

# Comparative Advantage

---

*Stanford Undergraduate Economics Journal*

*Spring 2015*

*Volume 3, Issue 1*

## Note from the Editor

---

On behalf of the Editorial Board, I am honored to present the 2015-2016 Stanford Economics Journal.

The Journal has undergone considerable change this year, as it continues to evolve to better achieve its mission of showcasing the best undergraduate economic research and analysis. In addition to transitioning to an annual rather than a biannual publication, we relaxed our submission criteria to allow for papers that lie at the intersection of economics and other social sciences. We also allowed younger students to submit shorter articles on economic topics that interested them. These changes were intended to expand the Journal's relevance and audience and to make the Journal more diverse and inclusive.

We truly appreciate all the submissions we received from around the globe. We enjoyed reading the incredible research that students have undertaken, and we could not be more thrilled with the 11 papers that we selected for the Journal. The topics examined in these 10 papers covered a wide spectrum, from microfinance and economic crises to fashion design and online dating. One feature they all shared, however, was the academic rigor that formed a foundation for thoughtful analysis. We could not be more proud to share these papers with the academic and online community as a whole.

Thank you to the Stanford Economics Association and the Stanford Department of Economics for their continued guidance and support. We are grateful for their commitment to the Journal and its mission of inspiring student interest in the field of economics.

Kimberly Tan

*2015-2016 Editor-in-Chief*

## Editors and Staff

---

### Editor-in-Chief

*Kimberly Tan*

### Production and Design Editor

*Michael Yu*

### Associate Editors

*Kay Dannenmaier*

*Andrew Granato*

*Rohit Kolar*

*Thomas Leung*

*Xianming Li*

*Surya Narayanan*

*David Schmitt*

*Daniel Wright*

*Laura Zhang*

## Contents

<b>1</b>	<b>Living in the Shadows or Government Dependents: Immigrants and Welfare in the United States / <i>Charles Weber</i></b>	<b>4</b>
<b>2</b>	<b>Crude Oil Price: An Indicator of Consumer Spending / <i>Natalie Li, Eva Lin</i></b>	<b>21</b>
<b>3</b>	<b>Romantic Unemployment: Neoclassical Economic Theory in Online Dating and Labor Markets / <i>Felix Carreon III</i></b>	<b>38</b>
<b>4</b>	<b>Contraception Employed: Using Economic Models to Predict the Effect of Employment on Condom Usage in Brazil / <i>Seth Merkin Morokoff</i></b>	<b>56</b>
<b>5</b>	<b>Creativity of Fashion Design: An Economic Creative Lifecycle Analysis of Three 20th Century Fashion Designers / <i>Benjamin Nickerson</i></b>	<b>84</b>
<b>6</b>	<b>The Dangers of Letting Money Float / <i>Jaewoo Jang</i></b>	<b>113</b>
<b>7</b>	<b>Afterlife Beliefs, Work Ethic, and Economic Outcomes / <i>Nicholas Cornell</i></b>	<b>123</b>
<b>8</b>	<b>Evaluating the Impact of Microfinance on Health Decisions / <i>Robert Fluegge</i></b>	<b>149</b>
<b>9</b>	<b>Probit Analysis of Non-Ferrous Metals as Leading Indicators: Testing Industrial Metal Prices in a Binary-Response Model of Pre-Crisis Thailand, 1997 / <i>Admund Tay</i></b>	<b>164</b>
<b>10</b>	<b>The Reaction Function of The Federal Reserve Post 2008 Financial Crisis / <i>Michel Cassard</i></b>	<b>179</b>

# Living in the Shadows or Government Dependents: Immigrants and Welfare in the United States

Charles Weber

Harvard University

May 2015

## **Abstract**

Are immigrants in the United States more likely to be enrolled in welfare programs than natives and how has this comparative usage changed over time? To address this question, I pool four panels from the 1990, 1991, 2001, and 2004 Survey of Income and Program Participation and regress different measures of welfare usage on binary migrant variables as well as including time fixed-effects. I find three trends: first, there is no statistically significant difference between the welfare use of similar immigrants and natives, second, immigrant welfare use decreased after the enactment of the 1996 Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA), which is probably driven by great decreases in Medicaid use over that time period, and third, if Medicaid use is excluded from the measurement of welfare use, immigrants use more of the remaining programs than natives and at an increasing rate after PRWORA. I conclude by proposing several directions for future research.

## **1 Introduction**

Are immigrants in the United States more likely to be enrolled in welfare programs than natives? There is no shortage of media coverage proclaiming the costs that migrants impose on public programs in the U.S.. Even seminal works in immigration economics have found that U.S. immigrants

tend to be negatively selected from their origin countries, which have much wider income distributions than the redistributive U.S. (Borjas, 1987). This implies that these less-wealthy, less-educated immigrants migrate in order to receive a greater standard of living through, among other benefits, welfare. However, there is also evidence that shows migrants with similar socioeconomic characteristics as their native counterparts use certain welfare programs at a lower rate likely due to fear of deportation (Watson, 2014). This evidence seems to resonate with the intuition that immigrants "live in the shadows" and seldom use public programs. Ultimately, the theoretical evidence as well as colloquial opinions on immigrant welfare use seem to contradict each other.

In order to resolve this two-part contradiction, this paper will focus the above question by examining the comparable welfare use of immigrants and natives over time. Particularly, I will focus on changing welfare trends since the passage of the 1996 Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) as it could be possible that the contradicting evidence above is the result of changes in welfare use over time. By using Survey of Income and Program Participation (SIPP) data from before and after this influential legislation, I will examine the difference between immigrant and native use of welfare programs as well as the difference in welfare use before and after the enactment of PRWORA. The empirical models will also highlight the potential differences in use of different welfare programs. Using this approach, I will provide evidence for three findings about immigrant welfare use. First, there is no statistically significant difference between the welfare use of similar immigrants and natives. Second, immigrant welfare use decreased after the enactment of PRWORA, which is probably driven by great decreases in Medicaid use over that time period. Third, if Medicaid use is excluded from the measurement of welfare use, immigrants use more of the remaining programs than natives and at an increasing rate after PRWORA. These findings imply that it is key to explore how immigrants and natives may differ in their use of specific welfare programs. After exploring these differing trends in welfare use, I will also propose research to examine whether or not these welfare programs are allowing immigrants to assimilate and contribute to the national economy.

The rest of the paper proceeds as follows. The second section gives background knowledge on the important details of PRWORA as well as the literature focused on immigrant welfare. The

fourth section presents the data and summary statistics. The third section outlines the empirical framework. The fifth section discusses the results. The last two sections conclude by discussing future research possibilities and the implications of this paper's analysis.

## 2 Background

### 2.1 The Impact of PRWORA

Much of the literature on this topic has stemmed from the impact of PRWORA. This welfare reform bill, while allowing refugees and political asylants to access benefit programs, dramatically limited the programs legal permanent residents (LPRs) could access and completely disallowed undocumented immigrants to access any welfare program (Levinson, 2002). Approximately 935,000 noncitizens lost benefits after the passage of PRWORA, an indication of how many immigrants are poor (Fix & Passel, 2002). It is also important to consider the political atmosphere in which this law was passed. In 1996, two other laws—the Antiterrorism and Effective Death Penalty Act and the Illegal Immigration Reform and Immigrant Responsibility Act—were also passed limiting noncitizens' rights of residence and judicial appeal and the ability of undocumented immigrants to obtain legal status. This arguably hostile political atmosphere towards migrants led to the first discussions of "chilling effects" in immigration academia. For example, Zimmermann & Fix (1998) found that noncitizen use of public benefits in Los Angeles County fell following welfare reform and was declining at a faster rate than that of citizens. In the next sub-section, I will further discuss Watson's (2014) analysis of chilling effects.

Amidst the backdrop of PRWORA, it seems plausible to presume that immigrant use of welfare programs could have dropped precipitously after its passage. Nonetheless, a majority of states continued to provide federally funded programs to immigrants if they were given the option to do so (Fremstad, 2004) and some high-immigrant states, e.g. California, provide significant assistance in addition to what is permitted federally (Fix & Passel 2002). In addition, differing program application and enrollment policies could also lead to differing effects than what would be expected after the passage of PRWORA. For instance, it is generally easier to apply for children's medical

assistance programs like SCHIP and Medicaid than for cash assistance or food stamps (Holcomb et al., 2003). The evidence suggests that PRWORA likely had direct impacts on immigrant welfare use, but the extent and manner of that impact is more ambiguous.

## 2.2 Contradicting Findings

There is limited research on immigrants welfare use much before the passage of PRWORA, but the research that does exist tends to support the hypothesis that immigrants use welfare programs more than natives. Borjas & Trejo (1991) originally found that more recent immigrant cohorts use the welfare system more intensively than earlier cohorts and that immigrants who had been in the U.S. for longer were more likely to receive welfare. Nonetheless, Borjas & Trejo use the 1970 and 1980 U.S. Censuses, which do not contain information on non-cash benefit programs including Medicaid, food stamps, and subsidized housing and energy programs. Borjas & Hilton (1996) fittingly use data from the Survey of Income and Program Participation (SIPP), which does contain observations on the use of non-cash benefit programs. The SIPP also is a panel dataset, which records data on families every four months for 32 months in earlier panels and as much as 48 months in the most recent panels. Using data from the 1984, 1985, 1990, and 1991 SIPP, Borjas & Hilton show that the average immigrant uses welfare at a rate of 21 percent whereas the native rate is 14 percent and that immigrants also experience more and longer welfare spells (1996, pg. 583). They also show that although much of this welfare gap diminishes after controlling for socioeconomic differences between migrants and natives, there still is a statistically significant difference between migrants and natives welfare use (1996, pg. 592). Borjas (1999) also finds that immigrant welfare recipients are more often located in relatively high-benefit states, further showing the strong correlation between welfare use and immigrants.

The above papers, although they show a correlation between welfare use and immigrants, all work with data predominantly from before PRWORA. Watson (2014) begins a discussion that is in great contrast with the findings of Borjas & Hilton (1996). Focusing solely on Medicaid, public health insurance for the poor and children, Watson finds that just 30 percent of eligible noncitizen adults were enrolled in Medicaid compared with 57 percent of eligible citizens, indicating



that immigrants use Medicaid much less than natives. Noting that immigration enforcement has dramatically increased since the early 1990s, Watson shows that this increase in enforcement has caused noncitizen mothers to enroll their children in Medicaid far less than what PRWORA would cause on its own. These "chilling effects," according to Watson, have caused immigrant Medicaid use to drop greatly since the early 1990s. Although Watson is only looking at Medicaid, this stands in great contrast to the findings of Borjas & Hilton who assert that even similar immigrants use welfare programs more often than their native counterparts.

The contradicting findings of Borjas & Hilton (1996) and Watson (2014) are perplexing. However, both would seem to fit the two intuitive narratives previously mentioned that dominate the media. The findings of Borjas & Hilton (1996) resonate with the idea that immigrants migrate to the U.S. in order to take advantage of the country's generous welfare programs. Yet the findings of Watson (2014) resonate with the idea that immigrants tend to live in the shadows in the U.S. and do not use welfare if they feel threatened by enforcement policy. It is important to emphasize that Borjas & Hilton (1996) are measuring the total use of all welfare programs whereas Watson (2014) focuses on Medicaid. As noted earlier, varying welfare program application and enrollment policies could certainly create different trends for different programs. Furthermore, Borjas (2003) shows that varying Medicaid cutbacks across states do not reduce health insurance coverage of immigrants as immigrants tend to find insurance through their employers. It may be possible that the decrease in Medicaid enrollment Watson (2014) finds was countered by immigrants finding other insurance and most other immigrant welfare use was not affected after PRWORA. However, the distribution of health insurance between the public and private sector is not the focus of this paper. The contradictory findings from above emphasize the ambiguity of PRWORA's impact on welfare program use. Both Medicaid and other program use must be analyzed more rigorously.

### **3 Data and Sample Statistics**

In the previous section I examined the contrasting findings of Borjas & Hilton (1996) and Watson (2014). It seems likely that the passage of PRWORA had an impact on immigrant welfare use,

but the extent of that impact is unclear. In order to clarify the true use of welfare programs by immigrants compared to the use by natives, it is necessary to reproduce the empirical analysis done by Borjas & Hilton (1996, Section III) using more current data.

Borjas & Hilton originally use data from 1984, 1985, 1990, and 1991 panels of the SIPP. This paper will use the 1990 and 1991 panels as well, but also merge those with new 2001 and 2004 panels. These additional panels will likely account for even the lagged effects of PRWORA. Using two panels each a similar distance in time before and after the enactment of PRWORA, it is most likely that any effects from this legislation would be found within the timeframe of the four panels. The SIPP is a nationally representative survey that interviews households at four-month intervals for a period between 2 and a half to 4 years and includes information on where each respondent was born. For the regression analysis below, I will define the observation's householder as an immigrant if they are either a naturalized citizen or not a citizen at all. As with all immigration economics, it is difficult to measure the differences between documented and undocumented migrants. Nonetheless, the SIPP is unlikely to include many observations from undocumented migrants because the federal Census Bureau conducts it. The results in this paper then will mainly pertain to documented migrants. The SIPP also lends itself well to analysis of individuals' use of welfare as it contains information on a respondent's amount of assistance received either directly through monetary benefits or through in-kind benefits.<sup>1</sup>

Table I presents the mean and standard deviation of the most useful variables in our pooled dataset. The table separates the data by time period (before and after PRWORA) then it also separates the data for natives and migrants. Table II highlights the statistics comparing native and immigrant households' use of specific welfare programs separated for the 1990 and 1991 panels and the 2001 and 2004 panels. Columns 3 and 6 each measure the welfare gap, or the difference in welfare use between natives and immigrants, during each time period before and after PRWORA respectively. In the first row measuring total welfare use, there is a 2.73 percent welfare gap with migrants using more welfare than natives before the passage of PRWORA. While both migrants and

---

<sup>1</sup>The Census Bureau also releases data from a survey called the SIPP Synthetic Beta (SSB). The SSB integrates person-level micro-data from the SIPP to Social Security Administration and Internal Revenue Services records. Although this dataset may provide a more accurate measurement of household's use of welfare, it is accessible by application only. Future research may improve upon this paper by using the SSB.

natives use welfare at a much higher rate after PRWORA, that gap diminishes to only 0.33 percent. Both of these gaps are statistically significant at the one percent level giving greater support to the hypothesis that the dynamics of total welfare use are changing over time for migrants and natives.

In the second row measuring Medicaid use only, there is evidence of what may be driving these changing welfare gaps. Before the passage of PRWORA, the Medicaid gap was 1.18 percent with migrants using more Medicaid than natives (gap again statistically significant at the one percent level). After PRWORA however, that gap flips to 1.88 percent with natives using more Medicaid than migrants. This tremendous decrease in migrant use of Medicaid relative to native use of Medicaid is a statistically significant difference at the one percent level. This is evidence that the dynamics of migrant enrollment in Medicaid must have changed dramatically after the passage of PRWORA while other welfare program enrollment generally increased (see third row measuring welfare excluding Medicaid).

Ultimately, while changing trends in specific program use by both migrants and natives are observable in the summary statistics, these trends could be driven by multiple other factors rather than post-PRWORA program changes. It will be important to control for socioeconomic differences between natives and migrants as well as for state and time fixed effects that could treat migrants and natives differently. Through a more rigorous empirical analysis, I will explore the true significance of these changing gaps across time.

**Table I: Summary Statistics**

	<b>Pre- PRWORA</b>			<b>Post- PRWORA</b>			
	<b>(1) Natives</b>	<b>(2) Migrants</b>	<b>(3) Both</b>	<b>(4) Natives</b>	<b>(5) Migrants</b>	<b>(6) Both</b>	<b>(7) Full Panel</b>
Receiving welfare in a given month (%)	0.12 (0.33)	0.15 (0.36)	0.12 (0.33)	0.18 (0.38)	0.18 (0.38)	0.18 (0.38)	0.16 (0.37)
Receiving Medicaid in a given month (%)	0.09 (0.29)	0.10 (0.30)	0.09 (0.29)	0.14 (0.35)	0.13 (0.33)	0.14 (0.35)	0.13 (0.33)
White (%)	0.78 (0.41)	0.33 (0.47)	0.75 (0.43)	0.73 (0.44)	0.29 (0.45)	0.69 (0.46)	0.71 (0.45)
Black (%)	0.12 (0.32)	0.06 (0.24)	0.11 (0.31)	0.13 (0.34)	0.07 (0.26)	0.13 (0.33)	0.12 (0.33)
Hispanic (%)	0.08 (0.27)	0.41 (0.49)	0.10 (0.30)	0.09 (0.29)	0.41 (0.49)	0.12 (0.32)	0.11 (0.32)
Other (%)	0.02 (0.15)	0.20 (0.40)	0.04 (0.19)	0.05 (0.21)	0.23 (0.42)	0.06 (0.24)	0.05 (0.23)
Female (%)	0.52 (0.50)	0.54 (0.50)	0.52 (0.50)	0.52 (0.50)	0.52 (0.50)	0.52 (0.50)	0.52 (0.50)
Married (%)	0.41 (0.49)	0.63 (0.48)	0.43 (0.49)	0.39 (0.49)	0.64 (.48)	0.41 (0.49)	0.42 (0.49)
Educational attain- ment	2.02 (1.09)	2.17 (1.14)	2.03 (1.09)	2.65 (1.14)	2.47 (1.30)	2.63 (1.16)	2.41 (1.17)
Age	33.27 (22.12)	43.11 (17.45)	33.99 (21.96)	35.43 (22.88)	43.66 (16.63)	36.09 (22.55)	35.43 (22.39)
Total family earned income (monthly)	2698.10 (2578.97)	2592.75 (2551.62)	2690.30 (2577.10)	4260.57 (5181.08)	4169.62 (4913.28)	4253.24 (5160.09)	3764.86 (4572.36)
Observations	2,687,718	214,930	2,902,648	5,872,129	514,453	6,386,582	9,289,230

Table I reports sample means (and standard deviations) for the dependent welfare variables and control variables used in the regression analysis for sub-groups (1-6) as well as the whole panel (7). Educational attainment variable is measured from 1-5 with 1 representing less than high school completion, 2: high school graduate, 3: some college education, 4: college graduate, and 5: post-college education. In

regression analysis, I use binary variables representing the various levels of education rather than what would be implied by the continuous variable above. One observation represents one survey month for an individual, thus there is 9,289,230 observations for the 348,884 individuals in the full panel.

## 4 Empirical Framework

The empirical analysis of this paper will follow very closely to that of Section III in Borjas & Hilton (1996). I will attempt to measure the change in welfare use before and after PRWORA by highlighting the difference between the 1990/1991 and 2001/2004 panels. I will use the four SIPP panels from 1990, 1991, 2001, and 2004 to estimate the regression,

$$P_{ist} = \alpha * I_{ist} + \beta * (I_{ist} * T_{ist}) + \gamma' * X_{ist} + \delta_s + \theta_t + \epsilon_{ist} \quad (1)$$

where  $P$  gives the fraction of time that household  $i$  received a particular type of welfare in a particular state  $s$  and year  $t$ ,  $I_{ist}$  is a binary variable equal to one if the householder is an immigrant and  $T_{ist}$  is a binary variable equal to one if the observation is from either the 2001 or 2004 panels (data that is drawn post-PRWORA). In certain models, I will include  $X_{ist}$ , a vector of socioeconomic and demographic characteristics of the household as well as  $\delta_s$  and  $\theta_t$ , state and time fixed effects respectively. The key coefficient of interest is  $\beta$  as it will measure the change in welfare use by immigrants before and after PRWORA.

**Table II: Changing Trends in Welfare Gaps**

	<b>Pre- PRWORA 1990/1991 Panels</b>			<b>Post- PRWORA 2001/2004 Panels</b>		
	<b>(1) Natives</b>	<b>(2) Migrants</b>	<b>(3) Welfare gap</b>	<b>(4) Natives</b>	<b>(5) Migrants</b>	<b>(6) Welfare gap</b>
All welfare	12.09%	14.82%	+2.73%	17.58%	17.91%	+0.33%
Medicaid only	9.06%	10.24%	+1.18%	14.42%	12.54%	-1.88%
Welfare excluding Medicaid	3.02%	4.58%	+1.56%	3.16%	5.37%	+2.21%
Observations	105612	7632	-	219280	16360	-

Table II highlights the statistical welfare gaps between natives and migrants and across time periods. It is most interesting to see the polar opposite welfare gap for Medicaid use across time periods.

In order to further delve into the contrasting findings of Borjas & Hilton (1996), who use all welfare programs as their dependent variable, and Watson (2014), who uses only Medicaid use as her dependent variable, I will run three main iterations of the above regression equation. The first will measure P for any and all welfare programs used by households. The second will measure P by all welfare programs not including Medicaid used by households. The third and final will measure P by only recording Medicaid use by households. Any differences in the coefficient of interest  $\beta$  between the three regressions could reconcile the contrasting views of Borjas & Hilton and Watson as well as the differences in welfare gaps seen in Table II.

## 5 Results and Discussion

Table III presents five models of interest. In column one, I present the most basic model regressing all welfare use on the two key variables, the binary migrant variable and the interaction term between migrant and the binary post-PRWORA variable. I find significant coefficients at the one percent level indicating that migrants on average use 1.03 percent less welfare than natives and that migrants' welfare use increased by 3.08 percent after PRWORA. It is likely however that this most basic model without any control variables has little internal validity however as indicated by its very

low adjusted R-squared value. Including state and year fixed-effects in the model shown in column 2 leads to the opposite results from the first model. This model indicates that migrants on average use 2.05 percent more welfare than natives while their welfare use decreases by 2.05 percent after PRWORA. Both coefficients are significant at the one percent level. From our discussion above, it is important to examine how greatly socioeconomic differences between migrants and natives could be driving these results. In the model in column 3, I control for a multitude of socioeconomic differences, which drops the significance of the migrant indicator variable implying that there is no significant difference between migrant and native welfare use. The interaction term remains significant at the one percent level however still showing that migrant use of welfare dropped by 1.61 percent after PRWORA. This model will ultimately be the most useful one as including the socioeconomic controls increase the adjusted R-squared value significantly indicating how much socioeconomic differences drive the welfare gaps seen in the summary statistics.

The models in columns 4 and 5 use different dependent variables than the first three models, which regress all welfare use on the specified independent variables. In the column 4 model, which only measures Medicaid use as the dependent variable, I see similar results to those in the model in column 3. The migrant indicator variable remains insignificant, while the interaction term increases in significance and absolute value. This is likely indicative of the great extent that migrant Medicaid use dropped. As Medicaid use makes up the majority of welfare use by both natives and migrants, the results seen in the column 4 model could be driving the results in the column 3 model. There is further evidence of the large role these Medicaid use drops play in the column 5 model. This model excludes Medicaid use in the measurement of welfare leading to dramatically different results. First, for the remaining welfare program use, I find that migrants' use these programs more than natives by 0.49 percent, an estimate that is significant at the five percent level. Further, we find that migrant use of the remaining programs increases by 0.77 percent after PRWORA.

Table III: Differing Trends in Specific Welfare Program Use

	(1) All Welfare	(2) All Welfare	(3) All Welfare	(4) Only Medi- caid	(5) Welfare (No Medi- caid)
<i>Migrant</i>	-0.0103** (0.0039)	0.0205** (0.0039)	-0.0002 (0.0038)	-0.0051 (0.0033)	0.0049* (0.0021)
<i>Migrant*Post- PRWORA</i>	0.0308** (0.0047)	-0.0205** (0.0049)	-0.0161** (0.0045)	-0.0238** (0.0040)	0.0077** (0.0025)
<i>Female</i>			0.0496** (0.0010)	0.0315** (0.0009)	0.0181** (0.0005)
<i>Married</i>			-0.0561** (0.0013)	-0.0432** (0.0011)	-0.0129** (0.0007)
<i>Total family earned income (monthly)</i>			-0.0000** (0.0000)	-0.0000** (0.0000)	-0.0000** (0.0000)
<i>Age</i>			-0.0004** (0.0001)	-0.0022** (0.0001)	0.0018** (0.0001)
<i>Age-squared</i>			-0.0000** (0.0000)	-0.0000** (0.0000)	-0.0000** (0.0000)
<i>High school</i>			-0.0833** (0.0019)	-0.0827** (0.0017)	-0.0006 (0.0010)
<i>Some College</i>			-0.1163** (0.0019)	-0.1085** (0.0017)	-0.0078** (0.0010)
<i>College</i>			-0.1277** (0.0019)	-0.1071** (0.0016)	-0.0206** (0.0009)
<i>Post-grad</i>			-0.1233** (0.0024)	-0.1018** (0.0021)	-0.0215** (0.0011)
<i>Racial controls</i>	yes	yes	yes	yes	yes
<i>State fixed-effects</i>	no	yes	yes	yes	yes
<i>Year fixed-effects</i>	no	yes	yes	yes	yes
Adjusted R-squared	0.00	0.02	0.17	0.15	0.03
Observations	9,289,230	9,288,917	9,288,917	9,288,917	9,288,917

\* p &lt; 0.05; \*\* p &lt; 0.01

Table III reports the OLS coefficients from a regression of all welfare program use (columns 1-3), just



Medicaid use (column 4), and all program use excluding Medicaid (column 5) on the listed independent variables. Column 1 is the most basic model only including the binary migrant variable and the interaction term. Column 2 adds state and year fixed-effects. Column 3 adds the listed socioeconomic controls. Standard errors are adjusted for the 348884 clusters for each individual household.

