industry." Retrieved from University of California Irvine Website: <http://www.socsci.uci.edu/~jkbrueck/groupings%20paper.pdf>.
- Carlton, D. W., Landes, W. M. and R.A. Posner. 1980. "Benefits and Costs of Airline Mergers: A Case Study." *Bell Journal of Economics*. Vol. 11(1): 65-83.
- Douglas, G.W. and J.C. Miller III. 1974b. "Quality Competition, Industry Equilibrium, and Efficiency in the Price-Constrained Airline Market." *American Economic Review*. Vol. 64(4): 657-669.
- Fageda X. and J. Perdiguer. 2011. "An empirical analysis of a merger between a network and low-cost airlines." Working Paper No. XREAP2011-01, Universitat de Barcelona. Retrieved from <http://www.pcb.ub.edu/xreap/aplicacio/fitxers/XREAP2011-01.pdf>
- Flouris, T. and T.J. Walker. 2005. "The Financial Performance of Low-Cost and Full-Service Airlines in Times of Crisis." *Canadian Journal of Administrative Sciences* Vol. 22: 3–20.
- Goetz, A.R. 2002. "Deregulation, competition, and antitrust implications in the US airline industry." *Journal of Transport Geography*. Vol. 10: 1-19
- Hofer, C., Windle, R. J. and M.E. Dresner. 2008. "Price premiums and low cost carrier competition." *Logistics and Transportation Review*. Vol. 44(5): 864-882.
- Ito, H. and D. Lee. 2003. "Low Cost Carrier Growth in the U.S. Airline Industry: Past, Present, and Future." Working Paper No. 2003-12, Brown University Economics Department. Retrieved from http://www.brown.edu/Departments/Economics/Papers/2003/2003-12_paper.pdf
- Kim, E. H. and V. Singal. 1993. "Mergers and Market Power: Evidence from the Airline Industry." *The American Economic Review*. Vol. 83(3), 549-569.
- Klein, J. et al. 1999. "Predation 'strategy' at American airlines: A Change of Heart at the Justice Department?" *Antitrust Law and Economics Review*. 93-110.
- Kwoka J. and E. Shumilkina. 2010. "The Price Effect of Eliminating Potential Competition: Evidence From An Airline Merger." *Journal of Industrial Economics*. Vol. 58(4): 767-793.
- Lichtenberg, F. R. and M. Kim. 1989. "The Effects of Mergers on Prices, Costs, and Capacity Utilization in the U.S. Air Transportation Industry" Working Paper No. 3197, National Bureau of Economic Research.
- Morrison, S. A. 2000. "The Proposed United-US Airways Merger: Hearing before the Committee on Transportation and Infrastructure." *United States House of Representatives*. Retrieved from