The differing findings across the last three models seem to resonate with hypotheses proposed earlier in this paper regarding the differing effects on welfare use different welfare programs may have. In the models in columns 3 and 4, I find that there is no statistically significant difference between the welfare use of similar immigrants and natives, but that immigrant welfare use decreased after the enactment of PRWORA, a trend likely driven by decreases in Medicaid use over that time period. In the last model, if Medicaid use is excluded from the measurement of welfare use, immigrants use more of the remaining programs than natives and at an increasing rate after PRWORA. These findings are consistent with the differing findings of Borjas & Hilton (1996) and Watson (2014). If one considers solely means-tested programs excluding Medicaid, it appears that immigrants use either the same or a greater amount of these programs than their native counterparts, a finding consistent with those of Borjas & Hilton. However, if one considers Medicaid along with other welfare programs, it appears that drops in Medicaid use since PRWORA have driven a decreasing trend in migrant welfare use, a finding consistent with those of Watson. Ultimately, I have shown that it is important to consider how one measures welfare use and what programs one considers when comparing migrants and natives use of these programs.

## 6 Long-term Impacts of Welfare

Now that there is a clearer picture of the current trend of welfare use for immigrant households, it is then important to discuss whether or not welfare is doing its job for immigrants; that is, are welfare programs helping immigrant households move out of poverty and be socially mobile? This question is most important as it will show whether the current state of welfare is helping immigrants assimilate into the nation's economy or if it is mainly a source of deadweight loss. Even Borjas & Hilton discuss in their conclusion, "little is known about the long-run impact of welfare dependency in the immigrant generation in terms of the economic and social outcomes of second-generation

Americans” (1996, pg. 602). Nonetheless, data from the SIPP will not be useful to answer this question about long-term impacts of welfare as the time period of observation is 48 months at a maximum.

A possible dataset that could allow for measurements of the long-term impact of welfare programs for immigrants in the National Longitudinal Study of Youth (NLSY). The NLSY 1979 cohort (NLSY79) follows the lives of a sample of American youth, ages 14-22 in 1979, born between 1957 and 1964. The cohort of 12,686 respondents was interviewed a total of 25 times between 1979 and 2012. Using this dataset, it may be possible to track the long-term outcomes of children of immigrants who belonged to welfare-receiving families during their youth. Unfortunately, there are many problems with this dataset however. First, in the sample of 12,686, only 874 belonged to immigrant families, and only 268 of those ever belonged to a family that received welfare. This meager sample size could greatly discredit any findings of an empirical analysis of the sample. Second, the external validity of any results would also have to be questioned as the welfare system between 1957 and 1979 functioned much differently than it does today. With these limitations in mind however, it may still be useful to run some preliminary empirical analysis.

Using only the observations from the respondents who are children of immigrants, the regression equation that would estimate these long-term impacts of welfare is

$$w_{it} = \alpha + \beta * I_{it} + \gamma * (I_{it} * P_{it}) + \delta' * X_{it} + \epsilon_{ist} \quad (2)$$

where  $w_{it}$  gives the current wage of individual  $i$  during time  $t$ ,  $I_{it}$  represents a binary variable indicating that individual  $i$  is the child of an immigrant household,  $P_{it}$  represents a binary variable equal to one if individual  $i$  ever belonged to a family that was enrolled in a welfare program, and  $X_{it}$  is a vector of socioeconomic and demographic characteristics of the individual when they were first interviewed. The coefficient of interest,  $\gamma$ , will measure the effect of being on welfare as child on the wage of a child of an immigrant household.

While the data that would allow for this kind of long-term analysis is limited, larger and more detailed long-term longitudinal datasets will only become more prevalent in the future. It will not

only be of great interest to the American public whether or not welfare programs are leading to better outcomes for immigrants and their children, but also to prospective immigrants deciding which country will give their children the best future. The field of welfare and immigration economics is ripe with opportunity.

## 7 Conclusion

This paper examines the various welfare gaps between immigrants and natives across time paying particular attention to before and after the passage of the 1996 PRWORA. Through this analysis, I contribute that it is vitally important to show that different welfare programs, especially Medicaid, may have different trends in use patterns by natives and migrants. In particular, three findings motivate this assertion. First, there is no statistically significant difference between the welfare use of similar immigrants and natives. Second, immigrant welfare use decreased after the enactment of PRWORA, a decrease mainly driven by great decreases in Medicaid use over that time period. Third, if Medicaid use is excluded from the measurement of welfare use, immigrants use more of the remaining programs than natives and at an increasing rate after PRWORA. While the previous section discussed the large area of potential research focusing on the long-term impacts of welfare use by migrants, the findings of this paper could be further clarified as well. While this paper shows that PRWORA or other similar legislature likely had some effect on migrant use of welfare, it is difficult to show the exact mechanisms in which this effect occurred. For instance, Watson's finding of increased enforcement leading to lower rates of Medicaid enrollment could still be valid in playing a large role in discouraging migrant enrollment in Medicaid. It would be useful to isolate the effects of both PRWORA and increasing enforcement by controlling for enforcement in the empirical analysis done in this paper. Still, the findings of this paper indicate that dramatic changes in the way immigrants use welfare occurred after PRWORA.

## References

- [1] Borjas, G. J. (1987). Self-selection and the earnings of migrants. *American Economic Review*, 77(4), 531-53
- [2] Borjas, G. J., & Trejo, S. J. (1991). Immigrant participation in the welfare system. *Industrial & labor relations review*, 44(2), 195-211.
- [3] Borjas, G. J., & Hilton, L. (1996). Immigration and the welfare state: Immigrant participation in means-tested entitlement programs (No. w5372). *National Bureau of Economic Research*.
- [4] Borjas, G. J. (1999). Immigration and welfare magnets. *Journal of labor economics*, 17(4), 607-637.
- [5] Borjas, G. J. (2003). Welfare reform, labor supply, and health insurance in the immigrant population. *Journal of Health Economics*, 22(6), 933-958.
- [6] Center for Economic and Policy Research. 2014. SIPP Uniform Extracts, Version 2.1.7 . Washington, DC.
- [7] Fix, M., & Passel, J. (2002). Scope and Impact of Welfare Reform's Immigrant Provisions, The. *The Urban Institute*.
- [8] Fix, M., & Zimmermann, W. (2002). All Under One Roof: Mixed-Status Families in an Era of Reform. *International Migration Review*, 35(2), 397-419.
- [9] Fremstad, S. (2004). Recent welfare reform research findings: Implications for TANF reauthorization and state TANF policies. *Washington, DC Center on Budget and Policy Priorities*.
- [10] Holcomb, P. A. (2003). The application process for TANF, food stamps, Medicaid and SCHIP: Issues for agencies and applicants, including immigrants and limited English speakers. *The Urban Institute*.
- [11] Levinson, A. (2002). Immigrants and welfare use. *Migration Information Source*.

- [12] Watson, T. (2014). Inside the Refrigerator: Immigration Enforcement and Chilling Effects in Medicaid Participation. *American Economic Journal: Economic Policy*, 6(3), 313-338.

# Crude Oil Price: An Indicator of Consumer Spending

Natalie Li

Eva Lin

New York University

May 2015

## **Abstract**

As the most major source of energy, crude oil has a significant impact on political and economic dynamics around the globe. The price of oil, therefore, directly and indirectly affects our daily lives. Higher oil prices reduce discretionary income while driving up prices of other consumer goods and services, thereby impacting domestic consumption. In this paper, we hypothesize that the real price of imported crude oil negatively influences personal consumption expenditure within the United States. To test the theory, we present a least squares model by regressing oil prices as well as other macroeconomic sets of time series data against personal consumption expenditure. The results indicate a structural change in 1980, aligning with the introduction of regulations for crude oil price control following the OPEC embargo. Our results are significant at the 99% confidence interval that real imported crude oil prices negatively impact real domestic consumer expenditure.

## **1 Introduction**

According to the US Energy information Administration's (EIA) latest estimate for the first quarter in 2015, the world consumes around 91.13 million barrels of crude oil per day. As one of the major energy sources in the world, crude oil dominates as the subject of both domestic and foreign energy policy. Its importance lies in its chemical potential, which is harnessed by oil refineries to produce

gasoline, heating oil, jet fuel and propane. Crude oil allows for effective transportation and efficient heating and electricity generation, defining the experience of modern-day consumers. Because of its crucial role in modern life, consumers view refined oil, such as gasoline, as an essential expense and therefore inelastic in demand with regards to price. Lower oil prices increase consumers' discretionary income via the income effect, thereby encouraging consumption in the economy. Higher oil prices, however, prompt the opposite to occur and decrease consumption. The price of oil is determined through global supply and demand, a delicate balance that, if disrupted, can have severe economic repercussions in the forms of low consumer confidence, high inflation, and an overall decrease in consumer spending.

The objective of this paper, therefore, is to examine the relation between oil prices and consumer spending in an oil-importing country: the United States. We hypothesize that the import price of crude oil has a negative impact on domestic personal consumption expenditure due to its microeconomic implications. The empirical methodology is to utilize the Ordinary Least Squares method to create a multivariable regression model in order to examine the United States from 1970 to 1995. The time period is carefully selected so that it covers the two oil price shocks in the 1970s, capturing an economy that is reliant on foreign oil and the restrictions it poses on domestic consumption expenditure.

Furthermore, we make necessary modifications to our model by modifying variables to avoid high multicollinearity, detecting structural changes by running the Chow Test, and implementing a heteroscedasticity test. The general expectation here is a negative correlation between real crude oil prices and real personal consumption expenditure.

## **2 Literature Review**

As the major source of energy, crude oil is no doubt one of the most widely studied commodities in the global economy. It has been extensively researched and used as an indicator for economic forecasting. Numerous studies have examined the relations between the price of oil and the economy of oil-imported countries. Hamilton's (1983) study recognizes oil shocks as a contributing factor

to economic downturns in the United States after World War II, suggesting an increase in oil prices could lead to a vast decline in total output. Some authors base their studies on foreign economies, stressing the impact of oil prices in the economy on a macroeconomics scale. Hanabusa (2009) investigates the relation between oil prices and economic growth in Japan. He suggests a causal relationship and argues that the price of oil is an important predicator of economic growth. TovarJalles (2009), on the other hand, analyzes the impact of oil prices on economic performance in France and notes that the price of oil has exerts less influence on macroeconomic variables as time progresses. Regarding oil prices and spending specifically, Odusami (2010) explains the short-term deviations of consumption, asset wealth and labor income from their long-term trend as a result of crude oil price fluctuations. He further recognizes consumption as a proportion of aggregate wealth, noting that as oil prices increases, the proportion of aggregate wealth consumed decreases. Odusami's (2010) study parallels with our hypothesis and provides a different angle to examine the crude price-consumption relation. Having investigated the causes of post-WWII oil shocks as well as their effects in the aftermath, Hamilton (2013) notes that price control is an important factor to be considered with the events of drastic oil price increases. He further points out that after the oil price spike in the 1970s caused by the embargo of Organization of Petroleum Exporting Countries (OPEC), there were numerous sets of elaborate regulations particularly introduced for oil price control. Interestingly, Hamilton (1983), Odusami (2009), and Petersen's (2005) studies all suggest weakening effect of oil prices on consumption after 1980, when the regulatory environment's effect came into the picture as oil prices started to fluctuate in different directions. Namely, the United States introduced a multitude of measures for further oil independence as a way of increasing economic stability. This allows us to test our theory in two separate periods-before and after the change in environment and movements of crude oil prices.

### 3 The Model

The model we use for this paper is a multivariable linear regression model that examines time series data. We study the time period from 1970 to 1995 and employ quarterly instead of annual data



due to the volatile nature of crude oil prices. The model we will be estimating with Ordinary Least Squares method is introduced as the following equation:

$$PC_t = \beta_1 + \beta_2 OP_t + \beta_3 Z_t + \beta_5 P_t \quad (1)$$

where  $Z_t = CPI_t + 0.6345CCI_t$

#### **Variable Codes:**

Dependent Variable:

- $PC_t$  Real Personal Consumption Expenditures (billions of chained 2009 dollars)

Independent Variables:

- $OP_t$  Real Imported Crude Oil Prices (\$/barrel)
- $Z_t$  - CCI and CPI Combination Variable
- $P_t$  - Population Growth Rate (%)

**Real Personal Consumption Expenditures** Real Personal Consumption Expenditures is the dependent variable in our model. It is a measure of goods and services consumed by households, including durable goods, non-durable goods and services. There are two categories of personal consumption expenditures: real and nominal. Real values are inflation adjusted while nominal values are not. Here, for the convenience of comparison, we will be using Real Personal Consumption Expenditures in billions of chained 2009 dollars, seasonally adjusted.

**Real Imported Crude Oil Prices** Oil prices can be defined as the prices of crude oil, gasoline, or heating oil, etc. Since we are examining the movement in original oil price rather than the prices of refined products, which would reflect the price of manufacturing process, we will use real imported price of crude oil, in dollars per barrel, as an indicator of oil prices in our model. We expect its coefficient to be negative because consumers have less discretionary income as the price of oil increases.

**Population Growth Rate** Population growth rate is defined as the rate at which the number of individuals in a population increases. We assume that an increase in population growth rate will lead to an increase in personal consumption expenditure, as there are more people to consume in the economy; in other words, the coefficient of P is expected to be positive. However, in a rare case where population growth rate is decreasing, yet the total population still remains higher than the initial population, the coefficient of P could be negative. The formula to derive P is the following:

$$P = \frac{\text{Population of year } t - \text{Population of start year}}{\text{Population of start year}} \quad (2)$$

**Consumer Confidence Index** CCI is a surveyed indicator to measure consumer's optimism in state of the economy in the present and near future. Consumers have two options facing their disposable income: to consume or to save. When consumers feel more confident and optimistic, they tend to spend more on goods and services than they save. As a result, CCI is expected to have a positive impact on personal consumption expenditure.

**Consumer Price Index** CPI measures the weighted average change in the price level of consumer goods and services purchased by households. The US Bureau of Labor collects two types of CPI data: CPI for Urban Wage Earners and the Chained CPI for All Urban Consumers. Since CPI for Urban Consumers represents a larger population in the general public, we will be using this set of CPI data in the model. In theory, an increase in the price level leads to inflation, causing interest rates to fall and encouraging consumers to spend rather than save because of the relatively low return from savings. Therefore, we expect a positive relationship between CPI and personal consumption expenditure.

**Consumer Confidence Index and Consumer Price Index Combination Variable** Due to the strong correlation of CCI and CPI, we obtain estimate of  $\beta_4 = 0.6345\beta_3$  from cross section analysis and introduce a combined variable Z (=CPI+0.6345CCI) in our regression model to resolve the problem of multicollinearity. Doing so will allow us to estimate the combined effect of CCI and CPI on personal consumption expenditure rather than the individual effects.

## 4 Data

The data in this paper is shown in Table 4.1. The five sets of data were extracted from several database sources online. From the website of US Energy Information Administration (AIE), we accessed quarterly data of real imported crude oil price from 1974 to 1980 and annual data from 1970 to 1973, the latter of which was repeated four times to fit into our quarterly data format. We were also able to collect the data of real personal consumption expenditure and CPI on the website of the Federal Reserve Bank of St. Louis (FRED) under Real Personal Consumption Expenditures and Consumer Price Index for All Urban Consumers. The population growth rate data is available on the website of the World Bank in annual form. We intentionally repeated annual data four times as an adjustment in order to align our data sets. As for the CCI data from the Organization for Economic Co-operation and Development, which is collected monthly, it was manually transformed into quarterly data by averaging every three month's figures. Note that the data of Z was calculated as  $CPI_t + 0.6345CCI_t$ .

## 5 Results

The estimated regression line of our model from 1980Q1 to 1995Q4 is given as the following equation:

$$PC_t = 393.6859 - 5.70934OP_t + 28.1758Z_t - 4.6115P_t \quad (3)$$

As shown in Table 5.1, the  $R^2$  in this model is extremely high with a value of 0.980394, which suggests that the independent variables in this model explains roughly 98% of the variation in the real personal consumption. As for the estimates, the intercept is significant yet possesses no economically meaningful interpretation since the oil price, CCI, CPI, and population growth rate are unlikely to be zero at the same time. The signs of OP and Z coefficients are as expected, indicating a negative impact of oil prices and a positive combined CCI-CPI impact on personal consumption. Nonetheless, the results also indicate problems that need to be examined. With 104 observations, population growth rate's t value of -0.038988 implies that P is not statistically

significant in this model.

In order to improve the model, we ran a Chow Test for 1980Q1 as it is suspected to be the turning point in oil price history. As seen in Table 5.2, the test result is significant and thus implies the presence of an appropriate structural break point in our time series model. In other words, regression on subintervals of [1970Q1, 1979Q4] and [1980Q1, 1995Q4] gives a better modeling than the regression over the whole interval of [1970Q1, 1995Q4]. As a result, we will proceed by running the regressions of the two time periods separately. The estimated regression line of our model from 1970Q1 to 1979 Q4 is introduced in the equation below:

$$PC_t = -5.7027 - 6.0451OP_t + 39.2359Z_t - 772.7508P_t \quad (4)$$

As seen in Table 5.3, the regression of this subinterval also achieves a high  $R^2$  of 0.981958, or 98% of explained variation in PC. Furthermore, all the estimates for the exogenous variables are significant on 99% confidence interval according to the t-test results. The negative sign of OP confirms our hypothesis and indicates about a 6.05 billion dollar decrease in Personal Consumption Expenditure with every one-dollar increase in per barrel oil prices. As expected, the coefficient of Z indicates a positive combined CCI-CPI impact on consumption. The negative sign of the population growth rate, however, does not match up with our original expectation based on the theory of larger population makes more consumption. Yet as we compare Graph 6.1 and Graph 6.3 from 1970 to 1979, it is true that as Personal Consumption Expenditure exhibits a rising trend, population growth rate was decreasing during the period. This could be explained by the slower yet steady increase in total US population, referring to Table 6.4 from the US Census Bureau. The intercept in this regression is not significant; however, we are not going to further interpret it because it has no intrinsic meaning in our model. One potential problem in this regression could be heteroscedasticity. In order to investigate, we plot the residuals, as shown in Graph 5.1. Since the residuals are roughly evenly scattered, we conclude that heteroscedasticity is not a major problem in this regression. The overall performance of this model for the subinterval of [1970Q1, 1979Q4] is quite satisfactory.

For the second subinterval, we run regression with data from 1980Q1 to 1995Q4. The second estimated regression line of our model is given as the following equation:

$$PC_t = -567.8053 - 4.3630OP_t + 34.3119Z_t - 428.4218P_t \quad (5)$$

As seen in Table 5.4, we again get a large  $R^2$  of 0.979784. On average, every dollar per barrel increase in OP is estimated to lead to a decrease of around 4.36 billion dollars in Personal Consumption Expenditures. This is smaller than our estimation of the OP coefficient for the period of 1970 to 1979 of 6.05 billion dollars, demonstrating the effect of a structural break caused by the U.S. policy changed for greater energy independence in 1980. Every unit increase in Z, on the other hand, is estimated to have a positive impact of around 34.31 billion dollar increase on consumption expenditures. Although the coefficients of OP and Z have the right signs and are both significant at a 99% confidence interval, P turns out to be only significant at a 90% confidence interval. We then plot the residuals for this subinterval to test heteroscedasticity. Graph 5.2 displays an abnormal distribution; the residuals exhibit a heteroscedastic pattern. In order to improve on the model, Whites Correction for Heteroscedasticity is implemented. Comparing Table 5.5 and Table 5.4, however, one can see that the results are not immensely improved: the intercept and P remain insignificant. This might be a result of the irrelevance using population growth rate instead of total population or an outcome of multicollinearity among our macroeconomic exogenous variables. The regression results of the second subinterval in this model are not as satisfactory, yet we achieved our objective of examining oil prices-consumption relation. We have seen significant and negative OP estimates in all of the regressions we run throughout the study, which support our hypothesis that oil prices do exert an influence on personal consumption decisions.

## 6 Conclusions

This paper helps us understand one of the factors that affect consumer-spending decisions—the price of oil. It provides empirical evidence that increases in oil prices have a negative effect on consumer spending. The result is reached by running a multiple linear consumption model, with

relevant macroeconomic factors as exogenous variables. This study aligns its finding with other earlier research that a structural change occurs in 1980, altering the relationship between oil prices and consumer expenditure. Under the model with two subintervals before and after 1980 —the price of real imported crude oil negatively impacts personal consumption in both time periods. The latter period, however, shows oil price as a weaker influence on consumer spending as compared to the former. This is consistent with the increased regulatory changes regarding oil price control by the United States around 1980. The study does not capture the relation between population growth and consumption expenditure. Theoretically, population growth should have a positive effect on consumer spending, yet the model indicates that it has no significant impact. Perhaps using population instead of population growth rate as one of the dependent variables would improve the model and give better results. Overall, the model has produced a satisfactory result that demonstrates the impact of crude oil prices as an indicator of consumer spending.

## 7 Appendices

Figure 1: Table 5.1

Dependent Variable: PC				
Method: Least Squares				
Sample: 1970Q1 1995Q4				
Included observations: 104				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	393.6859	144.382	2.726696	0.0076
OP	-5.709338	0.738754	-7.728339	0
Z	28.17584	0.42875	65.71631	0
P	-4.611511	118.2801	-0.038988	0.969
R-squared	0.980394	Mean dependent var		4527.654
Adjusted R-squared	0.979806	S.D. dependent var		1089.026
S.E. of regression	154.7573	Akaike info criterion		12.9593
Sum squared resid	2394983	Schwarz criterion		13.061
Log likelihood	-669.8834	Hannan-Quinn criter.		13.0005
F-statistic	1666.828	Durbin-Watson stat		0.08093
Prob(F-statistic)	0			

Figure 2: Table 5.2

Chow Breakpoint Test: 1980Q1			
Null Hypothesis: No breaks at specified breakpoints			
Varying regressors: All equation variables			
Equation Sample: 1970Q1 1995Q4			
F-statistic	38.29507	Prob. F(4,96)	0
Log likelihood	99.19816	Prob. Chi-Sq.	0
Wald Statistic	153.1803	Prob. Chi-Sq.	0

Figure 3: Table 5.3

Dependent Variable: PC				
Method: Least Squares				
Sample: 1970Q1 1979Q4				
Included observations: 40				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5.702674	142.3384	-0.040064	0.9683
OP	-6.045118	0.832232	-7.263739	0
Z	39.23592	1.415224	27.72418	0
P	-772.7508	93.05922	-8.30386	0
R-squared	0.981958	Mean dependent var		3446.428
Adjusted R-sq	0.980455	S.D. dependent var		354.0456
S.E. of regres	49.49738	Akaike info criterion		10.73636
Sum squared	88199.66	Schwarz criterion		10.90524
Log likelihood	-210.7271	Hannan-Quinn criter.		10.79742
F-statistic	653.116	Durbin-Watson stat		0.96979
Prob(F-statist	0			

Figure 4: Table 5.4

Dependent Variable: PC				
Method: Least Squares				
Sample: 1980Q1 1995Q4				
Included observations: 64				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-567.8053	369.0648	-1.538498	0.1292
OP	-4.362893	1.434718	-3.040942	0.0035
Z	34.3119	2.24424	15.28887	0
P	-248.4218	150.7624	-1.647771	0.1046
R-squared	0.979784	Mean dependent var		5203.42
Adjusted R-sq	0.978773	S.D. dependent var		809.4565
S.E. of regres	117.9333	Akaike info criterion		12.43858
Sum squared	834496.2	Schwarz criterion		12.57351
Log likelihood	-394.0345	Hannan-Quinn criter.		12.49173
F-statistic	969.312	Durbin-Watson stat		0.144586
Prob(F-statist	0			

Figure 5: Table 5.5

Dependent Variable: PC				
Method: Least Squares				
Sample: 1980Q1 1995Q4				
Included observations: 64				
White heteroskedasticity-consistent standard errors & covariance				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-567.8053	310.5748	-1.82824	0.0725
OP	-4.362893	1.023266	-4.263694	0.0001
Z	34.3119	2.273199	15.0941	0
P	-248.4218	156.364	-1.58874	0.1174
R-squared	0.979784	Mean dependent var		5203.42
Adjusted R-sq	0.978773	S.D. dependent var		809.4565
S.E. of regres	117.9333	Akaike info criterion		12.43858
Sum squared	834496.2	Schwarz criterion		12.57351
Log likelihood	-394.0345	Hannan-Quinn criter.		12.49173
F-statistic	969.312	Durbin-Watson stat		0.144586
Prob(F-statist	0	Wald F-statistic		868.9007
Prob(Wald F-	0			

Figure 6: Graph 5.1

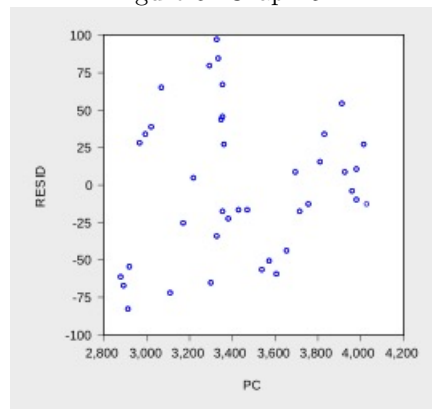




Figure 7: Graph 5.2

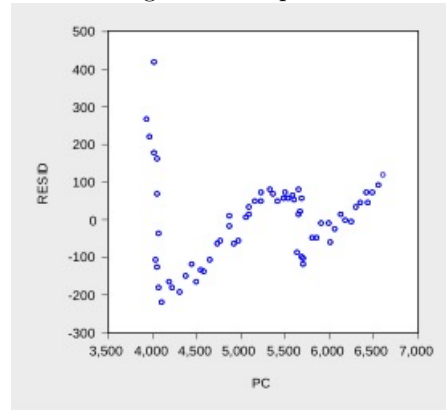


Figure 8: Graph 6.1

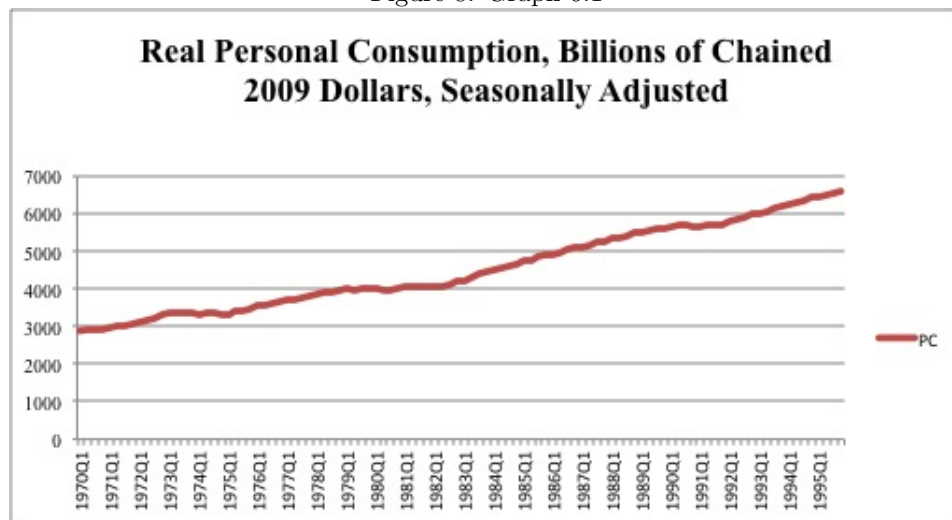


Figure 9: Graph 6.2

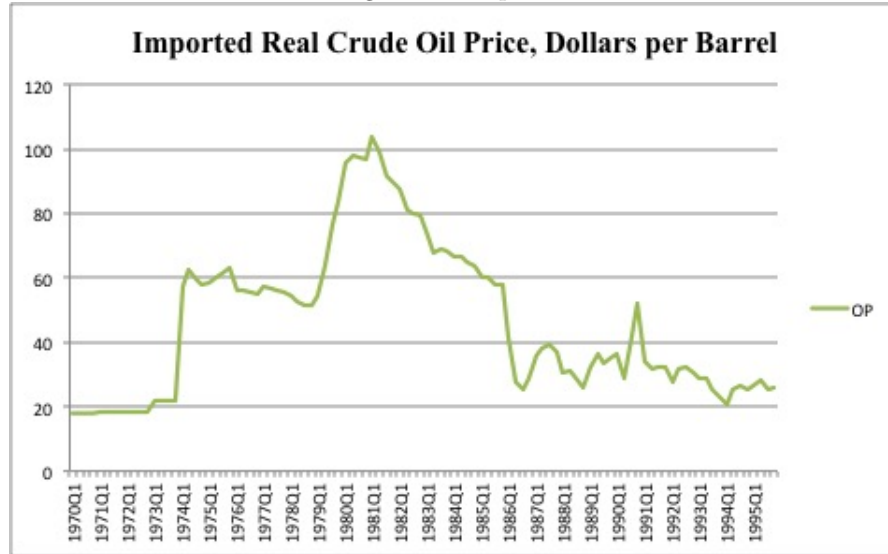


Figure 10: Graph 6.3

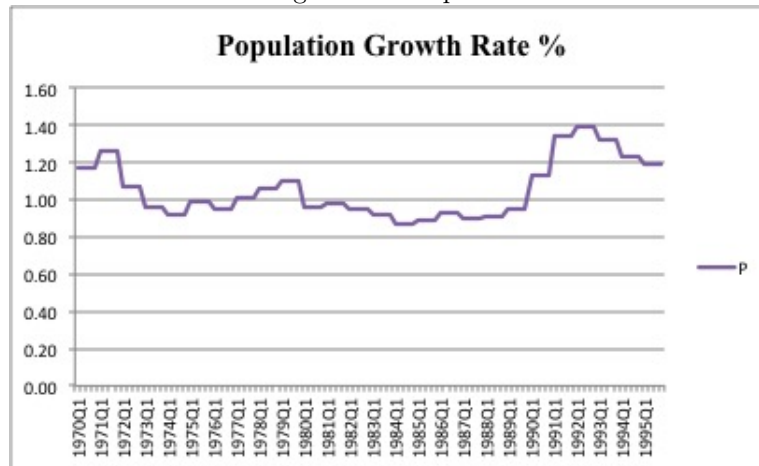


Figure 11: Graph 6.4

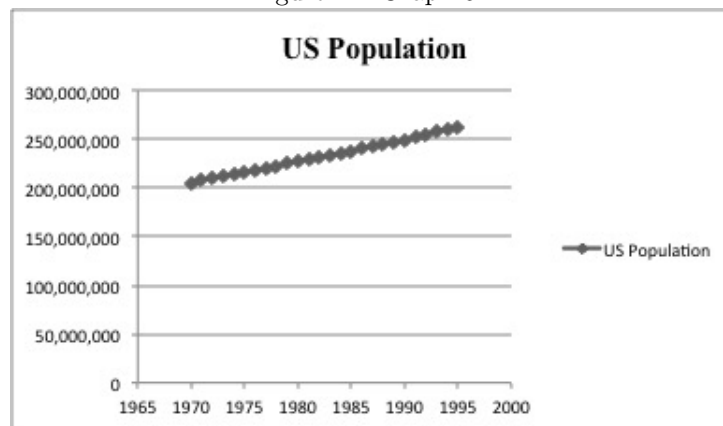


Figure 12: Table 4.1

Quarter	OP	P	CCI	CPI	PC	Z
1970Q1	17.970	1.165	98.926	38.100	2882.300	100.868
1970Q2	17.970	1.165	98.838	38.633	2895.600	101.345
1970Q3	17.970	1.165	98.736	39.033	2921.100	101.681
1970Q4	17.970	1.165	98.642	39.600	2913.100	102.189
1971Q1	18.430	1.264	99.108	39.933	2968.900	102.817
1971Q2	18.430	1.264	99.373	40.300	2996.100	103.332
1971Q3	18.430	1.264	99.499	40.700	3020.000	103.832
1971Q4	18.430	1.264	99.992	41.000	3070.200	104.445
1972Q1	18.140	1.071	100.553	41.333	3110.800	105.134
1972Q2	18.140	1.071	100.624	41.600	3170.200	105.446
1972Q3	18.140	1.071	100.850	41.933	3219.100	105.927
1972Q4	18.140	1.071	100.187	42.367	3294.600	105.936
1973Q1	21.640	0.955	99.315	43.033	3354.800	106.040
1973Q2	21.640	0.955	98.705	43.933	3353.400	106.561
1973Q3	21.640	0.955	98.552	44.800	3365.300	107.331
1973Q4	21.640	0.955	98.171	45.933	3355.500	108.272
1974Q1	57.425	0.914	97.643	47.300	3326.200	109.256
1974Q2	62.786	0.914	97.942	48.567	3337.900	110.711
1974Q3	59.705	0.914	97.244	49.933	3351.600	111.634
1974Q4	57.651	0.914	96.744	51.467	3302.500	112.851
1975Q1	58.353	0.906	97.283	52.567	3330.100	114.294
1975Q2	60.129	0.986	98.478	53.200	3385.700	115.684
1975Q3	61.232	0.986	98.752	54.267	3434.100	116.925
1975Q4	63.229	0.906	99.150	55.267	3470.300	118.177
1976Q1	56.256	0.950	99.716	55.900	3539.900	119.170
1976Q2	56.070	0.950	99.966	56.400	3572.400	119.029
1976Q3	55.567	0.950	100.274	57.300	3610.300	120.924
1976Q4	55.076	0.950	100.122	58.133	3657.500	121.660
1977Q1	57.226	1.006	100.251	59.200	3699.300	122.809
1977Q2	56.845	1.006	100.424	60.233	3719.700	123.952
1977Q3	56.064	1.006	100.146	61.067	3755.200	124.610
1977Q4	55.652	1.006	99.753	61.967	3811.800	125.262
1978Q1	54.194	1.060	99.537	63.033	3893.800	126.202
1978Q2	52.921	1.060	98.440	64.467	3915.600	127.562
1978Q3	51.746	1.060	99.299	65.967	3932.000	128.972
1978Q4	51.546	1.060	98.594	67.500	3963.500	130.050
1979Q1	56.160	1.104	98.130	69.200	3983.600	131.463
1979Q2	63.479	1.104	97.649	71.400	3981.300	133.359
1979Q3	76.777	1.104	97.358	73.700	4020.400	135.474
1979Q4	83.413	1.104	97.323	76.033	4031.200	137.784
1980Q1	95.737	0.960	97.107	79.033	4025.000	140.648
1980Q2	98.308	0.960	96.361	81.700	3934.100	142.811
1980Q3	97.534	0.960	97.788	83.233	3976.900	145.279
1980Q4	96.685	0.960	98.376	85.567	4029.600	147.987
1981Q1	103.706	0.981	97.908	87.933	4050.800	150.055
1981Q2	99.099	0.981	98.480	89.767	4050.100	152.253
1981Q3	91.718	0.981	98.586	92.267	4066.400	154.820
1981Q4	90.036	0.981	97.746	93.767	4035.900	155.787
1982Q1	87.320	0.953	97.599	94.600	4062.600	156.526
1982Q2	81.443	0.953	97.563	95.967	4077.600	157.871
1982Q3	79.977	0.953	97.772	97.633	4109.100	159.669
1982Q4	79.569	0.953	98.262	97.933	4184.100	160.280
1983Q1	72.810	0.914	98.871	98.000	4224.800	160.734
1983Q2	67.930	0.914	100.519	98.133	4308.400	162.912
1983Q3	68.883	0.914	100.606	100.100	4384.000	163.935
1983Q4	68.398	0.914	100.730	101.100	4453.100	165.013
1984Q1	66.354	0.866	101.433	102.533	4490.900	166.905
1984Q2	66.423	0.866	101.325	103.500	4554.900	167.791
1984Q3	65.140	0.866	101.430	104.400	4509.500	168.750
1984Q4	63.837	0.866	101.154	105.300	4650.600	169.482
1985Q1	60.407	0.886	100.950	106.267	4729.700	170.320
1985Q2	60.375	0.886	100.943	107.333	4774.100	171.282
1985Q3	58.015	0.886	100.762	107.900	4865.800	171.834
1985Q4	57.706	0.886	100.637	109.000	4878.300	172.967
1986Q1	41.867	0.924	101.093	109.567	4919.600	173.710
1986Q2	27.723	0.924	101.252	109.033	4974.600	173.277
1986Q3	25.506	0.924	101.100	109.700	5064.700	173.848
1986Q4	28.716	0.924	100.663	110.467	5007.100	174.337
1987Q1	35.531	0.894	100.513	111.800	5097.900	175.575
1987Q2	38.136	0.894	100.732	113.067	5168.600	176.981
1987Q3	39.292	0.894	100.801	114.267	5228.500	178.225
1987Q4	36.728	0.894	100.179	115.333	5239.500	178.897
1988Q1	30.782	0.908	100.810	116.233	5332.700	180.070
1988Q2	31.424	0.908	100.946	117.567	5371.800	181.617
1988Q3	28.346	0.908	101.134	119.000	5417.700	183.169
1988Q4	26.004	0.908	100.961	120.300	5479.700	184.360
1989Q1	32.476	0.944	101.024	121.667	5505.000	185.767
1989Q2	36.128	0.944	100.825	123.633	5530.900	187.480
1989Q3	33.281	0.944	100.715	124.600	5585.900	188.503
1989Q4	35.241	0.944	100.710	125.867	5610.500	189.773
1990Q1	36.322	1.130	100.635	128.033	5658.700	191.886
1990Q2	29.029	1.130	100.640	129.300	5676.400	193.096
1990Q3	41.335	1.130	99.059	131.533	5699.300	194.596
1990Q4	52.280	1.130	97.583	133.767	5656.200	195.684
1991Q1	34.990	1.136	98.575	134.767	5646.700	197.313
1991Q2	31.525	1.136	99.490	135.567	5684.000	198.693
1991Q3	32.093	1.136	99.499	136.600	5711.600	199.732
1991Q4	32.141	1.136	98.382	137.733	5710.100	200.141
1992Q1	27.430	1.187	98.194	138.667	5817.300	200.971
1992Q2	31.452	1.107	99.057	139.733	5857.200	202.585
1992Q3	32.495	1.187	98.846	140.800	5920.600	203.518
1992Q4	30.299	1.187	99.640	142.033	5991.100	205.255
1993Q1	28.554	1.119	100.183	143.067	6013.800	206.633
1993Q2	28.083	1.119	99.510	144.100	6067.800	207.234
1993Q3	25.381	1.119	99.046	145.767	6134.800	207.611
1993Q4	22.737	1.119	99.784	145.967	6189.100	209.280
1994Q1	20.086	1.226	100.751	146.700	6260.100	210.627
1994Q2	23.217	1.226	100.887	147.533	6308.600	211.419

## References

- [1] Hamilton, James D. "Oil and the Macroeconomy since World War II." *Journal of Political Economy* 91, no. 2 (1983): 228 - 48.
- [2] Hanabusa, Kunihiro. "Causality Relationship between the Price of Oil and Economic Growth in Japan." *Energy Policy* 37, no. 5 (2009): 1953-957.
- [3] TovarJalles, Joao. "Do Oil Prices Matter? The Case of a Small Open Economy." *Annals of Economics and Finance* 10, no. 1 (2010): 65-87.
- [4] Odusam, Babatunde Olatunji. "To Consume or Not: How Oil Prices Affect the Comovement of Consumption and Aggregate Wealth." *Energy Economics* 32, no. 4 (2010).
- [5] Hamilton, James D. "Historical Oil Shocks." In *Routledge Handbook of Major Events in Economic History*, 239-65. 2013.
- [6] He, Yanan, Shouyang Wang, and Kin Keung Lai. "Global Economic Activity and Crude Oil Prices: A Cointegration Analysis." *Energy Economics* 32, no. 4 (2010): 868-76.
- [7] Gounder, Rumania, and Matthew Bartleet. "Oil Price Shocks and Economic Growth: Evidence for New Zealand, 1989-2006." 2007.
- [8] OECD (2015), Consumer confidence index (CCI) (indicator). doi: 10.1787/46434d78-en (Accessed on 10 May 2015)
- [9] "Short-Term Energy Outlook." U.S. Energy Information Administration. Accessed May 10, 2015. <http://www.eia.gov/forecasts/steo/realprices/>.
- [10] Board of Governors of the Federal Reserve System (US), Effective Federal Funds Rate[FEDFUNDS], retrieved from FRED, Federal Reserve Bank of St. Louis <https://research.stlouisfed.org/fred2/series/FEDFUNDS/>, May 1, 2015.

- [11] US. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis <https://research.stlouisfed.org/fred2/series/CPIAUCSL/>, April 25, 2015.
- [12] "Historical National Population Estimates: July 1, 1900 to July 1, 1999." U.S. Census Bureau. April 11, 2000. Accessed May 10, 2015. <https://www.census.gov/popest/data/national/totals/pre-1980/tables/popclockest.txt>.
- [13] "Petroleum." Petroleum.co.uk. Accessed May 10, 2015. <http://www.petroleum.co.uk/>.

# Romantic Unemployment: Neoclassical Economic Theory in Online Dating and Labor Markets

Felix Carreon III

University of Michigan

May 2015

## **Abstract**

Both the online dating and labor markets are matching markets in which firms and agents must agree they're a good fit for one another. As a result, the model of the labor market can be used to illuminate factors in the online dating markets that are well explained by traditional neoclassical economic theory. Some of these characteristics include search theory, preference signaling, and thick markets. A large portion of this paper focuses on preference signaling, the ability of agents to signal to employers strong interest in a position. This model can also be used in the realm of online dating. Professors Soohyung Lee and Muriel Niederle from the University of Maryland and Stanford University, respectively, designed an experiment to determine the effect of virtual roses on accepting proposals from online daters. However, there are factors in both of these markets that run contradictory to neoclassical economic theory. Examples include the paradox of choice made famous by psychologist Barry Schwartz and the presence of asymmetric information, which through cooperative game theory can lead to efficient outcomes in the online dating market.

# 1 Introduction

*“I’m looking for someone to be the Kim to my Kanye, the Angelina to my Brad, the Beyonce to my Jay-Z, you get the picture. All kidding aside, I’m looking for someone who is spontaneous and adventurous. I’m looking for someone who is not afraid to try new things. My ideal match would be someone with a good sense of humor and who could engage in witty banter with me. If you feel that you meet many of these prerequisites please feel free to message me with a freestyle rap or haiku since I hear that’s how kids these days communicate.”*

This excerpt was taken from my profile on Match.com. Online dating has become a popular alternative to the traditional form of dating. Daily we are bombarded with advertisements for online sites such as eHarmony, Christian Mingle, and BlackPeopleMeet.com. Online dating provides an opportunity to sift through potential mates while at home - what could be better? Surprisingly, there are numerous economic concepts present in this complex market. The labor market has many similar facets to the online dating market. As a result, analyzing the labor market can illuminate characteristics of the online dating market well explained by neoclassical economic theory. However, there are instances in both markets that run contradictory to neoclassical economic thought. In these cases, online dating can help shed light on these facets observed in the labor market.

# 2 Background

The field of labor economics seeks to understand the markets for wage labor. This paper will focus on the interaction between workers and employers specifically explaining patterns of employment. The labor market is unique in that there is expected to have a persistent level of unemployment. Similarly, the existence and success of online dating sites is predicated on having consistent pool of available singles. Utility maximization is an important aspect of neoclassical economic theory, which presents a particular assertion of human behavior. These behaviors will then be assessed whether they confine to traditional neoclassical economics and their analysis will provide insight into the online dating and labor markets, respectively.



### 3 Search Theory

*There's no perfect substitute for "the one"*

The opportunity costs associated with searching for a job is high for both the agent and the firm. Additionally, dating is time consuming especially for those engaged in professional careers. The alternative of online dating provides decreased opportunity costs associated with the traditional form of dating. Yet we make these investments of our time because we believe they are worth it. In both of these markets, frictions are present that disrupt the efficiency of the traditional supply and demand model. This is referred to 'frictional unemployment.' But employment may also be the result of failed negotiations between two agents regarding wages. For example, say there's an agent X who has a reservation wage  $w_r$ , the lowest wage the agent is willing to accept. Agent X will be employed if he is offered a wage greater than  $w_r$  and unemployed if he is offered a wage less than  $w_r$ .

There are also frictions present in the online dating market. There isn't a single marketplace where agents are matched together; there are numerous online dating sites with prospective partners. High opportunity costs in the traditional dating market is a primary reason for the popularity of the online dating market. Ironically, the abundance of choice at the simple click of a mouse may lead to high "romantic unemployment." Much like in the labor market, online daters have reservation "utilities" before they can withdraw from the marketplace. In this context, withdrawing from market means no longer being single. In an article entitled *Matching and Sorting in Online Dating* in The American Economic Review, establishes reservation utilities for prospective matches. Let  $v_M(m)$  and  $v_W(w)$  be the reservation utilities for men and women, respectively. The following equations are based on Hiroyuki Adachi's theory on two-sided search (Hitsch 2010).

1.  $A_W(m,w) = [\text{woman } w \text{ accepts man } m] = [U_W(m,w) \geq v_W(w)]$
2.  $A_m(m,w) = [\text{man } m \text{ accepts woman } w] = [U_m(m,w) \geq v_M(m)]$

But what happens to reservation wages when the duration of unemployment increases? We would expect that as unemployment continues that a worker is willing to decrease his or her reser-

vation wage. The worker feels as though his or her current reservation wage may be pricing out prospective employers. A study analyzing the reservation wages of New Jersey workers for duration of 6 months illustrated a decrease of 5% to 14% in reservation wage per week of unemployment (Krueger 2013). An interesting insight discovered by Krueger and Mueller. Workers older than 50 appear less selective in the job market and adjust their reservation wage accordingly. “The apparent willingness of older workers to lower their reservation wage the longer they are unemployed is consistent with the view that job search is an investment, the cost of accepting a lower paying job is less for those who plan to spend less time in the labor market,” argues Kruger (Krueger 2013, pg. 16).

Much like in the labor market, online daters have certain reservation utilities that prospective matches must meet. As the length of being single (romantic unemployment) increases, we would expect that these reservation utilities to decrease similar to the labor market. In online dating, it is difficult to measure the reservation utilities in terms of a single unit such as dollars. As a consequence it is difficult to quantify and therefore test a decrease of reservation utilities as the duration of being single increases.

## 4 Preference Signaling

*Will you accept this rose?*

Economist Michael Spence first developed the idea of preference signaling in the late 1970s, long before the existence of online dating. In his model, the world consisted of skilled and unskilled workers. According to Spence, higher education served only to signal to employers whether workers were skilled or unskilled. In this model, talented workers can validate their skill level by graduating college. Unskilled workers may attempt to receive a college education but the course load will ultimately prove too difficult. This is the key aspect of Spence’s model of signaling, signals become meaning only if they’re costly (Oyer 2013). Receiving a college education is a costly investment. Owner and CEO of a midwestern cement block company echoed these sentiments when he was asked why he hires college graduates. “It doesn’t mean you’re smarter by having a college degree.

It means you've put up with a lot of stuff for four years and you were able to get through it." (Oyer 2013, pg. 104).

Spence could not have foreseen the implications that the Internet would have on his model of signaling. Now it is easy to submit many job applications with a click of a button. Similarly, it is just as easy to message numerous potential partners through online dating websites. According to an article in *The Economic Journal*, approximately 86% of those unemployed with home internet access used the web in one aspect of their job search (Kuhn et al 2013, pg. 1216). Comparatively, only 9% of the adult population has used an online dating site (Smith et al 2013, pg. 2). Of the entire online dating population, 66% have gone on at least one date with someone they encountered through this medium. Additionally, 23% of online daters report entering a long-term relationship or even meeting their spouse. While technology has increased the scope of both potential job opportunities and dating prospects, it also has created some challenges. Most notably, it is difficult to determine the level of interest in a particular person or job if the opportunity costs of completing a job application or messaging a prospective partner are quite low. As a consequence, Spence's model of signaling has become increasingly important in these matching markets.

The signaling model can be written more formally. Imagine a scenario where there are two firms  $f_1$  and  $f_2$  that are competing over two workers  $w_1$  and  $w_2$ . In this market, matches between groups yield a payoff of 1 while unmatched groups yield a payoff of 0 (Coles et al 2012). In addition, the probability that firm strictly prefers  $w_1$  to  $w_2$  is 50 percent and vice versa. In this game, workers are not aware of each other's preferences. If a firm receives a signal from its top worker, then that firm will certainly send a job offer to that worker knowing it will be accepted. If a firm receives no signal, then it will still send a job offer to its top worker not knowing if it will be accepted. What if the firm receives a signal from the second worker but not its top worker? If it offers the second worker a job offer then the firm is "responding" to the signal and if it chooses to still offer the job to its top worker then it is choosing to "ignore" the signal.