- Northeastern University Website: <http://www.economics.neu.edu/morrison/research/house000615.pdf>.
- Morrison, S. A. 1996. "Airline Mergers: A Longer View." *Journal of Transport Economics and Policy*. Vol. 30(3): 237-250.
- Pels, E. 2008. "Airline network competition: Full-service airlines, low-cost airlines and long-haul markets," *Research in Transportation Economics*. Vol. 24(1): 68-74.
- Peters, C. 2006. "Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry," *Journal of Law & Economics*. Vol. 49(2): 627-49.
- Richard, O. 2003. "Flight frequency and mergers in airline markets." *International Journal of Industrial Organization*. Vol. 21 (6): 907-922.
- Tan, K. M. 2010. "Incumbent Response to Entry by Low-Cost Carriers in the U.S. Airline Industry." Doctoral dissertation, The Ohio State University.
- The American Antitrust institute. 2008. "The Merger of Delta Air Lines and Northwest Airlines: An Antitrust White Paper." Retrieved from http://www.antitrustinstitute.org/files/AIWhite%20Paper_Delta_NW_071020081922.pdf
- Transportation Research Board. 1999. "Entry and Competition in the US Airline Industry: Issues and Opportunities." Special report 255, National Research Council.
- United States Department of Transportation. 2007. "Performance Measures in the Airline Industry." Retrieved from http://www.bts.gov/programs/airline_information/performance_measures_in_the_airline_industry/
- United States General Accounting Office. 2010. "Issues Raised by the Proposed Merger of United and Continental Airlines." Publication No. GAO-10-778T. May.
- United States General Accounting Office. 2008. "Airline Industry: Potential Mergers Acquisition Driven by Financial and Competitive Pressures." Publication No. GAO-08-845. July.
- United States General Accounting Office. 2004. "Commercial Aviation: Despite Industry Turmoil, Low-Cost Airlines Are Growing and Profitable." Publication No. GAO-04-837T. June.
- United States General Accounting Office. 1996a. "Airline Deregulation: Changes in Airfares, Service, and Safety at Small, Medium-Sized, and Large Communities." Publication No. GAO/RCED-96-79. Washington, DC. April.
- United States General Accounting Office. 1993. "Airline Competition: Higher Fares and Less Competition Continue at Concentrated Airports". Publication No. GAO/RCED-93-171. Washington, DC. July.
- United States Department of Transportation. 1990. "Secretary's Task Force on Competition in the US Domestic Airline industry." Washington DC.
- Vowles, T. M. 2006. "Airfare pricing determinants in hub-to-hub markets." *Journal of Transport Geography*. Vol. 14(1): 15-22
- Weinberg, M. 2007. "The Price Effects of Horizontal Mergers: A Survey." Working Paper No. 62, Princeton University.

The Rate of Returns to Schooling: A Case Study of Urban China

Xu Zhang*

Abstract

This study seeks to study the rate of return to schooling in urban China by addressing two questions: First, have the returns to schooling in China been underestimated? Second, should Mincer earning function be directly applied? The study suggests a modified Mincer earning equation with quadratic term of years of schooling to reflect the differences in average rate of returns to additional year of schooling among various levels of education. The OLS estimates of returns to education are lower than those from the IV approach. The overall returns to education from OLS estimation based on the standard Mincer equation are in the range of 7%-8%. The IV approach reports a much higher rate of returns to education, 16% for college graduate, based upon the modified Mincer earning equation, while the figure from the OLS estimation is 10.24%.

Introduction

China has been experiencing profound economic reforms and a sustained average annual growth rate of over 6% for a few decades. High-speed economic growth result in more human capital investment in China and the rate of returns to schooling has been widely studied.

Researchers find the studies are interesting and important due to several reasons. First, returns to schooling indicate the productivity of education and show the incentive for private investment in one's human capital, which contributes to economic growth especially for developing countries. Second, different levels of returns to schooling among various groups of people (e.g. men and women, different ethnic groups, city and rural areas, etc) may shed some light on the assessment of efficiency of resource allocation pertaining to education. Third, estimation on economic returns to schooling has important policy implications. Psacharopoulos and Patrinos (2002) suggest government should design policies and incentives to promote private investment in human capital and ensure investment for low-income families based upon the evidence from research. Finally, examining returns to schooling in China can help to assess the structure of the Chinese labor market and the extent of economic transformation. Historically, Chinese culture highly values education. The study of this issue can signal private demand for education. If the returns to schooling are too low, it may hurt economic growth by not accumulating enough human capital.