If each worker can only send one signal to one firm and each worker in the market uses this signaling mechanism then it is inferred that if one firm receives just one signal the other firm received a signal from one agent in the market. In the case of  $f_1$  and  $f_2$ , let's say that  $f_1$  prefers

$w_1$  to  $w_2$ . In the scenario in which  $w_2$  signals to  $f_1$ , then if  $f_1$  sends an offer to  $w_2$ , it will then receive a payoff of  $X$  (Coles et al 2012). In the case that  $f_2$  ignores the signal from  $w_1$ , there's a 50 percent probability that it was the firm's preference with. This would result in payoff of  $X$  for  $f_2$  and a payoff of 0 for  $f_1$ . Table 1 illustrates the payoff matrix for  $f_1$  and  $f_2$ . If  $X$  is greater than 0.5 then both firms respond to signals.

**Table 1: Firm  $f_1$ 's payoffs condition on receiving a signal from it second ranked**

$f_1/f_2$	worker <sup>1</sup>	
	Respond	Ignore
Respond	X	X
Ignore	0	1/2

Has the creation of the Internet led to inefficient outcomes in both of these markets? While there has been a substantial increase in the level of competition in both the job and online dating market, people with strong qualifications and interest in a particular job stand a far better chance than those who are not qualified and have little interest. But is there a way to signal strong interest in a particular opportunity or prospective partner? On Match.com, users are able to signal strong interest in a person by sending out VIP messages. Users are only able to send a VIP message once a week, thus illustrating a strong investment in that person. Other online dating sites allow users to purchase a virtual rose for some arbitrary amount of money.

Professors Soohyung Lee and Muriel Niederle from the University of Maryland and Stanford University, respectively, wanted to determine the effect of virtual roses on accepting proposals from prospective online daters. (Lee et al 2011). They conducted an experiment consisting of 613 participants, approximately of which 50% were female. Each participant was given 10 electronic messages, a proposal, to give out to potential partners. In addition, each participant was given two "virtual roses" but randomly selected 20% of participants were given eight. A virtual rose was a way to signal special interest in a person and was valued because of their limited allocation.

Through this experiment, Lee et al wanted to determine the effect the virtual roses had on accepting proposals. In order to do this, they ran multiple regressions on the sample set of partici-

<sup>1</sup>Coles et al. June 2012. Preference Signaling in Matching Markets. Pg. 7.

pants accounting for different factors. The only variable in this analysis was whether a virtual rose was attached to a proposal and the senders' desirability rating. The experiment consisted of two stages: a proposal stage and a response stage. During the response stage, it is important to note that participants did not know whether proposals were accepted or rejected. Lee et al discovered that a rose improved the chances of acceptance by 37% and that having eight roses improved the chances of a date for men by more 60% (Lee et al 2011, pg. 31).

## Acceptance and Receiving a Rose<sup>2</sup>

**TABLE XI ACCEPTANCE AND RECEIVING A ROSE**

	All (1)	Men (2)	Women (3)
Receive at least one rose	0.259* (0.139)	0.484** (0.237)	0.087 (0.156)
Female	-0.335** (0.133)		
Constant	0.488*** (0.097)	0.432*** (0.117)	0.231** (0.106)
No. obs	97	49	48
R-sq	0.080	0.081	0.007

Notes: OLS estimates. The dependent variable is the number of proposals that a recipient accepted. Standard errors are in parentheses. \*, \*\*, and \*\*\* indicate significance at 10, 5 and 1 percent, respectively.

Not surprisingly most of the 1,921 proposals were sent by men, compromising 66% of the sample (Lee et al 2011, pg. 11). Additionally, Lee et al studied the effect roses had on the acceptance of proposals from different desirability categories. A priori, we would expect the effect of roses to be greater for those in the middle or bottom desirability categories because these groups lack the confidence of top tier prospects. As a result, a rose is a gesture of reassurance for the recipient especially those in bottom desirability categories. Without attaching a rose to a proposal, an offer from a person in the top desirability index is approximately 18 percentage points more likely to be accepted than an offer from someone in the bottom desirability group (Lee et al 2011, pg. 21). The addition of a rose increases acceptance among all groups by 3.3 percentage points.

## Effect of Roses Across Desirability Categories<sup>3</sup>

TABLE VII EFFECT OF ROSES

Model	FE-R (1)	FE-R (2)	FE-R (3)	FE-R (4)	OLS (5)	FE Logit (6)	FE-R (7)
Recipients	All	Men	Women	Active	All	All	2 roses
<b>Model A</b>							
Rose	0.033** (0.016)	0.053 (0.033)	0.026 (0.018)	0.054** (0.025)	0.030* (0.018)	0.443** (0.201)	0.034* (0.018)
S_Middle	0.048** (0.019)	0.066* (0.037)	0.031 (0.022)	0.079** (0.031)	0.074*** (0.020)	0.811*** (0.298)	0.052** (0.022)
S_Top	0.178*** (0.020)	0.177*** (0.040)	0.175*** (0.023)	0.293*** (0.033)	0.191*** (0.021)	2.284*** (0.312)	0.181*** (0.023)
R-sq (log Lik.)	0.50	0.54	0.46	0.46	0.13	-242.37	0.49
<b>Model B</b>							
R_Bottom Rose	0.054 (0.047)	0.273 (0.218)	0.038 (0.046)	0.087 (0.071)	0.003 (0.047)	0.935 (0.598)	0.071 (0.052)
R_Middle Rose	0.078*** (0.027)	0.103 (0.076)	0.073*** (0.027)	0.097** (0.040)	0.082*** (0.029)	0.677** (0.317)	0.068** (0.031)
R_Top Rose	-0.001 (0.021)	0.035 (0.037)	-0.023 (0.027)	0.013 (0.035)	0.003 (0.024)	0.131 (0.292)	0.006 (0.024)
S_Middle	0.047** (0.019)	0.066* (0.037)	0.029 (0.022)	0.079** (0.031)	0.072*** (0.020)	0.815*** (0.299)	0.051** (0.022)
S_Top	0.176*** (0.020)	0.175*** (0.040)	0.174*** (0.023)	0.290*** (0.033)	0.189*** (0.021)	2.283*** (0.314)	0.180*** (0.023)
No. of proposals	1,902	657	1,245	1,153	1,902	796	1,516
No. of recipients	393	168	225	226	393	103	310
R-sq (log Lik.)	0.5	0.54	0.46	0.46	0.13	-241.18	0.49

Unfortunately virtual roses are not available in the labor market but preference signaling is still practiced. Because technology had made it significantly easier to apply to numerous jobs at low costs, competition is high for a limited number of positions. There are even formal structures in place that allow prospective employees to signal their interest in a company. The American Economic Association (AEA) operates a service in which economics graduate students can signal their interest in an interview with two firms at the annual Allied Social Science Associations (ASSA) meeting. Much like Lee and Nierdle's research, signaling may be necessary for top tier job candidates to demonstrate genuine interest to middle desirable firms. This may be the case for overqualified candidates who have geographic preferences and interest in second-tier employers. The preference

system was created for just this reason. The AEA recommends job applicants use their signals on employers “you like but that might otherwise doubt whether they are likely to be able to hire you” (Oyer 2013, pg. 116). This signaling mechanism increased chances of getting an interview from 25% to 32% and the effect was prominent for candidates for small liberal institutions compared to Ivy League teaching positions.

At the University, Ross School of Business students are able to bid for interviews for both internships and full-time positions. The system is not limited to just undergraduates but also includes MBA students as well. In order to obtain an interview with an employer, each Ross student must submit their resume and cover letter in iMPact, the school’s exclusive career portal. Each Ross student is allocated with 1,000 points, which they can use to bid for interviews. If a candidate is selected for an interview after a review of their information in iMPact then they will be charge 50 points if they chose to accept an interview invitation. Those who are not selected for an interview invitation can bid for an interview. In this scenario, if students who bid are successful then they’re charged a Market Clearing Price and placed on the interview schedule. If unsuccessful candidates are then placed on a waiting list, which is ranked by bid amount. Bid points are not deducted until the candidate has chosen to accept an invitation for an interview. In the appendix is a flowchart outlining this process.

The bidding system at the Ross School of Business is a prime example of signaling in the labor market. If a student is really interested in a prospective employer and is not selected for an interview after the initial round then the candidate will bid a significant portion of their points in order to secure an interview. However, an invitation for an interview does not guarantee being offered the position. The candidate must assess when and where to use the allocation of points to best maximize utility.

Within these markets, employers and potential dating partners are unable to sift through all potential candidates. The creation of the Internet has made search easier but as a consequence opportunity costs have decreased substantially. Michael Spence’s model of signaling becomes more important as a result of online search mechanisms. Through applications, resumes, cover letters and other forms of communication, candidates can signal their preferences to employers and potential

dating partners. Spence's framework asserts that signals only become meaningful if they're costly. Virtual roses are effective because online daters are unable to give them to all suitors, making them costly to give one prospect and not another. Employers and potential dating partners then respond to signals to make "hiring" decisions.

## 5 Thick vs. Niche Markets

### *Size Matters*

Neoclassical economic theory suggests that thick labor markets are preferable to thin markets. According to the barter model, "the probability of finding a trading partner depends on the number of potential partners available, so that an increase in the thickness of the market makes trade easier" (Moretti 2010, pg. 1286). As a result, metropolitan areas are attractive to both firms and agents because it's likely to lead to better matching between the two. This is known as the thick market effect where the more options in a market, the greater probability a buyer or seller is to find a good match. Economists Hoyt Bleakley and Jeffrey Lin illustrate that thick markets offer more job mobility for agents (Bleakley et al 2012). They also discovered that this greater job mobility leads to happier workers. Younger workers are more likely to shop around for the "best fit" because of the greater number of opportunities in thick markets. However, as agents progressed through their professional careers they are less likely to switch jobs because it becomes more costly to do so as a result of presumably better matching in thick markets. Additionally, these occupation changes occur more often in large, diverse local markets compared to small, specialized ones (Wheeler 2008).

A secondary benefit of thick markets is that it mitigates risk of idiosyncratic shocks to both workers and firms (Moretti 2010, pg. 1287). The numerous opportunities available in thick markets leave agents better off in terms of finding employment in the event of a negative demand shock. Moreover, firms benefit from having a large pool of candidates ensuring vacancies are filled.

But do thick markets also lead to better matching in the world of online dating? An online dating site would be of little use if only one person participated. Sites such as match.com and eHarmony benefit from network externalities, the value of using such sites is dependent on the



number of others using it similar to the proliferation of cellular telephones. The available pool of potential matches on these sites makes it more likely to find a suitable partner. After sorting through members that meet the agent's set of preferences, the pool of candidates becomes much smaller. As a result, starting with a larger supply of prospects is to the benefit of the consumer.

ChristianMingle, JDate, FarmersOnly, and BlackPeopleMeet are all niche markets. They consist of a single characteristic that defines that market. If a larger pool of candidates leads to better matching then why are these sites successful? These are sites successful because all candidates within these niche sites share the characteristic of the site that is usually of prominence for these candidates. Segregation through a specific religious affiliation is common for many couples so it comes to no surprise that many of these sites are characterized by certain religious beliefs. "People are joining ChristianMingle because that's the driving force in their life, and when you get there, you're already among those who share those same values," says ChristianMingle spokeswoman Ashley Reccord (Hill 2013).

But even for these niche characteristics, online daters benefit from using thick markets such as match.com and eHarmony. "I also pretended to be a vegetarian (I am not, but I'm willing to imagine for a few minutes? I would draw the line at pretending to be vegan). There were twenty-two women between forty and fifty living within twenty miles of me on VeggieDate.com" (Oyer 2013, pg. 173). Specialist dating sites may increase overall consumer welfare but size matters and a larger pool of candidates is better *ceteris paribus*. If vegetarians are only willing to date other vegetarians then the pool of possible matches becomes increasingly small. These agents stand to benefit from residing in a metropolitan area where it's more likely to find other vegetarians. However, vegetarians living in rural areas may have to compromise on this and other characteristics in prospective partners.

## 6 Paradox of Choice

### *More is Less*

While scrolling through prospective matches on Match.com, members can refine prospects through numerous characteristics such as appearance, interests, background, and lifestyle. Sites

such as Match.com and eHarmony.com offer an abundance of choice for consumers. Neo-classical economic theory would suggest that the increase in available options would result in greater consumer utility. However, psychologist Barry Schwartz argues more choice does not increase consumer welfare but decreases it. Schwartz, author of *The Paradox of Choice: Why More is Less*, believes that more choice sometimes leads to an inability to make a choice at all. “I feel stymied by the sheer numbers of men who are dating online now,” says New Yorker Ali, 32 (Singleton 2011). “It’s intimidating to figure out who I like, but even harder to consider that, for the average dude checking out online profiles, I’m just one of a million girls in the big city.”

Schwartz highlights that when one does make a choice, the abundance of choice often leaves us dissatisfied. This dissatisfaction is result of opportunity costs, the value of the next best alternative. “All these choices make it impossible not to constantly wonder if there’s someone better out there,” says Jenny, 28, a Los Angeles native (Singleton 2011). Jenny echoes the sentiments of many online daters; with so much choice we should be able to find the “perfect” match. Especially in western culture where the institution of marriage plays an important role in shaping our satisfaction within our lives. Social comparison theory suggests that this dissatisfaction may be a result of comparing ourselves with friends who not only are married but also have children. As a consequence, there is added social pressure in the realm of online dating.

A similar phenomenon occurs in the labor market. The presence of so much choice often leaves many employees wondering if they could land a better job with a better compensation package. Even those who are employed measure themselves amongst friends and neighbors as a tool of self-evaluation. While there are numerous opportunities in the labor market, these choices are limited by qualifications, education level, and skill set. This is one area in which the labor market differs from online dating. Employers seek the best candidates that make an immediate contribution to their organization. On the other hand, online daters have a set of preferences but decision-making is rarely dictated by economic theory.

## 7 Asymmetric Information

*“This is like writing a resume and trying to be honest..but not TOO honest :)...”<sup>2</sup>*

In both the online dating and labor markets, both suffer from the problem of asymmetric information. As a result, it’s difficult to mitigate risk in both markets. Especially in the world of online dating, it is difficult to verify if a person’s profile is completely truthful. As a consequence, these information asymmetries lead to inefficient outcomes. A potential partner may spend a great deal of time focusing on someone who may be a great fit but later realizes when physical interaction is involved that the person lied to increase their prospects. Because search is costly, this leads to a market failure. In the labor market, the same is also true. A potential job candidate may lie on their resume to enhance their chances of landing a job. If a firm does not due their due diligence, then the problem of asymmetric information becomes very costly. The problem of asymmetric information poses a challenge to finding the best match for both parties.

If asymmetric information is so costly, what can be done to resolve the problem? Some online dating sites go to extreme measures to validate the credentials of its members. They require potential members to submit government documents to verify their identity and age. Moreover, they require proof of employment to validate one’s personal income. Some even go as far as to take physical measurements of candidates to verify personal information that is often exaggerated on these sites.

What is the overall effect of asymmetric information in totality of these markets? Prospective employers and matches become increasingly suspicious of candidates’ qualifications. This is extremely prevalent in online dating. On average, members of the online dating community lie regarding their height, weight, age and income. It is typical for male members to increase their actual height by one to two inches and women to underreport their weight (Oyer 2013, pg. 46). As potential partners get older, age becomes a more significant characteristic in a match. As a result, many older members underreport their age by two years.

---

<sup>2</sup>Oyer, Paul. December 2013. A woman’s “self-summary on OkCupid. Everything I Ever Needed to Know About Economic I Learned from Online Dating. Harvard Business Review Press. pg 42.

The lack of perfect information in the online dating community hurts all members of the market. A member will wonder if a potential match has a body type that is “about average” or “a few extra pounds.” Members will also suspect other information of a candidate’s profile. They will assume that most members are over reporting their information to appear more attractive to potential partners. A male member will assume that a potential match who lists herself as “average” body type is actually “a few extra pounds” unless the member posts numerous photos illustrating otherwise. This lack of perfect information will lead to fewer matches between “good” matches especially if the agent deems one of these characteristics “non-negotiable”.

Much like “profile-inflation” in online dating, resumes are often padded to bolster qualifications of a candidate. Checks in this area make it difficult to tell blatant lies and thus these markets function more efficiently than those found in online dating. References can be used to verify previous employment. In addition, many employers require transcripts to demonstrate proof of education. A recent study found that economists are known to exaggerate information found on their resumes and as a result hiring managers discount such information accordingly (Synder 2011). “When people know that the information they provide will be compared with information gathered elsewhere, they have more incentive to tell the truth, and others are thus more likely to believe them,” states Oyer (Oyer 2013, pg. 53).

Game theory is a field of economics that focuses on strategic decision-making, which usually involves a winner and a loser between two agents. Specifically cooperative game theory involves decision-making when both agents’ interests are perfectly aligned. In the case of online dating, agents are looking for long-term partners. The success of cooperative game theory is dependent on perfect information between the two agents. Rarely does this occur in the world of online dating. As a consequence, sometimes it is better to lie regarding minor details on your profile. Initially, if a prospective partner discovered one of these details then there never would have been an opportunity for interaction between the two agents. Once there is interaction, the one agent now finds after spending some time with the prospective partner that detail is not a significant part of her preferences in her utility function. As a result, the relationship never would have formed under perfect information between the two agents.

However, there are instances where it's best to be honest. For example, if an agent has children from a previous relationship then it is best to be forthcoming about this information to prospective matches. "I can't be in a relationship with someone who will not feel comfortable with my children (they're with me half the time) or my dog (who is always with me)" (Oyer 2013, pg. 51). Lying, hiding, or misrepresenting this information would run against interests of prospective partners. In cooperative game theory, the more the interests of agents are aligned then lying becomes less of a concern.

This version of cooperative game theory runs contradictory to ideas of corruption and bad faith. Activities of corruption and bad faith not only lead to bad outcomes for US citizens but also played a significant part in each of the past three economic contractions. During the recession from 2007 - 2009, unemployment peaked at 10% in October 2009 (Bureau of Labor Statistics 2012). In December 2007, the national unemployment rate was just 5% and remained near or at the figure for the previous 30 months. While the scale of activities of corruption and bad faith in online dating are nowhere near to bring about a contraction in the US economy, it is interesting to note that these activities do result in efficient outcomes and increase the welfare of consumers.

## 8 Conclusion

Economists have studied no market as extensively than the labor market. Many behaviors exhibited in this market are dictated by traditional neoclassical economics specifically utility maximization. What makes the labor market unique is the market is defined by two-sided search; both employers and prospective employees must agree that they're a good fit for one another. Similarly, the on-line dating market shares this structure. As a result, the model of the labor market can highlight numerous characteristics in online dating that are well explained by neoclassical economic theory. However, they are instances in both markets that run contradictory to neo-classical economic thought. In these cases, online dating can help shed light on these characteristics observed in the labor market.

## References

- [1] Gunter J. Hitsch, Ali Hortacsu and Dan Ariely. March 2010. “Matching and Sorting in Online Dating”. *The American Economic Review*.
- [2] Alan B. Krueger and Andreas I. Mueller. November 2013. “A Contribution to the Empirics of Reservations Wages”.
- [3] Soohyung Lee and Muriel Niederle. October 2011. “Propose with a Rose? Signaling in Internet Dating Markets”.
- [4] Peter Coles, Alexey Kunshnir, and Muriel Niederle. June 2012. Preference Signaling in Matching Markets.
- [5] Hill, Logan. February 2013. At ChristianMingle and JDate, God’s Your Wingman.  
<http://www.bloomberg.com/bw/articles/2013-02-28/at-christianmingle-and-jdate-gods-your-wingman#p1> Accessed February 2015.
- [6] Hoyt Bleakley and Jeffrey Lin. 2012. Thick-Market Effects and Churning in the Labor Market: Evidence from US Cities.
- [7] Oyer, Paul. December 2013. *Everything I Ever Needed to Know About Economics I Learned from Online Dating*. Harvard Business Review Press.
- [8] Singleton, David. 2011. The Paradox of Dating Choice. Match.com  
<http://www.match.com/magazine/article/12175/The-Paradox-Of-Dating-Choice/> Accessed April, 2015.
- [9] Peter Kuhn and Hani Mansour December 2013. Is Internet Job Search Still Ineffective? *The Economic Journal*. Royal Economic Society. pg. 1213 - 1233.
- [10] Aaron Smith and Maeve Duggan October 2013. Online Dating & Relationships. Pew Research Center.

- [11] Wheeler, Christopher, 2008. Local market scale and the pattern of job changes among young men. Reserve Bank of St. Louis Working Paper 2005-033C. Revised.
- [12] US Bureau of Labor Statistics, February 2012. The Recession of 2007 - 2009.  
[http://www.bls.gov/spotlight/2012/recession/pdf/recession\\_bls\\_spotlight.pdf](http://www.bls.gov/spotlight/2012/recession/pdf/recession_bls_spotlight.pdf) Accessed April 2015.
- [13] Christopher Snyder and Owen Zidar, 2011. Resume Padding by Economists, working paper. Dartmouth College.
- [14] Moretti, Enrico, 2010. Local Labor Markets. *Handbook of Labor Economics*, Volume 4b.

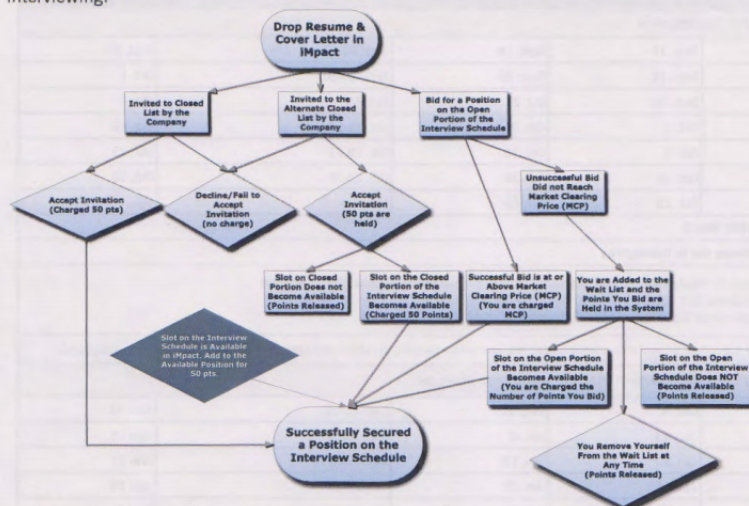
## Appendix

### **MICHIGAN** ROSS SCHOOL OF BUSINESS

## Career Services

### Bidding Overview – On Campus Interviewing: From Resume Drop to Interview Schedule

This flowchart shows the possible ways to be placed onto an interview schedule during on-campus interviewing.



- Start by submitting your resume and cover letter for a particular job. You will be either placed on the closed list, alternate closed list, or you will need to bid.
- If you are invited to the closed list, you can accept the invitation and be charged 50 bid points, which will place you on the interview schedule. You can decline the invitation for no charge.
- If you are invited to the alternate closed list, you can accept it for no charge. If a spot on the interview schedule opens, you will be placed on the interview schedule and will be charged 50 points. You will not know your rank on the alternate closed list.
- If you did not receive a closed list invitation, you may bid for a spot on the interview schedule.
- If you bid and are successful, you are charged Market Clearing Price and placed on the interview schedule. Remember that MCP is the lowest successful bid to get onto the interview schedule.
- If you bid and are not successful, you are then placed on the bidding wait list, which is ranked by the amount each student bid for that job. If a bidding interview slot comes open, the first wait listed student will be promoted onto the schedule and charged his or her bid price, not the Market Clearing Price.
- Your bid points are held while you are on the wait list. You may remove yourself from the wait list at any time without penalty and your bid points will be refunded immediately. If you are not promoted to the interview schedule, your bid points will be returned.
- If you run out of bid points, and are offered a closed list spot, RCS will accept the closed list on your behalf.

Career Services, Stephen M. Ross School of Business, University of Michigan, 2014



# Contraception Employed: Using Economic Models to Predict the Effect of Employment on Condom Usage in Brazil

Seth Merkin Morokoff

Princeton University

May 2015

## **Abstract**

Two separate theoretical frameworks suggest that employment may have a negative directional effect on the probability of condom usage in the developing world. However, this prediction initially seems counterintuitive. Due to the dearth of literature predicting condom usage in Brazil, despite its high incidence of HIV infection, I examine the estimated effects of various measures of employment on condom usage at last intercourse for both men and women. Using data from the 1996 Demographic and Health Survey, I estimate the effects of current employment and year-round employment on the probability of men using a condom at last intercourse as well as the effects of current employment, earning income through formal or informal work, and working outside of the home on the likelihood of women using a condom at last intercourse. I find that year-round employment decreases the likelihood of men using a condom at last intercourse by 31%. Further, working outside of the home increases the probability of women using a condom at last intercourse by 22%.

# 1 Introduction

The Brazilian National AIDS Program has been widely lauded as the leading example of an institutionalized strategy against HIV/AIDS in a middle-income country by various news outlets, non-governmental organizations, policymakers, and academics. Much of the praise focuses on the country's landmark 1996 decision to provide free universal antiretroviral therapy to any HIV-positive resident. However, prevention of HIV transmission through improved sexual education and promotion of safer sexual health behaviors that limit risk exposure (i.e., the proliferation of condom use) still plays a critical—if less exciting—role. Okie (2006) explains that widespread condom usage in a population not only cuts treatment costs for the government, but mathematical models predict that delivering large-scale antiretroviral treatment without effective prevention measures would worsen the AIDS epidemic in many countries due to the existence of healthy vectors in the population who can live normally with treatment but continue to transmit the virus.

Despite the successful proliferation of condom usage among the Brazilian population, with the rate of individuals reporting condom use at first sexual intercourse rising from 4% in 1986 to 48% in 1999 (Brazilian Ministry of Health, 2003), few studies focus on what factors predict condom use among a nationally representative sample. Two separate investigators who conducted recent studies on condom usage in Brazil cite a lack of national studies and a lack of attention paid to male sexual behavior in the literature focused on Brazilian sexual behavior (Calazans *et al*, 2005 and Merchan-Haman *et al*, 2002). In those studies that do exist, income and education are perhaps the strongest and most consistent correlates of condom usage. However, little analysis of the effect of employment on condom use has been conducted despite two prominent theoretical frameworks suggesting employment may have a negative directional effect on the probability of condom usage in the developing world—a prediction that feels counterintuitive and seems problematic for continued development.

The desired fertility model presented by Pritchett (1994) suggests that individuals in the developing world act rationally to achieve fertility targets based upon the utility they derive from having children and monetary or temporal constraints limiting the number of children they can

afford. In this model, stable employment would decrease rates of condom usage due to reduced fears of uncertainty regarding future employment increasing desired fertility or relaxing the budget constraint placed upon fertility and increasing the number of children affordable to the individual. In contrast, the family planning gap theory argued by Bongaarts and Sinding (2009) suggests many individuals in the developing world lack access to contraception. Therefore, people with greater resources necessary to obtain contraception including money, transportation, and time are more likely to use contraception. After controlling for income in this model, full-time employment would also decrease the rate of condom usage due to the workday diminishing the amount of time individuals have to procure contraception. I test these theories by estimating the effects of various measures of employment on the likelihood of condom use at last intercourse.

Using nationally representative data from the 1996 Demographic and Health Survey, I create probit models to analyze a sample of sexually active men and women of reproductive age who are not otherwise impeded from conceiving a child through sterilization; infecundity; or, in the case of women, current pregnancy. Reproductive age for women is defined as age 15 through 45 based on the background literature. No such age restriction exists for the male sample, which includes individuals between age 15 and 59. Individuals in the analysis sample are also not currently attending school. A full derivation of the analysis sample is presented in the data section of the paper.

I find that two measures of employment are consistently significant across all iterations of the model: year round employment for men and working outside the home for women. As predicted by the economic theory, year round employment has a negative association with condom use at last intercourse and decreases the likelihood of a man using a condom by 31%. However, working outside the home has a positive association with condom usage at last intercourse, increasing the probability of a woman using a condom by 22%. Finally, I check the robustness of the employment coefficients by controlling for a separate mechanism in which risky sexual behavior prompts individuals to use condoms with greater frequency.

## 2 Background

### 2.1 Specifics of the Brazilian Context

Several factors make Brazil a unique context to examine condom usage and sexual health behavior. The three primary factors repeated in the literature are: (1) the decentralized nature of the HIV/AIDS program which focuses on community-based strategies rather than national federal policy, (2) the legal status of prostitution and sex work which increases the number of members in at-risk populations but also increases their visibility, and (3) the overwhelming majority of the population that identifies as Roman Catholic - a religion that officially condemned condom usage during the period of when the DHS data were collected. A final consideration is the availability of condoms to Brazilian consumers and their preferences regarding alternate forms of contraception.

Gómez (2010) writes that the Brazilian Ministry of Health began to promote decentralization within its National AIDS Program in the 1990s. In 1994 under the administration of a new president focused on building Brazil's international reputation by successfully responding to the AIDS crisis, the country began receiving loans from the World Bank to execute HIV prevention policy. The federal government distributed that money to states which implemented regional policies. However, aware of the fact that a number of states had proven incapable of rendering effective HIV prevention policy, the Ministry of Health also pursued formal partnerships with NGOs to supplement funds and revamp strategy in the regions making the least progress. Further, Levi and Vitória (2002) write that certain policies focus exclusively on specific municipalities and cities. For instance, one major use of funding in Rio de Janeiro—the second largest city in Brazil—is free condom distribution and HIV prevention campaigns during Carnival. Therefore, while models of condom usage in Brazil including state indicators as part of a geographic fixed effects model to avoid problems of endogeneity through controlling for omitted factors specific to certain areas of Brazil ought to be considered, state indicators may over-control for condom availability since HIV-prevention policy is determined at the state level. The majority of the economic literature related to Brazilian condom usage fails to test models with state indicators because most of the studies only analyze data from one city.

A second consideration comes with the legal status of prostitution in Brazil. Okie (2006) writes that by 2001, due to targeted efforts by the National AIDS Program, 74 percent of Brazilian sex workers reported consistent use of condoms with clients. The prevalence of HIV infection among female sex workers has remained low and stable for an at-risk population. According to a 2005 report by the Brazilian Ministry of Health, about six percent of female sex workers are infected with HIV. Evidence from studies of health economics indicates that Brazilian sex workers are rational economic actors. Miranda *et al* (2011) find commercial sex workers were 9.01 times more likely to have used a condom at last sexual intercourse than other women in a logistic regression model of women under 30 years old in Vitória, Brazil. Shah (2013) finds that male sex workers within developing regions of Latin America demonstrate rational economic behavior by charging a compensating differential for disease risk for unprotected sex in areas with high STI prevalence. Certain states within Brazil have notably higher rates of HIV prevalence than others. Specifically, states in the Northern, Northeastern, and Southern regions of Brazil have consistently exhibited higher rates of HIV than states in other regions, and those in the Southeastern region exhibit much lower rates of HIV prevalence (United Nations General Assembly Progress Report, 2012). Therefore, clear regional differences in STI prevalence present the possibility to control for endogeneity with state indicators, although the potential for state indicators to over-control remains present.

The third consideration related to condom usage in Brazil is the strength of the Roman Catholic Church in society. 74% of Brazilians identify as Roman Catholic, in a 2007 census conducted by the Brazilian Institute of Geography and Statistics. Martine (1996) provides a history of the key factors in Brazil's declining fertility. He writes in the mid-1980s the Catholic Church began to support natural methods of contraception, which steered women towards sterilization and illegal abortions. However, liberal factions of practitioners of Catholicism began to clash with the traditional leadership by counseling community members to use methods like the birth control pill and condoms. In an overview of Brazilian HIV policy, Levi and Vitória (2002) find that during the 1990s religious associations began to partner with community groups and NGOs to form civil society organizations that focused on the prevention of HIV, thereby partially removing the stigma associated with condom usage within the Catholic community of Brazil. Despite the official stance

of the Church against condoms, Calazans *et al* (2005), Gupta (2000), Juarez and LeGrand (2005), and Juarez and Martin (2006) all estimate being Catholic has a positive effect on condom usage either at last sexual intercourse or first sexual intercourse, depending on the paper.

The final consideration concerns the availability of condoms to Brazilian consumers and their preferences regarding alternate forms of contraception. Condoms carried significant monetary cost for those in low income brackets in Brazil relative to necessary goods like food and other forms of contraception like sterilization. As mentioned, Martine (1996) finds that female sterilization was the most widely used method of birth control through the mid-1980s with women often consenting to sterilization to reduce their lifetime fertility. Sterilization was also practiced involuntarily as part of caesarian section deliveries on women from favela slum areas who had already birthed three children. However, beginning in the 1970s, fertility rates were already in rapid decline, and contraceptives—notably the birth control pill—were widely available for those who could afford them. Gómez (2010) argues that condom usage would not proliferate among the general population until 1994, when a new Brazilian president promoted HIV prevention policy. Statistics comparing studies by the Brazilian Ministry of Health over time show that while only four percent of the Brazilian population had used a condom during their first sexual encounter in 1986, that rate had increased to 48% in 1999. Martine (1998) adds a description of the three varieties of condom available in Brazil during the 1990s: imported condoms, Jontex condoms produced in Brazil under a license from Johnson and Johnson, and local brands of condoms. Imported condoms did not represent a significant portion of the market, and were sold only in specialty shops. Jontex condoms were widely available in supermarkets and pharmacies, but cost more than \$1 USD apiece—a price that was prohibitively expensive for those Brazilians living in favela slums. In neighborhood pharmacies, the local brands were available within the price range of the favelados. However, quality control was inconsistent, and the condoms were prone to breaking. Therefore, the proliferation of free government-sponsored health clinics during the 1990s, which supplied Jontex condoms as part of their HIV prevention strategy, greatly increased condom usage, especially in the lower socio-economic classes. Juarez and Martin (2006) note that even in their study of low-income adolescent males from slum areas of Brazil using data from 2000, less than one percent of respondents described

condoms as prohibitively expensive or unavailable, meaning condoms were widely available to those who wanted them during the period.

## 2.2 Review of Past Literature Modeling Condom Usage in Brazil

Calazans *et al* (2005), Merchan-Hamann *et al* (2002), and Miranda *et al* (2011) are the only three studies out of the seven reviewed that use current employment of the individual survey respondent as a covariate in their models. Calazans *et al.* (2005) find a statistically significant positive association between having never worked and lack of condom usage at last intercourse when using current employment as the reference category; this corresponds to a negative association between never having worked and condom use at last intercourse—the dependent variable most commonly used in the literature. Merchan-Hamann *et al* (2002) find current employment has no effect on the dependent variable “risky sexual behavior,” which the authors define as less than a frequent level of condom usage (frequency of condom use is self-reported in their survey as always, frequent, occasional, or rarely/never). Miranda *et al* (2011) initially find a slight negative association between formal employment (defined as either part-time or full-time income-generating work) and condom use at last intercourse using not being employed as the reference; however, the variable is not statistically significant, and the investigators drop it from subsequent models.

The conflicting direction of the results acts as motivation to further investigate the effect of employment on condom use at last intercourse. Further, each of the studies only includes current employment as a covariate, whereas other aspects of employment such as working full-time versus part-time or working outside of the home potentially affect condom use at last intercourse; more variables of interest ought to be explored under the umbrella of employment. Finally, both Calazans *et al* (2005) and Miranda *et al* (2011) have flaws in their implementation of using current employment as a covariate. Calazans *et al* (2005) only includes current employment when estimating a subgroup model for individuals with steady partners. The authors choose to exclude current employment from the model of individuals with casual partners without explanation. Miranda *et al* (2011) drop employment as a covariate from the final model as part of their variable elimination scheme, thus obscuring its estimated effect on condom use at last intercourse after controlling for

other covariates. I have not found any previous studies that correct for these issues. I address this gap in the literature by estimating the effects of five measures of employment on condom usage at last intercourse, and including these employment statistics as the variables of interest in each model variation I test.

Two other gaps in the literature mentioned previously stem from the limited analysis samples used in prior studies. The most obvious gap is the lack of studies using nationally representative data. Out of the seven studies reviewed, only Calazans *et al* (2005) analyze a nationally representative sample, and the study does not include a model with state indicators, which as mentioned previously, could limit bias of the estimated effects in the model or could over-control for the availability of condoms. Gupta (2000) uses data from the 1996 Demographic and Health Survey that I use as well; however, she subsets the data to target only adolescent women in nine states of the northeastern region of Brazil as her population of interest. Juarez and LeGrand (2005), Juarez and Martin (2006), Merchan-Hamann *et al* (2002), Miranda *et al* (2011), and Silveira *et al* (2005) all use data collected from a single city. Therefore, to the best of my knowledge, no study of condom usage in Brazil attempts to control for endogeneity from omitted variables with geographic fixed effects. I address this gap in the literature by incorporating state indicators into two iterations of my model, remaining cognizant of the potential issues.

The second gap arises due to the dearth of literature regarding male sexual behaviors in Brazil. Although four of the seven studies reviewed here include men in their sample, most target a narrow subgroup of men that fails to represent the typical Brazilian population. Further, three of the studies target groups that are unlikely to be employed, and therefore cannot examine the effect of my variable of interest: employment. Juarez and LeGrand (2005) and Juarez and Martin (2006) both examine condom use among males under age 20 in favela slums outside of the city Recife, a group unlikely to hold formal employment of any kind. Merchan-Hamann *et al.* (2002) analyze a sample of male and female students currently enrolled in secondary school, but current students are less likely to simultaneously hold employment. Calazans *et al* (2005) uses a nationally representative sample of men and women, ages 15 through 24. Only Calazans *et al* (2005) and Merchan-Hamann *et al* (2002) compare the likelihood of men reporting condom use at last sexual intercourse to the



likelihood of women reporting condom use at last sexual intercourse. Calazans et al. (2005) finds being male is positively associated with condom use at last intercourse. Merchan-Hamann *et al* (2002) finds being male is a positive correlate of risky sexual behavior—meaning men are less likely to report using condoms frequently than women. I address this gap in the literature by using a nationally representative sample of men and women after excluding groups that are unlikely to be employed like students to obtain a cleaner sample.

In general, the studies reviewed here use multivariate logistic regression models with coefficients expressed in terms of odds ratios to ease interpretation. Juarez and Martin (2006) and Miranda *et al* (2011) each regress condom usage at last intercourse on various demographic covariates and covariates related to an individual’s sexual history using logistic models. Calazans *et al* (2005) regresses lack of condom usage at last intercourse on these factors using a logistic model; since condom use is a binary variable, the direction of his results can be flipped to represent condom use at last intercourse to ease comparison with the previously mentioned studies. Silveira *et al* (2005) use the same dependent variable as Juarez and LeGrand (2006) and Miranda *et al* (2011), but the authors choose a Poisson regression to analyze their data because the outcome prevalence in their study (i.e., condom usage at last sexual intercourse) is greater than 10%, which can lead to overestimation of odds ratios. Gupta (2000) uses contraceptive use at first intercourse as the dependent variable in a logistic model, but notes that condoms represent 87% of the contraceptives used at first intercourse in the sample. Juarez and LeGrand (2005) similarly use condom usage at first sexual intercourse as the dependent variable in a logistic model. Merchan-Hamann *et al* (2002) uses Wilcoxon rank-sum tests to identify statistically significant covariates to the dependent variable “risky sexual behavior,” which the authors define as less than a frequent level of condom usage (frequency of condom use is self-reported as always, frequent, occasional, or rarely/never).

The studies converge on the following four points: (1) a significant difference in condom usage among young people according to the type of partnership at last sexual intercourse, with use being more frequent with casual partners than with steady partners; (2) a significant decline in condom usage with increasing age after reaching a certain threshold around age 20 (during adolescence, increasing age may predict higher prevalence of condom usage); (3) a significant positive association

between education and condom usage; (4) a positive association between identifying as Catholic and condom use at last intercourse. Other demographic variables have conflicting directional effects and significance within the group of studies.

### **3 Theoretical Framework**

Two major theoretical frameworks exist to evaluate contraception use in the developing world: the desired fertility model and the family planning gap model. Pritchett (1994) provides evidence for the economic argument that couples behave rationally to achieve fertility targets, thinking of children somewhat like durable goods. Couples jointly choose their bundle of children based on the utility derived from each additional child, constrained by the number of offspring they can afford. To test the theory, Pritchett (1994) combines data from the World Health Survey and Demographic and Health Surveys to compare women's self-reported desired fertility with their actual fertility, as well as their access to family planning clinics and the availability of contraception. Pritchett (1994) finds that changes in individual demand for children rather than the proliferation of family planning clinics fuel the demographic transition. In his model, income and education are statistically significant factors that decrease a woman's desired fertility.

Although employment is not included in the model, the previous two results suggest that it would also reduce a woman's desired fertility since all three factors signal the advancing status of a woman in society. Additionally, employment opportunities induce a substitution effect in which time spent raising children becomes more costly since the woman could use that time to earn income. The substitution effect would cause women to reduce their desired fertility. Finally, employment would reduce the time a woman could spend raising children, reducing her temporal budget constraint and thus limiting fertility. For men, employment ought to increase desired fertility, since men are typically expected to work formally rather than raise children. Stable employment would increase desired fertility due to reduced fears of uncertainty regarding future employment. Employment could also expand the monetary budget, allowing couples to meet previously unattainable fertility desires. Decreased desired fertility in women predicts a higher likelihood of condom usage at last

intercourse, meaning that for women employment ought to have a positive relationship with condom use, according to the desired fertility framework. Increased desired fertility in men predicts a lower likelihood of condom use, meaning that for men employment ought to have a negative relationship with condom usage.

Pritchett (1994) notes that the emergence of AIDS—an important consideration in Brazil—introduces complications into contraceptive decision-making that his model does not address. Further, the desired fertility framework downplays contraceptive costs in low-income regions as well as potential incongruences introduced by gender power dynamics in assuming men will comply with women's desired fertility.

Other economists argue a family planning gap exists due to the inaccessibility of contraception. Bongaarts and Sinding (2009) cite the simple statistic that 137 million women reported to the 2002 WHS that they did not hope to get pregnant but failed to practice any form of modern contraception as evidence that the family planning gap exists. Bongaarts and Sinding (2009) provide a brief overview of the theoretical factors that contribute to an individual's unmet desire for contraception: monetary costs of the contraception, monetary and temporal costs of transportation to providers of contraception, and social barriers like shame or spousal resistance.

In testing the theory, Bongaarts (1997) finds subsidized condoms and birth control pills in family planning clinics of developing countries contributed to half of the overall decline in fertility rates between 1960 and 1990. Tsui and Herbert (2011) contribute a more exact model showing for every 16 percentage point increase in contraceptive use, national fertility rates in developing countries fall by one child per woman. After controlling for income in this model, employment predicts a decrease in the rate of condom usage for both men and women due to the reduced time with which they have to procure contraception. Again, the family planning gap framework does not explicitly incorporate the effects of potential transmission of sexually transmitted infections, which could potentially alter individual preferences if individuals attempt to minimize their health risks by adopting condom usage. In this scenario, those who previously could not afford to buy contraception would substitute money or time away from other goods to buy protection from the risk of STIs.

Juarez and LeGrand (2005) model the estimated effect of a variety of demographic factors on condom use of low income males at first intercourse, and although the authors do not include employment in the regression, they do discuss potential negative mechanisms specific to the Brazilian context that align with the family planning gap framework. Discussing adolescents, Juarez and LeGrand (2005) argue that although the government and various non-governmental organizations distribute condoms for free, the exchange only occurs at health clinics, which have weekday working hours and require registration. Because adolescents are less likely than adults to plan ahead, younger people often do not have access to condoms as they need them. However, a similar argument applies to the employed: Because fully employed individuals work during business hours, they may not be able to participate in these programs. If they are low-income earners, alternative sources of condoms may be prohibitively expensive.

One relevant consideration when using employment as the variable of interest is the diverging characteristics of employment between men and women. Employment often conflicts with child-bearing desires of women, and it is possible that women are less likely to work if they are members of households within a high-income bracket. Further, differing questions related to female employment signal different potential mechanisms. For instance, if women who earn income are more likely to use condoms, then the additional monetary resources potentially triggered the substitution effect under the desired fertility framework or removed barriers to contraception when considering the family planning gap theory. Alternatively, if working outside the home predicts increased condom usage for women, then diminished temporal resources to devote to children decreases fertility using desired fertility theory, whereas increased mobility through improved transportation removes a barrier to contraception under the family planning gap framework.

Given these considerations, I posit that an individual's likelihood for condom use at last intercourse is a function of his desired fertility, his demographic characteristics that will act as constraints under the family planning gap framework, and his partners demographic characteristics:

$$CU_i = f(DF_i, DC_i, DC_p) \quad (1)$$

Although neither framework accounts for the implications of sexually transmitted infections, risk of transmission does represent a possible alternative mechanism through which individuals alter their sexual health behavior based on an perceived metric of their level of risk. Although testing that mechanism is not the goal of this paper, adding a vector of covariates to the model indicating whether an individual participates in risky sexual behavior and understands how to mitigate that risk through condom usage. Controlling for these variables in the regression will help to test whether the estimated effect of employment on condom use at last intercourse is robust.

## 4 Data

The 1996 Brazil Demographic and Health Survey is a nationally representative cross-sectional study of 13,283 households in Brazil decomposed into individual interviews with 12,612 women; 2,949 men; and 1,319 couples. The two-stage clustered sampling technique first randomly selects districts to sample with weighting based upon seven regional populations before randomly subsampling households within those districts. Therefore, the study samples more clusters and more households from the northeastern region—which contains over 30% of Brazil’s population—than from any of the other regions. In general, clustered data may correlate at the community level, because a variety of endogenous factors present in the community that influence behavior, warranting the use of clustered standard errors in subsequent regression models.

The survey samples individuals ages 15 and older, meaning certain adolescents younger than 15 who have already reached the point of sexual initiation will not be represented in this study of condom use at last intercourse. Unlike other DHS studies, the sample of women includes individuals who are both married and unmarried. DHS studies from other countries often include only unmarried women in the sample.

Questions regarding sexual health behaviors such as condom usage are retrospective and self-reported in the 1996 Brazil DHS (i.e. Did you use a condom at the time of your last sexual intercourse?), which may seem intuitively unreliable. However, retrospective surveys are frequently used in sexual health research, and both Morris (1993) and Cleland and Ferry (1995) show that the

data gathered in retrospective surveys on sexual behavior are generally reliable in both developed and developing countries, respectively. Further, Becker and Costenbader (2001) analyze the data provided by couples in the 1996 Brazil DHS to compare husbands' reported contraceptive use compared to that of their wives. In Brazil, they find that husbands did not report significantly higher condom usage at the time of last intercourse than their wives—a conclusion which increases the validity of the accuracy of the self-reported responses recorded in the data set.

The DHS is highly appropriate for studying the association between employment and condom use at last intercourse because it: (1) is a nationally representative data set with a large sample size; (2) includes a variety of measures of employment (i.e., current employment, self employment, paid employment, work outside the home); (3) includes a rich set of relevant control variables related both to demographics and to the respondent's sexual history, which are thought to influence condom use based on economic theory and ease comparisons to prior studies using similar controls; (4) is highly regarded for reporting well-collected, quality data in contrast to the majority of the studies in the relevant literature whose authors conduct their own surveys to collect data; and (5) includes state indicators, allowing the model to control for potential omitted variable bias through geographic fixed effects.

The DHS program has conducted three studies in Brazil: one in 1986, another in 1991, and a third in 1996. Although all three meet the five benefits of using DHS data listed above, I choose to analyze the data from 1996 for two reasons. First, Gupta (2000) combines the three surveys and finds that while the number of adolescent women reporting condom usage at first intercourse in 1986 was negligible, the rate had increased to nearly half of women in 1996. Therefore, I choose the 1996 data to introduce heterogeneity into the dependent variable, condom use at last intercourse. Second, as mentioned in the literature review, the year 1994 represents a natural break in HIV prevention policy in Brazil due to the election of a new president who actively pursued creation of a National AIDS Program with funding from the World Bank. Therefore, awareness and availability of condoms ought to be greater after the 1994.

Despite these advantages, there are limitations to the DHS data. First, the survey asks no questions to discern whether individuals are employed part-time or full-time, thus limiting the

depth of my analysis. Second, although the survey asks participants to report their current income, the DHS obscures that data to preserve anonymity. Therefore, income—perhaps the most common of the demographic controls included in studies on sexual behavior in the developing world—cannot be used as a covariate when relying on the DHS. The DHS does construct a household wealth index for individuals based upon the combined incomes of household members and ownership of various appliances. Third, individuals clearly bias certain responses due to social stigma. Only 2 men out of the 2,949 sampled report ever engaging in homosexual behavior; 38 individuals do not respond to the question. Populations like men who have sex with men are often separated into subgroups due to their high risk of contracting STIs. However, due to the homogeneity of the responses, I can neither exclude them from the analysis sample, nor control for them with an indicator variable in regression models. Finally, the survey does not ask male or female participants whether they are currently trying to conceive a child. Questions related to the respondents desire for children were asked and answered with a high response rate; however, because this data lacks a timing component, it cannot distinguish individuals who did not use a condom at last intercourse in an attempt to conceive.