* Farmingdale State College, State University of New York - History, Economics and Politics Department
2350 Broadhollow Road, Farmingdale, NY 11735 - xu.zhang@farmingdale.edu

Psacharopoulos (1994) surveys empirical studies of the returns to education across 98 countries and finds a world average of 10.1% , 9.6% in Asia and a range between 11.2% to 11.7% in low and middle income countries (those with a per capita income of less than \$2449). However, existing literature shows a much lower economic returns to education in China. Johnson and Chow (1997), using the data from 1988 Chinese Household Income Project (CHIP-88), estimates a return of 3.29%. Liu (1998) suggests the returns to education are in the range of 2.8% to 3.6% by using the same data set—CHIP 88. The CHIP 1995 (CHIP-95) data have an advantage over 1988 data as the former contain information on working hours which can be used to calculate hourly wage rate, while the latter collect data on annual earnings. Using CHIP-95, Li (2003) reports the overall return for an additional year of schooling is 5.4%. Chen and Hamori (2009) study the China Health Nutrition Survey (CHNS) of 2004 and 2006 and present the overall economic returns to schooling are in the range of 7% - 8%.

The conventional approach to estimate the returns to schooling is the standard Mincer earning equation (Mincer, 1974). Although hundreds of papers have examined this issue in many countries, using various data sets, and with different estimation techniques, few studies produce the “true” rate of return to education (Heckman, Lochner, & Petra, 2005).

The main problem discussed in the previous studies on the economic returns to education is the endogeneity of the schooling variable. Chosen years of schooling are not exogenous and tend to be correlated with unobservables in the error term of the earning function. The unobservable can be ability, which is positively correlated with years of schooling and with earnings, giving rise to “ability bias” (Card, 1999). Thus, commonly used estimation method – ordinary least squares (OLS) tends to overestimate the returns to education. Some studies employ various instrumental variables such as parental education (Heckman and Li, 2004; Li and Luo, 2004), parental income (Fleisher, Li, Li and Wang, 2005), and spouse’s education (Trostel et al., 2002). Card (1999) and Card and Lemieux (2001) state that IV estimates of the rate of return to education can be higher or lower than OLS estimates depending on the instrument variables chosen in the analysis. Heckman and Li (2004) use the data from the China Urban Household Income and Expenditure Survey in year 2000 to study the returns to higher education in China. They find the IV estimates of the economic returns to four-year college education are higher than the OLS estimates of the Mincer model. Fleisher, Li, Li and Wang (2005) use data from CHIP 1988, 1995, and 2002 to examine how selection and sorting affected the change of the economic returns to schooling for college graduates during China’s reforms between 1988 and 2002. They find the IV estimates of the returns to college education are sensitive to the use of a proxy for ability.

Panel data estimation can be another approach to control for unobserved individual fixed effects. Unfortunately, panel data are often not available, particularly for developing countries. Therefore, individual cross-sectional data are most commonly used in the study of returns to education in developing countries.

Most existing literatures directly follow the Mincer earning function to study the economic returns to education. The variable “years of schooling”, which is a proxy for investment in human capital, is the only

measure of education entering the regression analysis. Thus, by employing a Mincer earning function, previous studies simply assume the marginal returns to an additional year of schooling is constant and ignore the possibility that returns to schooling can vary at different level of education. In addition, previous studies estimate work experience, a proxy for general and specific job training in human capital theory (Mincer, 1974), on the basis of age and years of schooling but ignore the possibility that some group of people such as Intellectual Youth may spend time on totally unrelated jobs, for example, moving rocks, or growing plants. Intellectual Youth were sent to countryside for “re-education” during the Cultural Revolution when they were students in middle school and high school. Thus, estimates of job experience may be overestimated and therefore underestimate the returns to job experience, which influence the estimation of the returns to schooling.

Therefore, this study seeks to answer two questions: first, should the Mincer earning function be directly applied to estimate the returns to schooling in China? Second, whether the returns to schooling in China have been underestimated. The major findings are as follows. The study suggests a modified Mincer earning equation with quadratic term of years of schooling to reflect the differences in average rate of returns to additional year of schooling among various levels of education. The OLS estimates of returns to education are lower than those from the IV approach. The overall returns to education from OLS estimation based on the standard Mincer equation are in the range of 7%-8%. The IV approach reports a much higher rate of returns to education, 16% for college graduate, based upon the modified Mincer earning equation, while the figure from the OLS estimation is 10.24%. The household registration system plays a very important role in the determination of wage. Rural households tend to earn 30%-40% less than urban households. Workers working in state-owned enterprises tend to earn more than private sector.