My analysis sample is separated into a data set for men and another for women both of which include only sexually active individuals of reproductive age who are not otherwise impeded from conceiving a child through sterilization; infecundity; or in the case of women, current pregnancy. Reproductive age is defined as age 15 through 45 for women based on the background literature. No such age restriction exists for the male sample, which includes individuals between age 15 and 60. Although male ages could have been restricted for symmetry with the female data, they were not for four reasons: (1) no theoretical reason exists to exclude men over 45 when considering condom usage as men remain fertile throughout their life; (2) 14 percent of men over age 45 sampled in the DHS reported using a condom at last intercourse—although this number is lower than the rate of condom usage for younger age groups, it still shows prevalence of use among older demographics; (3) regression models described later in this paper that restricted male age to a maximum of 45 (but were ultimately not included in this paper) lost valuable evidence of trends in condom usage related to age; and (4) restricting the age of men in the analysis reduces the sample size of an already small group.

Although all sexually active individuals can benefit from condom usage to reduce the incidence of STI transmission, those who are not currently capable of conceiving a child cannot reap the family planning benefits of condom use—an important consideration given 36 percent of the male DHS sample and 25 percent of the female DHS sample report using condoms only for family planning purposes and not for STI protection. Both samples also exclude respondents who are currently enrolled in school, as full-time students are less likely to be employed. Professions that would systematically decrease access to condoms due to geographic constraints like serving in the military are not present in the raw DHS data because surveys are only conducted in residential households. A step-by-step derivation for each set is listed below.

Of the 2,949 male participants in the DHS, only 2,701 had previously engaged in sexual intercourse. Of those, 15 respondents were dropped from the analysis because they reported no knowledge of condoms. An additional 823 were dropped because they reported using an alternate form of contraception at last intercourse instead of using a condom. 370 were dropped because they were sterile—either naturally or due to vasectomy, leaving 1493 individuals who meet the requirements of these restrictions related to the dependent variable condom use at last intercourse. An additional 354 participants were excluded from the male analysis sample because they were still enrolled in school, which would somewhat impede current employment, leaving 1,139 observations. All remaining participants responded both to the survey question regarding condom use at last intercourse and the question regarding current employment. An additional 43 participants were dropped from the analysis sample due to non-response to survey questions used as covariates. The final analysis sample includes 1,096 observations.

Those in the analysis sample are more likely to have used a condom at last intercourse, more likely to be currently employed, more likely to live in rural areas, more likely to fall into the lowest wealth quintile, less likely to fall between the ages 15 and 19, and less likely to be married.

Of the 12,612 female participants in the DHS, only 9,952 had previously engaged in sexual intercourse. Of those remaining, 67 respondents were dropped from the analysis sample because they reported no knowledge of condoms. An additional 3,158 were dropped because they reported using an alternate form of contraception at last intercourse instead of using a condom. 548 were



dropped because they were currently pregnant. 2,497 were dropped because were infertile or had undergone a sterilization procedure. Finally, 306 women were dropped from the sample because they were over the maximum reproductive age of 45, leaving 3,378 individuals who meet the requirements of these restrictions related to the dependent variable condom use at last intercourse. An additional 536 participants were excluded from the female analysis sample because they were still enrolled in school, which would somewhat impede current employment. After these exclusions based on theory had been made, an additional 5 women were dropped from the sample due to non-response to the survey question regarding condom use at last intercourse. Finally, 6 participants were dropped due to non-response to survey questions regarding their employment. An additional 337 participants were dropped from the analysis sample due to non-response to survey questions used as covariates. The final analysis sample includes 2,494 observations.

Those in the analysis sample are more likely to have used a condom at last intercourse, more likely to fall within the lowest wealth quintile, more likely to have at least one child, less likely to fall between the ages 15 and 19 and less likely to have had multiple sexual partners.

Table 1 presents summary statistics of the dependent variable condom usage at last intercourse and the variables of interest related to employment for both the male and female analysis samples separately.

## 5 Methodological Approach

In modeling the impact of current employment on condom use at last intercourse among men and women in Brazil, I estimate the following equation:

$$CU_i = g(E_i, DF_i, DC_i) \quad (2)$$

where condom use at last intercourse is a function of an individual's current employment  $E$ , his desired fertility  $DF$  which controls for the desired fertility theory of contraceptive use, and his demographic characteristics  $DC$  which control for variables included in the family planning gap framework. As noted, few prior studies examine the effect of employment on condom usage.

Those studies that do include employment as a covariate in regression models do not compare the differences in the effect employment has on condom use among men and women in Brazil.

My study begins the process of addressing this gap by examining various measures of employment through probit models for a male analysis sample and a female analysis sample. The male analysis sample and the female analysis sample are examined separately to more fully understand the effect of employment on each gender due to the many covariates included in the regression that are expected to have diverging directional effects between genders based on the literature review.

Five differing measures of employment are introduced into separate models as variables of interest. The variables of interest introduced into the models of male behavior are: (A) an indicator variable signaling whether the individual is currently employed and (B) a factor variable measuring whether the respondent is employed year-round as opposed to seasonally, occasionally, or not employed. These are not included in the same model, but rather run independently due to collinearity between the variables. The variables of interest introduced into the models of female behavior are: (C) an indicator variable signaling whether the individual is currently employed, (D) an indicator variable representing whether the individual earns income, and (E) an indicator variable signaling whether the individual works outside of her home as opposed to working inside the home or not working. Again, these are treated as three independent variables of interest and tested individually in separate models.

The analysis estimates multivariate probit models for the effects of each measure of employment on condom use at last intercourse, controlling for a host of potentially confounding variables that account for desired fertility framework and the family planning gap framework. A number of different model specifications were considered to assess robustness of the results and explore patterns within the estimates. Two of these modifications to the general model are included in the results of the paper. The first adds state indicators to the model. The second adds a vector of covariates related to participation in risky sexual behaviors and sexual education, which as argued earlier present a different mechanism by which an individual would alter his sexual health behavior and therefore can test the robustness of the estimated effects for each measure of employment.

## 6 Multivariate Analysis

Table 2 presents a summary of probit estimates of the effects of the five measures of employment on the probability of condom use at last intercourse using four differing models with increasing adjustments. Model 1 includes no covariates. Model 2 includes a measure of desired fertility and a vector of covariates representing demographic characteristics. Model 3 adds state indicators to the controls in Model 2. Model 4 adds covariates related to participation in risky sexual behaviors and sexual education, which as argued earlier present a different mechanism by which an individual would alter his sexual health behavior and therefore can test the robustness of the estimated effects of employment. Each cell contains the estimated probit coefficient, the standard error corrected for state clustering of observations (in parentheses), and the marginal effect [in brackets].

In all models, the varying measures of employment exhibit diverging effects for men and women. The direction of the effect of increased employment for men is negative across all four models and both variables of interest, even when not statistically significant. The direction of the effect of increased employment for women is positive across all four models and the three variables of interest, even when not statistically significant. The magnitude of marginal effects generally declines as more covariates are added regardless of the variable of interest, with the exception of the marginal effect of being employed year-round for men which remains relatively constant. The largest declines in the magnitude of the marginal effects occur between Model 1 and Model 2 (i.e., when adding only the basic covariates: desired fertility and the vector of demographic characteristics). Introducing state indicators in Model 3 does not substantively affect the estimates. Therefore, although the decentralized Brazilian HIV-prevention policy is determined at the state level, introducing state indicators into the model does not over-control for differences in the availability of condoms. Model 4 introduces the vector of covariates indicating risky sexual health behaviors and sexual education. Again, the addition of these covariates does not reduce the magnitude of the marginal effects greatly. That said, while the marginal effects remain virtually unchanged between the latter three models with either variable of interest using the male analysis sample, the marginal effects do decrease by about 0.5 percentage points in each step between Models 2, 3, and 4 with all three measures

of employment when analyzing the female analysis sample, indicating potential over-control from addition of state indicators or a lack of robustness in the effect of these measures of employment on condom use at last intercourse when faced with other possible mechanisms that modify sexual health behavior.

Two of the five variables of interest maintain statistical significance through each of the model iterations: employed year-round (B) for men and working outside of the home (E) for women. Both remain significant at the 1% level across all four models. Current employment (A) for the male sample is not statistically significant in any of the models. In the analysis of the female sample, current employment (C) and earning income (D) are statistically significant at the 1% level in Model 1 without covariates; however, both diminish in significance falling to the 5% level after the basic covariates are added in Model 2. Thus, the effects of employment and earning income in Model 1 were potentially confounded by the effects of wealth and education, which both have a significant positive relationship with condom use at last intercourse correlate positively with employment measures in the female analysis sample. Both C current employment and D earning income lose statistical significance above the 10% level after adding the risky sexual behavior indicators in Model 4.

Examining employed year-round (B) and working outside the home (E) specifically, the sizes of the estimated effects are quite large. In the most adjusted model, Model 4, holding year-round employment (B) decreases the probability of a man using a condom at last intercourse by 11.3 percentage points. From the relevant sample means presented in Table 1, 36% of men in the analysis sample used a condom at last intercourse. Therefore, an 11.3 percentage point decrease represents a 31% reduction in the probability of a man reporting condom use at last intercourse. Working outside the home (E) increases the probability of a woman using a condom at last intercourse by 4.9 percentage points. 22% of women in the sample used a condom at last intercourse. A 4.9 percentage point increase represents a 22% increase in the likelihood of a woman reporting condom use at last intercourse.

The remaining discussion of the results will focus on Models 3 and 4, using employed year-round (B) and working outside the home (E) as the variables of interest for the male sample and

the female sample, respectively. The results of Model 3 and Model 4 are similar to each other in terms of maintained levels of statistical significance of covariates across both models, consistent direction of the estimated effect of each covariate, and similar magnitude of the marginal effects of each covariate. Therefore, variation between the models when considering the analysis of either subgroup is not an issue.

Statistically significant positive demographic predictors of condom usage at last intercourse for both men and women are higher levels of education (specifically, the completion of secondary school and pursuing higher education) and greater levels of wealth. The only significant negative demographic predictor in common for both men and women is being married. The only demographic covariate that is significant for both populations with diverging directional effects is age. In the models of the male sample, each 5-year age group under 35 years of age is a statistically significant positive predictor of condom use at last intercourse using the group Age 15–19 as the baseline reference category. The direction on the effect of age changes to negative in each of the groups above 40 years of age, and becomes significant with the groups Age 50–54 and Age 55–59. Models based on the female sample estimate the effect of age as consistently negative and generally significant, again using the group Age 15–19 as the baseline.

Certain covariates are significant for only one of the samples. When using the female analysis sample, desired fertility is a significant predictor of condom use at last intercourse at the 1% level. Wanting children after 2+ years, wanting children but being unsure of desired timing, and wanting no more children all increase the likelihood of a woman using a condom at last intercourse by between 35.4% and 57.3% compared to the relevant sample mean when using those women who want children within the next 2 years as the baseline. None of these categories are significant predictors of condom use at last intercourse for men. Using the male subgroup, identifying as Roman Catholic is a positive predictor of condom use, significant at the 10% level compared to a baseline of respondents who are not religious.

The directions of the estimated effects in these models are largely consistent with those in the relevant literature, on variables where prior results converge. Higher levels of education are generally reported as a positive predictor of condom use in the literature reviewed previously,

whereas being married has a negative association with condom use at last intercourse. Increasing age in our models is a positive predictor of condom use for young men, as found Juarez and Martin (2006) who restrict the target population to males under age 20, but negative for women as seen in the findings of Miranda *et al* (2011) and Silveira *et al* (2005). Being Catholic has typically been reported as a positive predictor of condom use for both men and women; however, Gupta (2000) is the only study that finds a highly significant positive relationship between being Catholic and condom use through an indicator variable that signals whether an individual is Catholic and groups respondents who are practitioners of other religions with respondents who are not religious in the category non-Catholic. In our data, 7% of respondents in the male subset and 12% of those in the female subset report practicing Evangelism. These are large portions of the sample population and being Evangelist has a negative association with condom use at last intercourse in the regression models. Therefore, grouping them with respondents who are not religious likely distorts the estimated effect of being Catholic, explaining the disparity in the findings with those of Gupta (2000).

For the male sample, being married has the largest marginal effect of any of the demographic covariates, reducing the probability of condom use at last intercourse by 23.8 percentage points, or 34% compared to the relevant sample mean. For the female sample, desired fertility has the largest marginal effect, with women who desire no more children being 12.6 percentage points, or 57.3% compared to the relevant sample means, more likely to use a condom at last intercourse.

Although adding the vector of covariates related to risky sexual behavior and sexual education into Model 4 does not alter the estimated effects of the demographic covariates, most of the additional covariates are statistically significant themselves. Both of the risky behavior indicators included in the model (having casual sex at last intercourse and having multiple sexual partners in the 12 months preceding the administration of the survey) are positive indicators of condom use at last intercourse. The two variables are correlated and therefore problems with collinearity may arise. However, this does not reduce the predictive power of the model or bias the estimates of individual coefficients, but does increase the standard errors on the correlated variables.

Being able to correctly identify that condoms prevented the transmission of HIV—a covariate

related to the individual’s level of sexual education—increases the likelihood of both men and women using condoms at last intercourse, and is significant at the 1% level for both groups. Finally, a covariate which indicates whether the respondent had been diagnosed with a sexually transmitted infection in the 12 months preceding the survey has a negative association with condom use at last intercourse in the male sample, probably due to reverse causality. The variable is not statistically significant. The STI indicator variable is a significant positive predictor of condom use at last intercourse for women; however, as mentioned previously, an unrealistic proportion of women report having an STI in the 12 months preceding the survey. Therefore, we ought to be cautious in interpreting the estimated effect of the variable when using the female subgroup; however, it may serve as a useful proxy for recent medical attention and sexual health counseling.

## 7 Conclusion

Using nationally representative data from the 1996 Demographic and Health Survey, I find robust evidence that certain measures of employment are strongly associated with condom use at last intercourse. The directional effects of employment diverge between men and women, with men experiencing negative associations between employment and condom usage while women experience positive associations between the two.

The effects of the study have interesting implications for the two theoretical frameworks considered. Although my probit models do not fully test whether the desired fertility model or the family planning gap model better predicts health behavior related to contraceptives, the inclusion of the controls indicating a respondent’s desired fertility do suggest that women act as rational economic agents, increasing their use of contraceptives when they have no desire for more children, or only want more after two or more years. In contrast, desired fertility was not a significant predictor of male behavior. The desired fertility framework may better predict women’s actions because they have more agency over their pregnancy or the actions of married men who desire a child at the same time as their partner.

Future research is needed complement the findings of this study by continuing to investigate the

effects of other measures of employment such as working part-time on condom usage. Additionally, more rigorous economic theory to account for the implications of sexually transmitted infections via the risky sexual behavior mechanism I briefly introduce to check for robustness ought to be considered due to the high level of significance and magnitude of the results.

## References

- [1] Amaral, Ernesto & Daniel Hamermesh (2007). “Macroeconomic and Policy Implications of Population Aging in Brazil.” NBER Working Paper, No. 13533. Cambridge, MA.
- [2] Becker GS (1960). “An Economic Analysis of Fertility, Demographic and Economic Change in Developed Countries.” National Bureau Committee for Economic Research Conference Paper, 209–231.
- [3] Becker, Gary S (1991). *A Treatise on the Family*. Cambridge, MA: Harvard University Press.
- [4] Bongaarts, J. (1997). “The Role of Family Planning Programmes in Contemporary Fertility Transitions: The Continuing Demographic Transition.” Oxford: Clarendon Press, 422–444.
- [5] Bongaarts, John & Steven Sindling (2009). “A Response to Critics of Family Planning Programs.” *International Perspectives on Sexual and Reproductive Health*: 35(1).
- [6] Calazans, Gabriela, Teo Araujo, Gustavo Venturi, & Ivan Franca (2005). “Factors Associated with Condom Use among Youth Aged 15-24 Years in Brazil in 2003.” *AIDS* 19(4): 42–50.
- [7] Carvalho, J. & L. Wong (1999). “Demographic and Socioeconomic Implications of the Rapid Fertility Decline in Brazil: a Window of Opportunity.” *Reproductive Change in India and Brazil*: 208–239
- [8] Flrez, C.E. & J. Nez (2003). “Teenage Childbearing in Latin American Countries.” *Critical Decisions at a Critical Age*: 47–92.



- [9] Gmez, Eduardo (2010). “What the United States Can Learn From Brazil in Response to HIV/AIDS: International Reputation and Strategic Centralization in a Context of Health Policy Devolution.”
- [10] Gupta, Neeru (2000). “Sexual Initiation and Contraceptive Use among Adolescent Women in Northeast Brazil.” *Studies in Family Planning* 31(3): 228–238.
- [11] Jorgensen, O.H. (2011). “Macroeconomic and Policy Implications of Population Aging in Brazil.” World Bank Policy Research Working Paper, No. 5519
- [12] Juarez, Fatima & Thomas LeGrand (2005). “Factors Influencing Boys’ Age at First Sexual Intercourse and Condom Use in the Shantytowns of Recife, Brazil.” *Studies in Family Planning* 36(1): 57–70.
- [13] Juarez, Fatima & Teresa Martin (2006). “Safe Sex Versus Safe Love? Relationship Context and Condom Use Among Male Adolescents in the Favelas of Recife, Brazil.” *Archives of Sexual Behavior* (35)1: 25–35.
- [14] Lam, D. & S. Duryea (1999). “Effects of Schooling on Fertility, Labor Supply, and Investments in Children, with Evidence from Brazil.” *Journal of Human Resources* 34(1): 160–90.
- [15] Lam, D., & L. Marteleto (2005). Small Families and Large Cohorts: The Impact of the Demographic Transition on Schooling in Brazil. *Growing Up Global: The Changing Transitions to Adulthood in Developing Countries*, Washington, D.C., United States: National Academies Press.
- [16] Martine, George (1996). “Brazil’s Fertility Decline, 1965-95: A Fresh Look at Key Factors.” *Population and Development Review* 22(1): 47–75.
- [17] Martine, George, Monica Das Gupta, and Lincoln C. Chen (1998). *Reproductive Change in India and Brazil*. Delhi: Oxford University Press.

- [18] Mason, A. (1995), “Demographic Transition and Demographic Dividends in Developed and Developing Countries.” Department of Economics, University of Hawaii at Manoa and Population and Health Studies.
- [19] Merchan-Hamann, Edgar, Maria Ekstrand, Esther Hudes, & Norman Hearst (2002). “Prevalence and Correlates of HIV-Related Risk Behaviors among Adolescents at Public Schools in Brasilia.” *AIDS and Behavior* 6(3): 283–293.
- [20] Merrick, T. & E. Berqu (1983). *The Determinants of Brazil’s Recent Rapid Decline in Fertility*. Washington, DC, United States: National Academies Press.
- [21] Ministerio da Saude, Brasil (2002). “Acesso ao preservativo faz a diferenca: Prevention access to condoms makes all the difference.” Resposta positiva experiencia do programa brasileiro de AIDS: 28–36
- [22] Miranda, AE, NC Figueiredo, W McFarland, R Schmidt, & K Page (2011). “Predicting Condom Use in Young Women: Demographics, Behaviors, and Knowledge from a Population-Based Sample in Brazil.” *International Journal of STD & AIDS* 22(1): 590–595.
- [23] Okie, Susan (2006). “Fighting HIV - Lessons from Brazil.” *New England Journal of Medicine*.
- [24] Pritchett L (1994). “Desired Fertility and the Impact of Population Policies.” *Population and Development Review*, 20(1): 1–55.
- [25] Rutenberg, N., L.H. Ochoa & J.M. Arruda (1987) “The Proximate Determinants of Low Fertility in Brazil.” *International Family Planning Perspectives* 13(3): 75–80
- [26] Silveira, Mariangela, Ina dos Santos, Jorge Beria, Bernardo Horta, Elaine Tomasi, & Cesar Victora (2005). “Factors Associated with Condom Use in Women from an Urban Areas in Southern Brazil.” *Cadeirno Saude Publica* 21(5): 1557–1564.

## 8 Appendix

**Table 1: Relevant Sample Characteristics**

	Male Analysis Sample	Female Analysis Sample
Condom Used at Last Intercourse	.36	.22
Currently Employed	.82	.49
Employed Year-Round	.82	–
Earns Income	–	.57
Works Outside of Home	–	.44
Sample Size	1,096	2,494

Notes: Means are reported as proportions given the large number of dichotomous variables, unless otherwise indicate. Standard errors are presented below the means of continuous variables parenthetically.

**Table 2: Effects of 5 Measures of Employment on Condom Use at Last Intercourse**  
**Using Male and Female Analysis Samples**

	<b>No Covariates (1)</b>	<b>Basic (2)</b>	<b>+ State Indicator (3)</b>	<b>+ Risk Behavior (4)</b>
	<i>Coefficient (SE) [ME]</i>	<i>Coefficient (SE) [ME]</i>	<i>Coefficient (SE) [ME]</i>	<i>Coefficient (SE) [ME]</i>
<b>Male Sample, N=1096</b>				
Currently Employed, (A)	-.141 (.088) [-.052]	-.114 (.109) [-.033]	-.125 (.110) [-.034]	-.120 (.121) [-.032]
Employed Year-Round, (B)	-.336** (.147) [-.129]	-.396*** (.129) [-.116]	-.384*** (.129) [-.107]	-.417*** (.134) [-.113]
<b>Female Sample, N=2494</b>				
Currently Employed, (C)	.267*** (.046) [.078]	.135** (.053) [.036]	.116* (.060) [.030]	.087 (.065) [.023]
Earns Income, (D)	.336*** (.098) [.098]	.132** (.059) [.035]	.109** (.060) [.028]	.077 (.065) [.020]
Works Out of Home, (E)	.391*** (.051) [.113]	.226*** (.054) [.061]	.211*** (.054) [.055]	.192*** (.052) [.049]

Notes: \*\*\* p<0.01; \*\* p<0.05; \* p<0.10. SE = Standard Error. ME = Marginal Effect. Basic Covariates include all Demographic Characteristics and Desired Fertility from Table 2. Risky Behavior Mechanism includes all Risky Sexual Behaviors and Understanding of Risk covariates from Table 2.

# Creativity of Fashion Design: An Economic Creative Lifecycle Analysis of Three 20<sup>th</sup> Century Fashion Designers

Benjamin Nickerson  
University of Chicago

May 2015

## **Abstract**

Innovations and inventions are often cited as sources of economic growth. These drivers can be thought of as functions of creativity. To study creativity is therefore to study aspects of economic growth. This paper builds upon research on creativity done by Professor David Galenson of the University of Chicago, who has found that innovators can be categorized into two groups: conceptual and experimental. Conceptual innovators often produce their most innovative works early in their careers while experimental innovators dedicate their careers towards perfecting their craft and create later innovations. This paper analyzes the artistic and creative careers of three 20th century haute couture fashion designers in order to better understand what drives innovation in the fashion industry. Each artist's career is systematically measured by three key metrics including museum exhibition records, auction price data and museum collection dates. Essays and reviews written by notable critics of the era further support claims derived from the data. This paper finds that Cristóbal Balenciaga and Madame Grès are experimental designers while Yves Saint Laurent is conceptual.

# 1 Introduction

Innovation is the diffusion of an idea. Innovative art changes a discipline by shifting practices, altering preconceived notions, and establishing new styles or methods by which the work is made. Prior empirical studies on creativity have found that artistic innovations tend to occur in two distinct patterns: sudden, specific, and precise; or gradual and steady (Galenson, 2009). The first type of innovations, those that are sudden and specific, are conceptual. Conceptual innovators tend to work deductively, using ideas as the foundation for their innovation, and make their greatest contributions earlier in their careers (Galenson, 2009). The second type of innovations, those that occur more gradually over time, are experimental. Experimental innovators tend to work inductively, forming their innovations through observation, and produce their innovations later in their careers (Galenson, 2009).

This paper will analyze innovations specific to fashion design. This paper will use museum exhibition checklists and auction price data to empirically study the creative lifecycles of three fashion designers. This paper will also refer to criticism and interviews by designers and museum curators to further study each designer. By analyzing the lifecycles and contributions of major 20th century fashion designers, this paper will define what it means to be a conceptual or experimental innovator in the context of the discipline and come to understand which 20th century fashion designers were conceptual and which were experimental. The designers studied include experimentalists Cristóbal Balenciaga and Madame Grès and conceptualist Yves Saint Laurent (see Table 1).

# 2 Couture Fashion Design

Haute couture is the term used to define the high-end fashion industry of the 20th century (Martin & Koda, “Haute Couture”) The origins of haute couture can be traced to the House of Worth in Paris during the mid-19th century. The term refers to the construction of clothing for individuals from start to finish by a designer or couturier. The creative ability and artistic independence of the designer makes this discipline particularly relevant for the study of innovation. In high-end fashion design, a designer creates a finished, signature product that becomes part of a larger body

of work over the course of a designer’s career. Analysis of the complete body of work can lead one to determine unique characteristics of the designer, which works were most innovative, and what influence the works had on other designers. By doing so, this paper looks to lead to a new understanding of creativity in the context of haute couture.