The rest of the paper is organized as follows. The following section discusses the empirical model and data. Section 3 presents estimation results. Section 4 concludes.

The Empirical Framework and data

This study begins with the basic human capital earning function (Mincer, 1974):

$$\ln w_i = \alpha_0 + \beta_0 S_i + \beta_1 X_i + \beta_2 X_i^2 + \epsilon_i \quad (1)$$

where $\ln w_i$ denotes the natural logarithm of the hourly wage for individual i ; S_i denotes years of schooling of individual i ; X_i represents number of years of job experience (or age) for individual i ; ϵ_i is the error term, and β_0 shows the average rate of return to one additional year of schooling.

The ordinary least square (OLS) estimation procedure has been extensively used in previous studies. However, the OLS estimates may bias the rate of return to education due to the omission of the ability variables. Standard critique emphasizes an upward bias because of the positive correlation among the ability, earning and years of schooling. Controversies still exist regarding whether the omitted ability

variables bias the OLS estimates, and if they do, how much the potential bias would be. This paper will follow this standard approach and present the result from the OLS estimation first. In order to address the endogeneity problem, an instrument variable approach is also adopted in this study by using spouse's education as an instrument (Trostel et al., 2002). Spouse's education is correlated with one's schooling but not with one's wage rate and education is one of the positive assortative traits in marriage (Jepsen and Jepsen, 2002). The following two-equation framework is adopted to address the endogeneity problem of schooling:

$$\ln w_i = \alpha_0 + \beta_0 S_i + \beta_1 X_i + \beta_2 X_i^2 + \epsilon_i \quad (2)$$

$$S_i = \gamma Z_i' + u_i \quad (3)$$

where Z_i represents the vector of the instrument variable and the other exogenous variables.

This study uses urban sample of the third wave (in year 2002, the first and the second wave of data were collected in 1988 and 1995, respectively) of data from China Household Income Project (CHIP), a national cross-sectional study collectively designed by a team of Chinese and western economists and conducted by the Institute of Economics at the Chinese Academy of Social Science. The CHIP-2002 survey covers about 6835 households and 20632 individuals in urban areas of China.

CHIP-2002 contains very rich information about the earnings, which include regular salary/wage, income from secondary jobs, all kinds of transfer income and all these earning measures are in annual terms. Since this study focus on wages determination, I only look at regular salary/wage for annual earnings and exclude all the other sources of income. In separate survey questions, the respondents are interviewed to indicate the number of work-hours in an average day and the average number of work days per week. Thus, the hourly wage rate can be calculated by using annual earnings and the number of work-hours. Before CHIP-95, due to data limitation, annual earning, instead of hourly wage rate, was used as dependent variable in the analysis. Li (2003) states using annual earnings will underestimate earning differences between different educational groups as highly educated people tend to work fewer hours than less educated people.

Another advantage shared by both CHIP-2002 and CHIP-95 is respondents were interviewed to indicate at which year they found their first jobs, which makes it possible to calculate the actual job experience of the respondents, while the earlier studies estimate job experience through age and years of schooling. According to the human capital theory, job experience is a proxy for human capital investment on the job. At the beginning of one's career, human capital accumulation dominates depreciation. As years of experience increases, human capital depreciation will eventually dominate accumulation, thus, job experience enters the earning function in linear form as well as quadratic form. However, there is still a problem in estimating the rate of return to schooling when estimated actual job experience is employed in the regression as it may not reflect the possibility that one may be idle for a while to land a job. Thus, job experience may be overestimated while the returns to schooling may be underestimated if one has been unemployed for some time between the year he got his first job and year 2002.

There is an improvement in the measure of education as CHIP-2002 collects data on eight education levels based upon degrees as well as years of schooling through a separate interview question. If education levels are employed to estimate the returns to schooling, a particular number of years should be assigned to each education level. The existing literature takes the following transformation scheme for each education level: Graduate (Yanjiusheng, 19 years), college/university (Daxue, 16 years), Junior college (Dazhuan, 15 years), Technical secondary school (Zhongzhuan, 12 years), Senior middle school (Gaozhong, 12 years), Junior middle school (Chuzhong, 9 years), elementary school (Xiaoxue, 6 years) and Classes for eliminating illiteracy (2 years). Since some students may stay at the same grade for more than one year, using education level as measure of education may underestimate the years of schooling, and, therefore, overestimate the returns to schooling. In this study, years of schooling directly from the interview question will be used to measure the education background of the respondents.

A problem may arise in the standard Mincer earning equation as it assumes the average returns to schooling is constant and the measure of education is assumed to be linear with the earning variable, which neglects the possibility that contributions of various education levels to earning can be different. Thus, in this study I adopt a modified version of Mincer earning equation as follows:

$$\ln w_i = \alpha_0 + \beta_0 S_i + \beta_1 S_i^2 + \beta_2 X_i + \beta_3 X_i^2 + \epsilon_i \quad (4)$$

where years of schooling S_i and job experience X_i appear in both linear and quadratic terms.

Since this study focus on the contribution of schooling to wage, I restrict the sample to fulltime workers who were legally employed for the whole year at the time of interview. I also exclude those people younger than 16 years old as students are mandated to be in school before 16 years old by compulsory education law. Observations with inconsistent information on age and actual job experience and self-employed workers who report zero salary/wage are excluded as well. After exclusion, the sample contains about 8925 observations.

A few dummy variables are generated to capture the impact on earnings from the perspective of sex, ownership, household registration system and ways to find the jobs. D_female is set as 1 for females and 0 for males. D_SOE is set as 1 if the employers are state/provincial/local government enterprises. D_hukou is set as 1 for rural household, and as 0 for urban household registration. D_guanxi is set as 1 if the job is introduced by relatives or friends, and as 0 if otherwise. Intellectual Youth are a group of students who were sent to countryside for “re-education” by conducting heavy-labor work during the Cultural Revolution when the value of education was highly ignored. Thus, it is interesting to examine the wage determination for this special group of people and dummy variable D_IYouth is generated as 1 for being Intellectual Youth during Cultural Revolution, 0 for not. Table 1 reports the definition for each variable.

Results and Discussion

Table 2 presents the descriptive statistics of all the variables.

As discussed in previous section, the Mincer earning equation is estimated for hourly wage rate through OLS analysis. Two proxies are employed for the work experience: the age and the actual work experience. Household registration system (Hukou) is strictly implemented in China and becomes a constraint for a few issues such as job-hunting, school enrollment, healthcare benefits. A dummy variable for household registration system D_{hukou} is added to examine the impact of household registration system on wage rates. In addition, market condition reflected by the ownership of the enterprises can also influence the wage rate. Thus, a dummy variable for state owned enterprises D_{SOE} is incorporated in the regression. The wage difference between sex is another focus of study and a dummy variable D_{female} is generated to capture this effect. The results based on hourly wage by using OLS estimation procedure are given in Table 3.

The results from Table 3 show that the OLS estimates of return to education in 2002 are in the range of 7%-8%, which are below the world average. Households registered as rural earned about 40% less than urban households. Females earn about 13%-16% less than males while people work in state-owned enterprises earn around 32%-33% more than other enterprises. The impact of having experience as Intellectual Youth is insignificant in this regression and thus ignored. Only 879 observations reported their jobs were introduced by relatives/friends (through Guanxi) and the impact is insignificant and thus ignored as well.