Among the many notable artists in the history of fashion design, the designers studied in this paper were chosen because of their innovative contributions to couture as a form of art. Many famous designers, such as Charles Frederick Worth and Paul Poiret of the early-20th century or Ralph Lauren and Calvin Klein of the late-20th century, are well known today because of their significant contributions to fashion as a business. (Martin & Koda, “Haute Couture”) This paper differentiates between fashion as a business and fashion as an art and studies the designers who made their most important contributions to fashion as a form of art. By focusing on the art rather than the business, one can determine what techniques, styles, sources of inspirations, and final products correspond with conceptual or experimental innovations in the art of fashion design.

### 3 Selection Criteria

Cristóbal Balenciaga, Madame Grès, and Yves Saint Laurent were selected by comparing lists of designers who have the most number of items in the Costume Institute of the Metropolitan Museum of Art in New York (Met) and the list of retrospective exhibitions hosted by the Costume Institute. These lists can be seen in Table 2 and Table 3, respectively. The Costume Institute was founded in 1937 and contains over 35,000 costumes dating back to the 15th century. In 2009, the Brooklyn Museum joined its costume collection with the Met, creating the “largest and most comprehensive costume collection in the world” (“Costume Institute”). The historical and geographical span of the Met’s costume collection makes for a thorough and objective process of selecting the most important fashion designers to study.

In addition to holding many of the world’s most important costume items, the Costume Institute also hosts retrospective exhibitions that are useful to study the career path of major fashion designers. According to Andrew Bolton of the Costume Institute, the museum criteria for hosting an

exhibition on an individual designer is “whether the designer changed the course of fashion history” (Menkes, 2011). Given that innovations are pieces of art that diffuse and significantly change a discipline, the designers featured in monographic retrospectives must have made at least one major innovation over the course of their career. Using both the number of items in the collection and the list of individual retrospectives will ensure selecting designers with significant, innovative careers.

## 4 Measurement

Retrospective museum exhibitions offer a unique opportunity to examine the important pieces and periods of an artist’s career. Prior studies have found that museum exhibitions serve as a way for curators to “tacitly reveal their judgments of the importance of an artist’s work at different ages” (Galenson, “Old Masters and Young Geniuses”). Thus, the most important ages of an artist’s career will tend to have the most works in an exhibition. However unlike painting or other forms of art, monographic fashion retrospectives are a relatively new phenomenon. The Met hosted their first monographic exhibition in 1973 with “The World of Balenciaga” and their second in 1983 on Yves Saint Laurent (see Table 2). This raises an issue concerning the validity of these early exhibitions on accurately measuring the career of an artist. Since the concept of these types of shows is new, there is perhaps a gap in the scholarship when compared to other artistic disciplines that would lead to the same quality of exhibitions painters receive. To overcome this issue, the paper will analyze multiple exhibitions of each artist from other museums to obtain a larger sample size to overcome any idiosyncrasies of a given museum.

The 1983 Met exhibition on Yves Saint Laurent was particularly unique for the purposes of this study because it occurred while Saint Laurent was still designing. In a 2011 New York Times article, Katell le Bourhis of the Met stated that at the time of the exhibition Yves Saint Laurent’s couture house had only “a single brown lace dress available” for the exhibition; the museum had to find donors to lend their personal collections to the museum for the exhibition (Menkes, 2011). However, Saint Laurent’s partner Pierre Bergé established a foundation soon after to preserve Saint Laurent’s work. The foundation hosted an exhibition after Saint Laurent’s death in 2002 with the



Fine Arts Museums of San Francisco and Montreal Museum of Fine Arts. This is the exhibition used in this study. Nathalie Bondil of the Montreal Museum of Fine Arts stated she “insisted on having an independent curator and expert” on the team preparing the show because she believed in the importance of freedom from commercial branding of exhibitions (Menkes, 2011). This increases the usefulness of the 2002 exhibition as evidence for measuring Saint Laurent’s career path. John E. Buchanan Jr., director of the Fine Arts Museums of San Francisco, makes an argument that further supports the use of exhibitions as a tool for measuring a fashion designer’s career:

In considering a monographic exhibition, we look for the genius factor.’ We want the designer who is seminal - who has created a singular vision, silhouette, technique or style unlike that which came before and who has a broad-reaching oeuvre that inspires and influences successive generations of designers (Menkes, 2011).

Therefore, this paper will analyze the number of items and the year of each item produced across several individual retrospective exhibitions as a way of measuring the career of innovative fashion designers. The Victoria and Albert Museum in London, Musée Gallarllia in Paris, and National Gallery of Victoria in Australia have also hosted exhibitions pertinent to the designers studied this paper.

In addition to using exhibitions, this paper will also look at auction price data from several different auctions over the past ten years to study each designer’s career. This data will be used in conjunction with the criticism analysis and exhibition profiles. Prior studies on creativity have shown auction markets to be an accurate predictor of a designer’s most innovative works, as the most important works tend to be the most expensive pieces of an artist’s collection (Galenson, “Old Masters and Young Geniuses”). While the selection size of the auction data might not be large enough to do a complete analysis, the range of the years gathered for each designer will contribute to the analysis from the exhibitions and critical essays.

## 5 Conceptual Designers

Conceptual fashion designers tend to create a single design or style that breaks away from current trends. Like other conceptual artists, conceptual designers often find inspiration for their designs from their ideas or dreams. Furthermore their designs serve as a means to convey the artist's ideas. Conceptual designers often plan their work ahead of times with sketches and then work in teams or groups to produce their designs. They typically receive recognition more for the concept or idea behind the design than the craftsmanship or tailoring elements of the work. Their designs tend to be radical and controversial, as well. Conceptual designers also tend to be more public figures with louder, more outgoing personalities. Yves Saint Laurent is the archetypal conceptual designer studied in this paper.

## 6 Experimental Designers

Experimental designers tend to be committed to a specific style of design and pursue complete perfection within his or her style. They are known primarily for their extremely strong craftsmanship and tailoring skills. Like experimental painters, experimental designers tend to be committed to creating beauty and grace in their designs. Like architects, experimental designers tend to use nature as a source of inspiration. Experimental designers also often return to their finished products, repairing and reworking their designs because of their strong tailoring skills. Experimental designers tend to work in solitude and remain away from the public sphere, often letting their work speak for itself. Cristóbal Balenciaga and Madame Grès are the two experimental designers studied in this paper.

### 6.1 Cristóbal Balenciaga (1895-1972):

“The master of us all” - Christian Dior

Cristóbal Balenciaga is a famous Spanish experimental designer who transformed the couture industry with his unparalleled tailoring and craftsmanship abilities. His contemporary Gabrielle

Coco' Chanel once said, "Of all twentieth-century designers, the only one who could create a garment from beginning to end was Balenciaga" (Healy, 1992). Balenciaga was born in 1895 in a small town in the Basque region of Spain. He worked from the age of 12 as a seamstress for his mother. Using what he learned from his mother, he opened a small house in San Sebastian in 1919 where he began to develop his talents. In 1931 Balenciaga opened a second house, and by 1937 he had become well known and had sold a dress to the Duchess of Westminster the following year. He led one of the most successful and prominent couture houses in Europe until he retired in 1968 (Healy, 1992).

Balenciaga is most recognized for his tailoring skills and knowledge of fabrics. For the Balenciaga exhibition hosted by the National Gallery of Victoria, critic Robyn Healy stated, "Balenciaga's extraordinary tailoring skills and knowledge of fabrics set him apart from all other designers" (Healy, 1992). His skills were so strong he was often described as being like a sculptor: "Balenciaga uses fabric like a sculptor working in marble. He can rip a suit apart with his thumbs and remake or alter his vision in terms of practical, at-hand dress making" (Healy, 1992). This skill is commonly associated with experimental artists in other disciplines. Prior studies on creativity have shown experimental artists tend to be known for their craftsmanship abilities (Galenson, 2009). Developing mastery of fabric and tailoring requires patience and needs to be developed over the course of a career, which is consistent with experimental artists of other disciplines.

Balenciaga's main innovation was in the style and technical aspects of couture. As one exhibition curator wrote, his "style and technical innovation became part of fashion's everyday vocabulary he represented the best in design and construction" (National Gallery of Victoria). Balenciaga used his tailoring skills to create a new silhouette for women. This silhouette combined geometric principles with high-end fabrics to create a garment that flowed freely over a woman's body in a simple yet complex manner. His specific innovations include: "boxy, semi-fitted suits; tunic dresses; 7/8 length sleeves; stand-away collars; voluminous evening coats with dolman sleeves; and magnificent ball gowns" (Kellogg, 2002). The stand-away collars and 7/8 length sleeves represent his technical innovations while the semi-fitted suits and ball gowns represent his stylistic innovations.

Balenciaga's inspiration came from his travels and homeland. In a book by fashion professor

Noel Palomo-Lovinski titled *The World's Most Influential Fashion Designers*, the author describes how Spanish history and painters inspired many of Balenciaga's works. Palomo-Lovinski highlights Balenciaga's "clear connection with historical Spanish painters, such as Francisco de Zurbarán, Francisco Goya, and Diego Velázquez," which can be seen by Balenciaga's use of lace and embroidery in many of his works. In particular, she displays a 1951 Balenciaga dress that is reminiscent of Francisco Goya's paintings. Palomo-Lovinski also describes the innovative shape of the 1951 dress: "The shape of this Balenciaga gown is innovative for the time, with the skirt being cut away at the front with a long back." This is representative of the type of style Balenciaga brought to couture (Palomo-Lovinski, 2010).

Balenciaga was also inspired by earlier couturiers including Madeline Vionnet, Coco Chanel, and Edward Molyneux. According to Palomo-Lovinski, Balenciaga admired Molyneux's slow development of technique over time, learned the importance of cutting and draping from Vionnet, and developed his emphasis on comfort from Chanel (Palomo-Lovinski, 2010). Although a complete analysis has not been conducted, preliminary studies on Chanel and Vionnet suggest that two were experimental. If this were true, this would be consistent with prior theory on creativity, for experimentalists tend to work with and learn from experimentalists and vice versa (Galenson, 2010).

Balenciaga's reserved, quiet nature and methodical process of creating is similar to that of other experimental artists. Like other experimentalists, Balenciaga relentlessly pursued perfection in his work and would often revisit designs years after they were completed. His shy personality allowed his work to speak for itself. Diana Vreeland of the Met describes how an "aura of mystery" developed over the course of his career as a result of his reserved disposition (National Gallery of Victoria). He rarely gave interviews and stayed out of the public life. When offered the opportunity to meet Picasso, Balenciaga apparently replied, "He is always wearing disguisethe man is a clown."

In true experimental fashion, Balenciaga's style grew and developed over the course of his career. Charlotte Seeling writes in her book *The Century of the Designer* that "from year to year his technique became more refined and his designs simpler" (Seeling, 2000). Claire Wilcox, curator of the exhibit on couture at the Victoria and Albert Museum in London, describes how Balenciaga

developed his skills over his career:

This supreme ability, comparable to perfect pitch in a musician, allowed him technical and aesthetic freedom, the importance of which has perhaps never been properly gauged, but which took him many years to achieve. In fact, if only his pre-war designs had remained, the designer would barely have stood out from his fellows (Wilcox, 2007).

That his designs only became great later in his career further supports the claim that he is experimental. This pattern of late development appears in his exhibitions as well. For the National Gallery of Victoria exhibition, the curator writes that “by looking at various basic suit types over a ten-year period, one can clearly see the ideals he was moving towards and the dominant features of his work” (National Gallery of Victoria). In the introduction to the 2011 Fine Arts Museums of San Francisco exhibition on Balenciaga, curator Hamish Bowles writes that Balenciaga’s career path was unique among designers of the time: “Indeed, in many ways he reversed the designer’s traditional career trajectory, producing some of his most thoughtful and even provocative designs in the twilight of his career” (Bowles & De Young, 2011). Between Wilcox’s and Bowles’ descriptions of the growth and development of Balenciaga’s skills and influence over the course of his career, it becomes clear that the reason for this late success is because he was an experimental designer.

Tables 4.1 through 4.3 each support this conclusion. Each table represents the distribution of the designer’s age and the number of items in exhibitions hosted by the Metropolitan Museum of Art, Fine Arts Museums of San Francisco, and National Gallery of Victoria. In the case of the Metropolitan’s exhibition (Table 4.1), 57.3% of his designs occur from the years 1960-1968, when Balenciaga was aged 65-73. In contrast, only 7.3% of the work came from the 1940s (ages 45-54). In Table 4.2, the San Francisco exhibition contained 45.2% of items from 1960s. In Table 4.3, the number is 58.3%. Table 4.4 shows the age distribution of Balenciaga’s work in the complete collections of the Fashion Institute of Technology Museum in New York and the Victoria and Albert Museum in London. The median ages of these collections are 58 and 65, respectively. Each of these profiles is consistent with that of an experimental designer. Because Balenciaga was known for his tailoring skills, these were most strongly developed at the end of his career.

The auction data further support this conclusion. His most expensive item sold was a 1960 (age 65) cape that sold in 2014 for a Christie's auction for over 57,000. *His next most expensive item was also from the 1960s and sold for over 50,000.* As prior studies have shown on creativity, the most expensive work is the most important, which corresponds with what the exhibition and museum collection data show: Balenciaga was most successful at the end of his career.

By the end of Balenciaga's career, his influence had spread across most of the couture world. The designers Balenciaga influenced directly include André Courrèges, Christian Dior, Emanuel Ungaro, Francisco Costa, and Hubert de Givenchy. Courrèges was "influenced by Balenciaga's structural and architectural forms" while Costa was influenced by Balenciaga's use of "shape, simplicity, and structure" (Palomo-Lovinski, 2010). Ungaro remembers the first time he saw Balenciaga: "When I saw Balenciaga for the first time, it was a discovery something so important for my life and my mind" (Healy, 1992). Balenciaga's commitment to simplicity and elegance continues to inspire fashion designers today. Rather than sacrificing his values and changing his style to meet the demand of consumers of the day, Balenciaga remained committed to couture and chose to retire rather than change. As Carmel Snow, editor of Harper's Bazaar, wrote upon his retirement, "Nothing is so mysterious as simplicity As always we may expect to see Balenciaga's influence sink deeply, noiselessly, until it pervades the whole world of fashion" (Palomo-Lovinski, 2010).

## 6.2 Madame Grès (1903-1993)

### Sphinx of Fashion

Madame Grès was a French experimental designer often unknown by the public but extremely well respected and admired by practitioners and critics. Over the course of her 50-year career, Grès inspired the many designers who followed her to pursue pure beauty and simplicity in their designs. Grès dedicated her career to perfecting her Grecian gown, which became the unequivocal model for beauty for the fashion world. Madame Grès was born to a bourgeois family in Paris as Germaine Krebs. She began her career as a sculptor but did not find success in the field. She changed her name to Alix Barton and opened her first design house in 1934 but was soon forced to close because of the War. In 1942, she opened another house named Grès where she achieved much of the fame

for which she is known today. Curator for the Costume Institute's exhibition on Madame Grès Richard Martin wrote, "we tender our highest esteem for this designer of exceptional vision, whose work is untrammelled by commerce or compromise" (Martin &Koda, 1994).

Madame Grès' innovation for the fashion industry was in creating free-flowing, beautiful evening gowns. Known for her strong, self-taught craftsmanship abilities, Madame Grès would often design her gowns directly on the wearer or mannequin. In experimental fashion, she would "drape dresses directly on her customers, and like Nina Ricci she cut straight out from the material without making a pattern first" (Seeling, 2000). The concept of creating the material directly on the wearer without designing prior to creating is an experimental practice. Experimental painters, for instance, often do not plan their work prior to painting (Galenson, 2010). Richard Martin describes Grès' face-to-face method of design as "a Rodinesque practice that exudes a feeling of tangible proximity to flesh and a sense of surety in rendering the body that are similar to that modern sculptor's works" (Martin &Koda, 1994). Curator of the Fashion Institute of Technology's exhibition on Madame Grès Patricia Mears also writes on Grès' hand-tailoring capabilities: "Her unique, self-taught methods of construction may lack mathematical precision, but their technical consistency and graceful aesthetics come undeniably from the hand of a master couturier" (Mears, 2007).

Like other experimental artists, Grès primarily found inspiration for her main work from her travels and life experiences. For example, Grès's 1935 "Pagoda" jacket represents a combination of Eastern and Cubist influence, while her sari dress embodies her travels 1958 travels to India and her "lifelong interest in dominoes, caftans, and other untailored constructions that affected the liquid softness of her 1970s world" (Martin & Koda, 1994). In addition to her travels Grès was inspired by other fashion designers, including Madeleine Vionnet and Balenciaga. According to Palomo-Lovinski, although Vionnet was more "mathematical" in her draping while Grès was more natural, Grès observed the difference between the two styles and formed her own innovations as a result of learning from other designers (Palomo-Lovinski, 2010).

Madame Grès' personality was also consistent with that of an experimental artist. Like Balenciaga, she was a mystery to the public: "Madame Grès is an enigma whose life was shrouded in mystery, and whose persona defied the typical characteristics of a couturier" (Palomo-Lovinski,

2010). She frequently changed her name and remained out of the public sphere for most of her career. Her exact day of birth is unknown and her death was kept hidden by her family for almost a year until after she died. Like Balenciaga, she let her work speak for itself and did not resort to marketing or commercialization to promote her designs. Professor of fashion Ann Kellogg writes, “A perfectionist, she was not concerned with the functionality of dress, rather in sculpting fabric into exquisite pieces of art” (Kellogg, 2002).

Grès’ most innovative decade was the 1970s, when she was between the ages of 67 and 75. Tables 5.1 and 5.2 support this conclusion. In the Metropolitan’s exhibition of Madame Grès, 31% of the designs come from this time span. The next highest decade by percentages is the 1960s, which included 25.7% of the designs. Again, this is consistent with experimental innovation. Because Madame Grès worked to perfecting her specific style of the Grecian, pleated gown, she had the most success with this later in her career. The most reliable data for Madame Grès comes from Table 5.2, the Musée Galliera exhibition, which had 119 items. The Met’s exhibition was relatively smaller (only 35 items) and the collections held by the Fashion Institute of Technology and Victoria and Albert Museum are not that large either (33 and 24 items, respectively) but still are useful in determining the most important time period for the designer.

The auction data in Table 5.4 also supports the conclusion that Madame Grès was most successful later in her career. The most expensive dress sold was from when she was 57 years old and sold for 9,020. *Thesecondmostexpensivewasfromwhenshewas62andsoldfor8,450.* While these numbers are not nearly as high as the other designers’, Madame Grès was still respected by critics and other designers as the critical literature demonstrates. The auction data also shows that of the 31 dresses observed, only 4 items came from before the age of 47. This is consistent with the fact that experimental designers produce their best work later in their careers.

Madame Grès has inspired many designers today. One of her most prominent followers is Doo-Ri Chung who creates a similar type of draping jersey dress shown on runways today. Contemporary Japanese designer Yohji Yamamoto also includes the pleated techniques in his designs. Costello Tagliapietra, Alber Elbaz, Isabel Toledo, Ralph Rucci, and Halston are other designers who have incorporated Grès’ free-flowing, graceful style into their collections.



Madame Grès embodies a pure experimental designer. As Richard Martin writes, “Grès invented one model that she practiced, polished, perfected, and purified.” Her self-taught skills and beautifully graceful dresses are innovations in the fashion industry that remain today. Her persistent devotion to her work makes her an experimental artist that has had a tremendous impact on the history of couture.

### **6.3 Yves Saint Laurent (1936-2008)**

“For Yves - and herein lies his uniqueness - each collection is a means of bringing dreams to life, expressing fantasies, encountering myths, and creating out of them a contemporary fashion.” - Pierre Berge, fellow designer and partner (Fine Arts Museum of San Francisco).

Yves Saint Laurent was a French conceptual designer. At the young age of 17, he won third place in the prestigious International Wool Secretariat competition for design sketches. The next year, at the age of 18, Saint Laurent won. The award is an international recognition given to outstanding fashion designers as judged by the top designers in the field and served as the launching point for Saint Laurent’s career. Saint Laurent was born in Algeria to a wealthy French family. His family exposed him to the theater and arts at an early age where he developed an interest in design (Kellogg, 2002). Saint Laurent went on to apprentice for Dior after winning the International Wool Secretariat prize, and at the age of 21 became the lead designer for Dior upon his sudden death in 1957. It is from this young age, 21, that Yves Saint Laurent began his career continuously redefining fashion for women.

Saint Laurent’s innovation was establishing the idea that women could be empowered by the clothing he designed. Saint Laurent had a specific idea behind his designs; his clothes served a purpose other than to make women look graceful, like Balenciaga and Madame Grès. His idea was that he wanted women to feel powerful wearing his clothes, and he designed accordingly. As he once stated in an interview, “My dream is to give women the basis of a classic wardrobe, which, escaping the fashion of the moment, will give them greater confidence in themselves” (Fine Arts Museum of San Francisco).

His creative process was entirely conceptual. Saint Laurent would sketch out his designs in

great detail prior to designing. Like conceptual architects, Saint Laurent could hand the designs to his team and step back from the work. This is possible for conceptual artists because the creation of the final product is not what is important. What is important is the idea behind it. This would not have worked for Balenciaga or Madame Grès because their technical skills and abilities were unmatched by fellow designers. One of Saint Laurent's contemporaries remembers the thousands of drawings Saint Laurent would produce and the high quality of each sketch: "When he began, Yves Saint Laurent would produce up to one thousand drawings but what makes Yves Saint Laurent's sketches of such amazing quality are their closeness to their final destination." Saint Laurent himself even acknowledged his sketching capabilities: "When I give the ateliers a sketch they can immediately recognize the direction of the fabric they can read it like a road map" (Fine Arts Museum of San Francisco). This occurs for conceptual designers because the concept for the design is complete before it is even drawn; it is an idea conceived of by Saint Laurent himself and finished before he writes it down.

Palomo-Lovinski highlights the years from 1966-1968 as the most important years of Yves Saint Laurent's career. In 1966 he created the Mondrian line. In 1967 he created *Le Smoking*, his tuxedos for women. Also in 1967, he created his African collection. In 1968 he introduced the safari line. Each of these collections was as independent as it was radical. The frequent change in style is common for conceptual innovators, as style is a means for expressing ideas. For Saint Laurent, the idea he wished to express was empowering women. His long-time partner and fellow designer Pierre Berge describes this goal:

Much like Dior gave women the feeling of beauty, YSL gave women power. He gave them masculine dress "sliding masculine shoulders onto those of women, dressing them in tuxedos, reefer jackets, blazers, sport coats, and trench coats" - Saint Laurent, in his own way, transferred power to women.

This pattern of drastic, sudden change continued for Saint Laurent. In 1977 he created the *Ballets Russes* collection inspired by the Russian ballet. Kellogg describes this collection as being a "dramatic departure" from his prior work. This work was conceptually innovative, as it was well defined, immediate, and spread quickly to other designers: "Shortly thereafter, ethnic fashions at

every price point proliferated on the fashion scene” (Fine Arts Museum of San Francisco).

When looking at Tables 5.1 and 5.2, Saint Laurent’s most important years become clear. To begin, the exhibitions for Balenciaga and Madame Grès featured almost no work from either designer before the age of 40. For Saint Laurent’s 2002 exhibition, over 35% of the work comes from before the age of 40. His 2010 exhibition contains nearly 40% from before the age of 40. What is most important in these tables though are the years from 1966 through 1971. During this 5-year time span, Saint Laurent created each of the individual collections described above. This 5-year span also contains the most designs in a concentrated period over his entire career.

The auction data from Table 5.3 also demonstrates this phenomenon. This creative time period started with the Mondrian dress in 1966. The Mondrian dresses were a direct link to Piet Mondrian’s geometric patterns of design. In an interview reprinted for his 2002 exhibition, the couturier states clearly in regard to the Mondrian collection: “I understood that until then, the world was rigid and that the time had come to make it move” (Fine Arts Museum of San Francisco). One of these dresses sold for \$46,980 in a 2011 Christie’s auction. The most expensive item out of the auction data was a 1967 dress with ostrich feathers he personally designed for model Danielle Luquet de Saint Germain. This year matches with the year with the most number of designs in his two exhibitions. The dress sold for \$154,375. The second most expensive item comes from his 1977 Ballets Russes collection. His Picasso-inspired Cubist dress from this year was the second highest selling item on the list. This also was a unique, radically different type of dress from his other work.

Saint Laurent brought haute couture to the common woman. The elegance and exclusivity of Balenciaga and Grès was no longer possible as both giants had retired, and Saint Laurent was there to bring in the new wave. One could not separate Saint Laurent’s clothes from Saint Laurent the person and his ideas: “For twenty-five years, Saint Laurent has fully exemplified Jean Cocteau’s phrase: In every landscape or still life, a painter always portrays himself” (Fine Arts Museum of San Francisco). Berge continues to say, he “extended the realm of aesthetics to embrace social issues, using in a certain way the approach of a moralist.” One does not find such strong ideas in the work of Balenciaga and Grès. This is the difference between conceptual and experimental innovation in design.