Households who were not originally born in cities can obtain Hukou through the following ways: going to college, becoming officials, joining the army, land occupied by government, buying a house in the city and buying a Hukou. Therefore, the wage difference between rural and urban household may be overestimated as those people who obtain Hukou by going to college or becoming officials tend to earn more due to higher education background. In order to capture the pure effect from Hukou on the wage rate, OLS regression is conducted again in the sample excluding people who obtained their Hukou by going to college or becoming officials and the results are presented in Table 4. Clearly, the results from Table 4 show that household registered as rural earn less than urban household while the magnitude decreases by about 5%-6% when controlling for the way to obtain Hukou.

As explained in previous section, the Mincer earning equation ignores the possibility that the marginal returns to schooling can be different among various levels of education. To address this problem, a modified Mincer earning equation is proposed as equation (4) where a measure of education appears in both linear and quadratic forms. Table 5 reports the results of the OLS estimation on equation (4).

As Table 5 presents, the linear term of years of schooling becomes insignificant while the quadratic term of schooling is significant in the estimation on the modified Mincer earning equation. By using the model with actual experience, the average return to additional year of schooling is 0.0064 times number of years of schooling. Thus, it is predicted that the higher education one receives, the higher returns to

schooling. For example, for a college graduate, the average return to schooling is $0.0064 \times 16 = 10.24\%$; while for a junior middle school graduate, the average returns to schooling is $0.0064 \times 9 = 5.76\%$.

To address the endogeneity problem of schooling, this study follows the IV approach and matches the individual data with the household data. The data on head of household and spouse are pulled out. Spouse's education is used as a proxy for the years of schooling of head of household as education is a positive assortative trait. Spouse's education is related with years of schooling of head of household, but not connected with the wage rate. Table 6 presents the results from 2SLS estimation. Again, since quadratic term of schooling appears in the second stage of regression, the average return to schooling is about $0.01x$ years of schooling. For example, the average rate of return to schooling for a college graduate is $0.01 \times 16 = 16\%$, which is much higher than the OLS estimate 10.24%. Thus, the estimates of returns to schooling from the IV approach are higher than those from the OLS approach and the OLS estimates tend to underestimate the rate of returns to schooling.

CONCLUSION

This study uses the most recent data from the Chinese Household Income Project (2002) to examine the returns to education in urban China. The main goals are to investigate two questions: first, is standard Mincer earning equation outdated to estimate the returns to education in China? Second, has the rate of returns to education been underestimated from the OLS technique compared to the 2SLS approach?

The study suggests a modified Mincer earning equation with quadratic term of years of schooling to reflect the differences in average rate of returns to additional year of schooling among various levels of education. The OLS estimates of returns to education are lower than those from the IV approach. The overall returns to education from OLS estimation based on the standard Mincer equation are in the range of 7%-8%. The IV approach reports a much higher rate of returns to education, 16% for college graduate, based upon the modified Mincer earning equation, while the figure from the OLS estimation is 10.24%. Household registration system plays a very important role in the determination of wage. Rural households tend to earn 30%-40% less than urban households. Workers working in state-owned enterprises tend to earn more than private sector.

The study did not consider measurement errors, which can be an extension for future studies.

REFERENCES

- Card, D. (1999). "The causal effect of education on earnings." In O. Ashenfelter, & D. Card (Eds.) *Handbook of Labor Economics, Vol. 3b.* (pp. 1801- 1863)
- Card, D. & T. Lemieux. (2001). "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis." *The Quarterly Journal of Economics*, 116(2), 705 – 746.
- Chen, Guifu and Shigeyuki Hamori. (2009). "Economic Returns to Schooling in Urban China: OLS and the Instrumental Variables Approach." *China Economic Review* 20 (2009), 143 – 152.

- Fleisher, B.M.; Li, H.; Li, S. and X. Wang (2005). "Sorting, Selection, and Transformation of Return to College Education in China." Working Paper Nos. 05-07, Ohio State University.
- Heckman, J.J. and X. Li. (2004). "Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000." *Pacific Economic Review*, 9(3), 155 – 171.
- Heckman, J.; Lochner, L. and T. Petra. (2005). "Earnings Functions, Rate of Return, and Treatment Effects: The Mincer Equation and Beyond." NBER Working Paper: 11544.
- Jepsen, Christopher and Lisa Jepsen. (2002). "An Empirical Analysis of the Matching Patterns of Same-Sex and Opposite-Sex Couples." *Demography*, 39(3), pp.435-453
- Johnson, E.N. and G.C. Chow. (1997). "Rates of Return to Schooling in China." *Pacific Economic Review*, 2(2), 101 – 113.
- Li, H. (2003). "Economic Transition and Returns to Education in China." *Economics of Education Review*, 22(3), 317 – 328.
- Li, H. and Y. Luo. (2004). "Reporting Errors, Ability Heterogeneity, and Returns to Schooling in China." *Pacific Economic Review*, 9(3), 191–207.
- Liu, Z. (1998). "Earnings, Education and Economic Reforms in Urban China." *Economic Development and Cultural Change*, 46(4), 697–725.
- Mincer, J. (1974). "Schooling, Experience and Earnings." New York NBER.
- Psacharopoulos, G. (1994). "Returns to Investment in Education: a Global Update." *World Development*, 22(9), 1325–1343.
- Psacharopoulos, G., and A.P. Harry. (2002). "Returns to Investment in Education: a Further Update." World Bank Policy Research Working Paper 2881.
- Trostel, P.; Walker, I. and P. Woolley. (2002). "Estimates of the Economic Return to Schooling for 28 Countries." *Labor Economics*, 9(1), 1- 16.

Table 1: Definition of Variables

Variable	Definition
Wage	Natural Log of Hourly Wage Rate
Age	Age in Years
Act_EXP	Actual Work Experience by Year 2002
Schooling	Years of Schooling
D_hukou	1 for Household Registered as Rural; 0 for Household Registered as Urban
D_female	1 for Females and 0 for Males
D_SOE	1 for Ownership of Present Work Unit as State/Provincial/Local State Enterprises; 0 for Private/Share-Holding Companies
D_guanxi	1 if Job is introduced by Relatives/Friends; 0 if Otherwise.
D_IYouth	1 for Intellectual Youth, 0 for not.

Table 2: Descriptive statistics of all variables

Variable	Number of Observations	Mean	Standard Deviation
Wage	8925	1.472	1.8096
Age	8996	40.7023	9.0738
Act_EXP	8968	20.5467	9.5908
Schooling	8996	11.5728	2.9653
D_hukou	8996	0.011	0.1043
D_female	8996	0.4369	0.496
D_SOE	7239	0.5959	0.4907
D_guanxi	9445	0.0931	0.2905
D_IYouth	11398	0.2092	0.4067

Source of Data: China Household Income Project 2002 (CHIP-02), obtained from Inter-University Consortium for Political and Social Research

Table 3: Rates of Return to Education (Fulltime Workers, OLS)

Dependent Variable	Wage	Wage
Schooling	0.0774** (23.58)	0.0808** (24.68)
Age	0.0948** (12.98)	
Age_Square	-0.0010** (-11.34)	
Act_EXP		0.0373** (10.34)
Act_EXP_Square		-0.0005** (-6.02)
D_hukou	-0.4274** (-4.99)	-0.3823** (-4.44)
D_female	-0.1569** (-8.36)	-0.1341** (-7.17)
D_SOE	0.3311** (17.29)	0.3196** (16.74)
Constant	-1.6796** (-11.03)	-0.1535** (-2.77)
Summary Statistics	$\bar{R}^2 = 0.1806$ F=243.23 N=6597	$\bar{R}^2 = 0.1851$ F=250.16 N=6584

Numbers in parentheses are t-statistics

**Denotes that variables are significant at the 5% level

Table 4: Rate of Returns to Education (Full-Time Workers, Excluding Individuals who Obtained Hukou by Going to College or Becoming Officials, OLS)

Dependent Variable	Wage	Wage
Schooling	0.0820** (23.84)	0.0842** (24.63)
Age	0.0944** (12.49)	
Age_Square	-0.0010** (-10.89)	
Act_EXP		0.0373** (10.06)
Act_EXP_square		-0.0005** (-5.84)
D_hukou	-0.3712** (-4.30)	-0.3227** (-3.72)
D_female	-0.1552** (-8.02)	-0.1317** (-6.82)
D_SOE	0.3420** (17.43)	0.3307** (16.90)
Constant	-1.7299** (-11.04)	-0.1975** (-3.47)
Summary Statistics	$\bar{R}^2 = 0.1858$ F=236.49 N=6193	$\bar{R}^2 = 0.1897$ F=242.03 N=6180

Numbers in parentheses are t-statistics

**Denotes that variables are significant at the 5% level

Table 5: Rate of Returns to Education (Full-Time Workers, OLS, Based on Equation (4) Modified Mincer Earning Equation)

Dependent Variable	Wage	Wage
Schooling	0.0189 (1.30)	0.0121 (0.84)
Schooling_Square	0.0027** (4.15)	0.0032** (4.90)
Age	0.0962** (13.12)	
Age_Aquare	-0.0010** (-11.52)	
Act_EXP		0.0387** (10.71)
Act_EXP_Square		-0.0006** (-6.31)
D_hukou	-0.4325** (-5.05)	-0.3860** (-4.50)
D_Female	-0.1530** (-8.15)	-0.1290** (-6.89)
D_SOE	0.3394** (17.22)	0.3170** (16.62)
Constant	-1.412** (-8.55)	-0.1790** (-2.05)
Summary Statistics	$\bar{R}^2 = 0.1826$ F=211.45 N=6597	$\bar{R}^2 = 0.1879$ F=218.61 N=6584

Numbers in parentheses are t-statistics

**Denotes that variables are significant at the 5% level

Table 6: Rate of Returns to Education (Head of Household and Spouse, 2SLS)

	1 st (Schooling)	Stage (Wage)	2 nd (Wage)	Stage	1 st (Schooling)	Stage (Wage)	2 nd (Wage)	Stage
Schooling_Hat_Square			0.0047** (10.17)				0.005** (10.80)	
IV: Spouse's Education	0.4650** (29.37)				0.4600** (29.27)			
Age	-0.3743** (-6.62)		0.1492** (8.35)					
Age_Square	0.0039** (6.06)		-0.0016** (-8.24)					
Act_EXP					-0.1715** (-6.37)		0.0469** (5.50)	
Act_EXP_Square					0.0027** (4.78)		-0.0007 (-4.22)	
D_hukou	-1.6187** (-3.22)		-0.3515** (-2.32)		-1.9175** (-3.81)		-0.2636* (-1.73)	
D_Female	-0.2631** (-2.52)		-0.0336 (-1.09)		-0.2743** (-2.63)		-0.0144 (-0.47)	
D_SOE	0.7130** (7.33.)		0.3139** (10.13)		0.7333** (7.55)		0.2947** (9.48)	
Constant	14.5236** (11.52)		-2.5472** (-6.13)		8.2242** (22.11)		0.0528 (0.40)	
Summary Statistics	$\bar{R}^2 = 0.3016$		$\bar{R}^2 = 0.1139$		$\bar{R}^2 = 0.3058$		$\bar{R}^2 = 0.1084$	
	F=214.84		F=64.17		F=218.64		F=60.58	
	N=2972		N=2949		N=2965		N=2942	

Numbers in parentheses are t-statistics

**Denotes that variables are significant at the 5% level

*Denotes that variables are significant at the 10% level