## 7 Conclusion

Experimental fashion designers make their best work late in their careers. This phenomenon occurs because experimental innovations are most often made as a result of strong tailoring and craftsmanship skills. These skills are honed and perfected over the course of a career. As a result, one's best designs occur later. These designs tend to be graceful and beautiful, often appearing simple when in fact they are quite complex. Cristóbal Balenciaga was an experimental designer most well known for his gowns while Madame Grès was an experimentalist who invented the pleated, Grecian jersey gown. Both designers' styles are seen on runways today. Conceptual fashion designers tend to make their best work earlier in their careers. Conceptual designs often embody an idea that is captured by the design and style of the clothing or fabric. Conceptual designers sketch their designs in advance, producing nearly complete sketches before the designs are prepared. Yves Saint Laurent was a conceptual designer who brought the idea of empowering women to couture. He is most well known for putting women in pant suits and creating dresses inspired by other artists' work like Mondrian and Picasso.

The difference between the two forms of creativity is consistent with the dichotomy that exists in other disciplines. Couture as a form of museum art is a relatively new phenomenon when compared with other forms of art. Yet over the last 40 years there has been a significant movement among museums collect designs that transformed the discipline. Within the discipline itself, further studies on creativity should explore the change in fashion design over the latter half of the 20th century to determine if the conceptual revolutions that occurred in painting and songwriting also occurred in fashion design. It is possible that the rise of ready-to-wear collections has made it easier for designers to impart their ideas on consumers. Regardless, the study of creativity is essential towards better understanding the history and nature of fashion design.

## References

- [1] Bowles, H., & M.H. De Young Memorial Museum. (2011). *Balenciaga and Spain*. San Francisco

- [2] Bowles, H., Mller, F., Foundation Pierre Bergé-Yves Saint Laurent., Montreal Museum of Fine Arts., & M.H. De Young Memorial Museum. (2008). *Yves Saint Laurent style*. New York; London: Abrams.
- [3] “Costume Institute.” Metropolitan Museum of Art, New York.
- [4] Font, Lourdes. “Dior Before Dior.” *West 86th: A Journal of Decorative Arts, Design History, and Material Culture*, Vol. 18, No. 1 (Spring-Summer 2011), pp. 26-49.
- [5] Fashion Institute of Technology Collection, New York.
- [6] “Costume Institute.” Metropolitan Museum of Art, New York.
- [7] Fine Arts Museums of San Francisco. (2002). *Yves Saint Laurent*. San Francisco.
- [8] Galenson, “From White Christmas’ to Sgt. Pepper,” *Historical Methods* (2009), pp. 17-32.
- [9] Galenson, “The Greatest Architects of the Twentieth-Century,” NBER Working Paper (2008).
- [10] Galenson, *Old Masters and Young Geniuses*, Chapter 2.
- [11] Galenson, “Understanding Creativity,” *Journal of Applied Economics* (2010).
- [12] Grès, A., Saillard, O., Lécallier, S., Cotta, L., Musée Galliéra., & Musée Bourdelle. (2011). *Madame Grès : la couture l’œuvre*. Paris: Paris Musées.
- [13] Healy, R., & National Gallery of Victoria. (1992). *Balenciaga : masterpieces of fashion design*. Melbourne: National Gallery of Victoria.
- [14] Kellogg, A. T. (2002). *In An Influential Fashion: An Encyclopedia Of Nineteenth-And Twentieth-Century Fashion Designers And Retailers Who Transformed Dress*. Westport, Conn.: Greenwood Press.
- [15] Martin, R., & Koda, H. (1996). *Christian Dior*. Metropolitan Museum of Art, New York.
- [16] Martin, R. & Koda, H. (1994). *Madame Grès*. Metropolitan Museum of Art, New York.

- [17] Martin, R., & Koda, H. "Haute Couture." Metropolitan Museum of Art, New York.
- [18] Mears, P. (2007). *Madame Grès: the sphinx of fashion*. New Haven Conn.: Yale University Press.
- [19] Menkes, Suzy. "Gone Global: Fashion as Art?" *The New York Times*. 04 July 2011.
- [20] Metropolitan Museum of Art Collection, New York.
- [21] Metropolitan Museum of Art. (1973). *The World of Balenciaga*. New York.
- [22] Musée des Beaux-Arts de la Ville de Paris. (2010). *Yves Saint Laurent*. Paris.
- [23] Musée Galliera - Musée Bourdelle. (2011). *Madame Grès, La Couture L'œuvre*. Paris.
- [24] National Gallery of Victoria. (1992). *Balenciaga: Masterpieces of Fashion Design*. Melbourne.
- [25] "New Look." *In The Thames & Hudson Dictionary Of Fashion and Fashion Designers*. London: Thames & Hudson, 2007.
- [26] Palomo-Lovinski, Noel. (2010). *The World's Most Influential Fashion Designers Hidden Connections and Lasting Legacies of Fashion's Iconic Creators* (pp. 192 p.).
- [27] Saint Laurent, Y., & Bergé, P. (2010). *Yves Saint Laurent Haute Couture: l'œuvre intégral 1962-2002*. Paris: éditions de la Martinière.
- [28] Saint Laurent, Y., Vreeland, D., & Costume Institute (New York N.Y.). (1983). *Yves Saint Laurent*. New York: Metropolitan Museum of Art : C.N. Potter.
- [29] Seeling, C., Morris, N., Morris, T., & Waloschek, K. (2000). *Fashion : The Century of the Designer, 1900-1999* (English ed.). Cologne, Germany.
- [30] Victoria and Albert Museum Collection. London, England.
- [31] Wilcox, Claire. (2007). *Golden Age of Couture: Paris and London 1947-1957*. London, England.

## 8 Appendix

Table 1: 20th Century Fashion Designers

	<b>Designer</b>	<b>Date</b>	<b>Country</b>	<b># of Items</b>	<b>Met Exhibition</b>
<b>Experimental</b>	Cristóbal Balenciaga	1895-1972	Spain	450	The World of Balenciaga
	Madame Grès	1903-1993	France	326	Madame Grès
<b>Conceptual</b>	Yves Saint Laurent	1936-2008	France	404	Yves Saint Laurent

Source: The Costume Institute, Metropolitan Museum of Art, New York, NY

Table 2: Top 10 Designers by # of Items in Metropolitan Museum of Art

<b>Designer</b>	<b># of Items</b>
Charles James	727
House of Dior	668
House of Balenciaga	497
Cristóbal Balenciaga	450
Yves Saint Laurent	404
Yves Saint Laurent, Paris	352
Madame Grès	326
Elsa Schiaparelli	317
House of Chanel	304
Christian Dior	282

Source: The Costume Institute, Metropolitan Museum of Art, New York, NY

Table 3: Monographic Exhibitions

<b>Selected Met Exhibitions</b>	<b>Date</b>
The World of Balenciaga	3/23/73 - 6/30/73
Yves Saint Laurent	12/14/83 - 9/2/84
Madame Grès	9/13/94 - 11/27/94
Christian Dior	12/12/96 - 3/23/97
Gianni Versace	12/11/97 - 3/22/98
Adrian: American Glamour	5/14/02 - 8/18/02
CHANEL	4/5/05 - 8/13/05
Paul Poiret: King of Fashion	5/7/07-8/6/07
Alexander McQueen	5/4/11-8/7/11
Charles James: Beyond Fashion	5/8/14-8/10/14

Source: The Costume Institute, Metropolitan Museum of Art, New York, NY

Table 4.1: The World of Balenciaga (1973)

Year	Age	# of Items (n=191)	% Distribution	% by Decade
1938	43	1	0.5%	1.6%
1939	44	2	1.0%	
1940	45	1	0.5%	7.3%
1941	46	0	0.0%	
1942	47	0	0.0%	
1943	48	0	0.0%	
1944	49	0	0.0%	
1945	50	2	1.0%	
1946	51	4	2.1%	
1947	52	3	1.6%	
1948	53	1	0.5%	
1949	54	3	1.6%	
1950	55	10	5.2%	33.3%
1951	56	3	1.6%	
1952	57	5	2.6%	
1953	58	5	2.6%	
1954	59	6	3.1%	
1955	60	9	4.7%	
1956	61	4	2.1%	
1957	62	12	6.3%	
1958	63	5	2.6%	
1959	64	5	2.6%	
1960	65	15	7.8%	57.3%
1961	66	11	5.7%	
1962	67	10	5.2%	
1963	68	15	7.8%	
1964	69	15	7.8%	
1965	70	17	8.9%	
1966	71	10	5.2%	
1967	72	11	5.7%	
1968	73	6	3.1%	
<b>31</b>		<b>191</b>		

Source: Metropolitan Museum of Art. (1973). The World of Balenciaga. New York



Table 4.2: Balenciaga and Spain (2011)

Year	Age	# of Items (n=191)	% Distribution	% by Decade
1938	43	1	0.7%	3.0%
1939	44	3	2.2%	
1940	45	2	1.5%	
1941	46	1	0.7%	12.6%
1942	47	0	0.0%	
1943	48	1	0.7%	
1944	49	0	0.0%	
1945	50	2	1.5%	
1946	51	2	1.5%	
1947	52	2	1.5%	
1948	53	6	4.4%	
1949	54	1	0.7%	
1950	55	8	5.9%	39.3%
1951	56	11	8.1%	
1952	57	5	3.7%	
1953	58	7	5.2%	
1954	59	6	4.4%	
1955	60	1	0.7%	
1956	61	3	2.2%	
1957	62	6	4.4%	
1958	63	4	3.0%	
1959	64	2	1.5%	
1960	65	6	4.4%	45.2%
1961	66	7	5.2%	
1962	67	8	5.9%	
1963	68	2	1.5%	
1964	69	11	8.1%	
1965	70	7	5.2%	
1966	71	6	4.4%	
1967	72	10	7.4%	
1968	73	4	3.0%	
<b>31</b>		<b>135</b>		

Source: Bowles, H., & M.H. De Young Memorial Museum. (2011). Balenciaga and Spain. San Francisco

Table 4.3: Balenciaga: Masterpieces of Fashion Design (1992)

Year	Age	# of Items (n=191)	% Distribution	% by Decade
1938	43	0	0.0%	
1939	44	1	1.4%	1.4%
1940	45	1	1.4%	
1941	46	2	2.8%	
1942	47	0	0.0%	
1943	48	0	0.0%	
1944	49	0	0.0%	
1945	50	0	0.0%	
1946	51	0	0.0%	
1947	52	0	0.0%	
1948	53	0	0.0%	
1949	54	0	0.0%	4.2%
1950	55	5	6.9%	
1951	56	3	4.2%	
1952	57	3	4.2%	
1953	58	2	2.8%	
1954	59	3	4.2%	
1955	60	2	2.8%	
1956	61	2	2.8%	
1957	62	3	4.2%	
1958	63	1	1.4%	
1959	64	2	2.8%	36.1%
1960	65	5	6.9%	
1961	66	4	5.6%	
1962	67	4	5.6%	
1963	68	3	4.2%	
1964	69	7	9.7%	
1965	70	2	2.8%	
1966	71	7	9.7%	
1967	72	9	12.5%	
1968	73	1	1.4%	58.3%
<b>31</b>		<b>72</b>		

Source: National Gallery of Victoria. (1992). Balenciaga: Masterpieces of Fashion Design. Melbourne

Table 4.4: Balenciaga Museum Collection Age Distributions

Age	Fashion Institute of Technology (n=19)	Victoria and Albert Museum (n=129)
30s	0	0
40s	2	1
50s	8	21
60s	7	76
70s	2	31
Mean age:	58.89	64.74
Median age:	58	65

Source: Fashion Institute of Technology Collection, New York. Victoria and Albert Museum Collection, London.

Table 4.5: Balenciaga Auction Prices and Age Distributions

Designer Age	Price (\$ USD)	Item (n = 26)	Year of Sale	Auction House
55	3,126	Dress	2009	Christie's
55	2,709	Dress	2009	Christie's
55	1,700	Dress	2012	Christie's
55	1,250	Suit	2009	Christie's
56	4,584	Dress	2009	Christie's
59	2,244	Dress	2009	Christie's
60	3,126	Dress	2009	Christie's
60	1,667	Ensamble	2009	Christie's
60	1,346	Dress	2008	Christie's
64	3,116	Dress	2010	Kerry Taylor
65	4,902	Suit	2007	Christie's
<b>65</b>	<b>57,464</b>	<b>Cape</b>	<b>2014</b>	<b>Christie's</b>
66	1,320	Dress	2010	Augusta
67	1,140	Ensamble	2010	Augusta
67	1,571	Ensamble	2008	Christie's
65	6,888	Dress	2013	Kerry Taylor
69	3,141	Suit	2008	Christie's
65-75*	11,745	Dress	2011	Christie's
65-75	5,418	Dress	2009	Christie's
65-75	5,418	Gown	2009	Christie's
65-75	4,039	Coat	2008	Christie's
65-75	3,320	Jacket	2008	Christie's
65-75	1,795	Dress	2008	Christie's
65-75	1,188	Suit	2007	Christie's
65-75	3,000	Dress	2011	Augusta
65-75	1,800	Dress	2012	Augusta

Source: <http://www.christies.com/results/>; [www.invaluable.com](http://www.invaluable.com); [www.augusta-auction.com](http://www.augusta-auction.com)

\*Ages listed "65-75" correspond with items listed as designed in "1960s" since this was the age range for Balenciaga during those years and exact date of clothing was not provided

Table 5.1: Madame Grès (1994)

Year	Age	# of Items (n=35)	% Distribution	% by Decade
1935	32	2	5.7%	
1937	34	1	2.9%	
1938	35	1	2.9%	11.4%
1946	43	1	2.9%	
1947	44	1	2.9%	5.7%
1950	47	1	2.9%	
1952	49	1	2.9%	
1954	51	1	2.9%	
1956	53	2	5.7%	14.3%
1961	58	1	2.9%	
1965	62	4	11.4%	
1968	65	2	5.7%	
1969	66	2	5.7%	25.7%
1970	67	3	8.6%	
1971	68	1	2.9%	
1974	71	2	5.7%	
1975	72	2	5.7%	
1976	73	1	2.9%	
1978	75	2	5.7%	31.4%
1980	77	3	8.6%	
1985	82	1	2.9%	11.4%
21		35		

Source: Metropolitan Museum of Art. (1994). Madame Grès. New York

Table 5.2: Madame Grès, La Couture à L'œuvre (2011)

Year	Age	# of Items (n=119)	% Distribution	% by Decade
1933	30	4	3.4%	
1934	31	1	0.8%	
1935	32	2	1.7%	
1936	33	1	0.8%	
1937	34	3	2.5%	
1938	35	3	2.5%	
1939	36	3	2.5%	14.3%
1940	37	1	0.8%	
1942	39	1	0.8%	
1943	40	1	0.8%	
1944	41	2	1.7%	
1945	42	4	3.4%	
1946	43	5	4.2%	
1947	44	3	2.5%	
1948	45	5	4.2%	
1949	46	5	4.2%	22.7%
1950	47	4	3.4%	
1951	48	4	3.4%	
1952	49	3	2.5%	
1953	50	2	1.7%	
1955	52	2	1.7%	
1956	53	2	1.7%	
1958	55	1	0.8%	15.1%
1960	57	2	1.7%	
1962	59	1	0.8%	
1963	60	1	0.8%	
1964	61	1	0.8%	
1965	62	2	1.7%	
1966	63	1	0.8%	
1967	64	1	0.8%	
1968	65	1	0.8%	
1969	66	2	1.7%	10.1%
1970	67	11	9.2%	
1971	68	3	2.5%	
1972	69	3	2.5%	
1973	70	1	0.8%	
1974	71	2	1.7%	
1975	72	4	3.4%	
1976	73	5	4.2%	
1977	74	5	4.2%	
1978	75	2	1.7%	
<b>1979</b>	<b>76</b>	<b>2</b>	<b>1.7%</b>	<b>31.9%</b>
1981	78	3	2.5%	
1982	79	1	0.8%	
1985	82	1	0.8%	
1986	83	1	0.8%	
1989	86	1	0.8%	5.9%
<b>47</b>		<b>119</b>		

Source: Musée Galliera - Musée Bourdelle. (2011). Madame Grès, La Couture à L'œuvre. Paris

Table 5.3: Madame Grès Museum Collection Age Distributions

Age	Fashion Institute of Technology (n=33)	Victoria and Albert Museum (n=24)
30s	9	1
40s	10	3
50s	1	5
60s	8	10
70s	5	5
Mean age:	52.24	61.76
Median age:	47	65

Source: Fashion Institute of Technology Collection, New York. Victoria and Albert Museum Collection, London.

Table 5.4: Madame Grès Auction Prices and Age Distributions

Designer Age	Price (\$ USD)	Item (n=31)	Year of Sale	Auction House
37	4,100	Dress	2013	Kerry Taylor
44	5,576	Dress	2013	Kerry Taylor
45	1,937	Dress	1998	Christie's
52	6,778	Dress	1998	Christie's
54	3,873	Dress	1998	Christie's
55	3,873	Dress	1998	Christie's
56	968	Dress	1998	Christie's
56	1,937	Dress	1998	Christie's
57	1,937	Dress	1998	Christie's
<b>57</b>	<b>9,020</b>	<b>Dress</b>	<b>2013</b>	<b>Kerry Taylor</b>
58	6,560	Dress	2013	Kerry Taylor
58	968	Dress	1998	Christie's
58	1,937	Dress	1998	Christie's
59	1,937	Dress	1998	Christie's
59	968	Dress	1998	Christie's
59	1,937	Dress	1998	Christie's
60	1,646	Dress	1998	Christie's
<b>62</b>	<b>8,450</b>	<b>Dress</b>	<b>2013</b>	<b>Gros-Delettrez</b>
64	3,300	Dress	2011	Augusta
72	2,905	Dress	1998	Christie's
81	1,937	Dress	1998	Christie's
81	1,937	Dress	1998	Christie's
37-47*	2,214	Dress	2013	Kerry Taylor
47-57	5,676	Dress	2007	Christie's
47-57	6,197	Dress	1998	Christie's
47-57	1,549	Dress	1998	Christie's
57-67	1,549	Dress	1998	Christie's
<b>57-67</b>	<b>8,134</b>	<b>Dress</b>	<b>1998</b>	<b>Christie's</b>
57-67	1,937	Dress	1998	Christie's
67-77	1,937	Dress	1998	Christie's
67-77	1,549	Dress	1998	Christie's

Source: <http://www.christies.com/results/>; [www.invaluable.com](http://www.invaluable.com); [www.augusta-auction.com](http://www.augusta-auction.com)

\*Ages listed "37-47" (etc.) correspond with items listed as designed in "1940s" (etc.) since this was the age range for Madame Grès during those years and exact date of clothing was not provided

Table 6.1: Yves Saint Laurent (2002)

Year	Age	# of Designs (n=160)	% Distribution	% by Decade
1958	22	1	0.6%	
1962	26	5	3.1%	
1963	27	1	0.6%	
1964	28	2	1.3%	
1965	29	5	3.1%	8.8%
1966	30	6	3.8%	
<b>1967</b>	<b>31</b>	<b>11</b>	<b>6.9%</b>	
<b>1968</b>	<b>32</b>	<b>5</b>	<b>3.1%</b>	
<b>1969</b>	<b>33</b>	<b>7</b>	<b>4.4%</b>	
<b>1970</b>	<b>34</b>	<b>6</b>	<b>3.8%</b>	
<b>1971</b>	<b>35</b>	<b>5</b>	<b>3.1%</b>	
1973	37	1	0.6%	
1975	39	1	0.6%	26.3%
1976	40	7	4.4%	
1977	41	7	4.4%	
1978	42	2	1.3%	
1979	43	6	3.8%	
1980	44	7	4.4%	
1981	45	5	3.1%	
1982	46	2	1.3%	
1983	47	3	1.9%	
1984	48	3	1.9%	
1985	49	1	0.6%	26.9%
1986	50	3	1.9%	
1987	51	2	1.3%	
1988	52	9	5.6%	
1989	53	3	1.9%	
1990	54	10	6.3%	
1991	55	6	3.8%	
1992	56	3	1.9%	
1993	57	3	1.9%	
1994	58	1	0.6%	
1995	59	2	1.3%	26.3%
1996	60	1	0.6%	
1997	61	4	2.5%	
1998	62	1	0.6%	
1999	63	4	2.5%	
2000	64	6	3.8%	
2001	65	3	1.9%	11.9%
<b>44</b>		<b>160</b>		

Source: Fine Arts Museums of San Francisco. (2002). Yves Saint Laurent. San Francisco

Table 6.2: Yves Saint Laurent (2010)

Year	Age	# of Designs (n=160)	% Distribution	% by Decade
1958	22	7	2.3%	
1959	23	0	0.0%	
1960	24	1	0.3%	
1961	25	0	0.0%	
1962	26	8	2.6%	
1963	27	1	0.3%	
1964	28	2	0.7%	
1965	29	3	1.0%	7.2%
1966	30	9	2.9%	
<b>1967</b>	<b>31</b>	<b>23</b>	<b>7.5%</b>	
<b>1968</b>	<b>32</b>	<b>13</b>	<b>4.2%</b>	
<b>1969</b>	<b>33</b>	<b>11</b>	<b>3.6%</b>	
<b>1970</b>	<b>34</b>	<b>11</b>	<b>3.6%</b>	
<b>1971</b>	<b>35</b>	<b>19</b>	<b>6.2%</b>	
1972	36	1	0.3%	
1973	37	2	0.7%	
1974	38	3	1.0%	
1975	39	5	1.6%	31.6%
1976	40	11	3.6%	
1977	41	17	5.5%	
1978	42	7	2.3%	
1979	43	11	3.6%	
1980	44	9	2.9%	
1981	45	10	3.3%	
1982	46	8	2.6%	
1983	47	10	3.3%	
1984	48	11	3.6%	
1985	49	6	2.0%	32.5%
1986	50	2	0.7%	
1987	51	2	0.7%	
1988	52	11	3.6%	
1989	53	3	1.0%	
1990	54	9	2.9%	
1991	55	5	1.6%	
1992	56	2	0.7%	
1993	57	1	0.3%	
1994	58	4	1.3%	
1995	59	6	2.0%	14.7%
1996	60	2	0.7%	
<b>39</b>		<b>307</b>		

Source: Fine Arts Museums of San Francisco. (2002). Yves Saint Laurent. San Francisco



Table 6.3: Yves Saint Laurent Auction Prices and Age Distributions

Designer	Age	Price (\$ USD)	Item (n=33)	Year of Sale	Auction House
	20	7,439	Gown (Dior)*	2011	Christie's
	21	4,698	Dress (Dior)	2011	Christie's
	22	3,900	Trapeze (Dior)	2013	Kerry Taylor
	23	7,800	Dress (Dior)	2012	Augusta
	23	4,800	Dress (Dior)	2010	Augusta
	23	4,500	Dress (Dior)	2011	Augusta
	24	4,698	Dress (Dior)	2012	Christie's
	26	3,132	Tunic	2011	Christie's
	29	3,320	Gown	2008	Christie's
	<b>30</b>	<b>46,980</b>	<b>Mondrian Dress</b>	<b>2011</b>	<b>Christie's</b>
	31	12,724	African Dress	2011	Christie's
	32	16,600	Safari suit	2008	Christie's
	32	4,698	Suit	2011	Christie's
	<b>32</b>	<b>154,375</b>	<b>Ostrich Dress</b>	<b>2013</b>	<b>Gros-Delettrez</b>
	34	3,141	Dress	2008	Christie's
	37	5,090	Gown	2011	Christie's
	41	2,640	Jacket and skirt	2011	Christie's
	42	19,500	Suit	2013	Gros-Delettrez
	43	12,350	Suit	2013	Gros-Delettrez
	<b>43</b>	<b>71,500</b>	<b>Picasso Dress</b>	<b>2013</b>	<b>Gros-Delettrez</b>
	44	4,980	Evening wrap	2008	Christie's
	46	780	Dress	2013	Gros-Delettrez
	48	1,300	Dress	2013	Gros-Delettrez
	49	1,680	Dress	2011	Augusta
	50	780	Dress	2013	Gros-Delettrez
	51	4,750	Suit	2011	Christie's
	51	4,000	Suit	2011	Christie's
	53	1,170	Dress	2013	Gros-Delettrez
	53	1,560	Dress	2013	Gros-Delettrez
	23-33**	27,500	Suit	2011	Christie's
	34-44	2,760	Dress	2009	Augusta
	45-55	40,000	Coat and Gown	2011	Christie's
	56-66	2,501	Jacket	2009	Christie's

Source: <http://www.christies.com/results/>; [www.invaluable.com](http://www.invaluable.com); [www.augusta-auction.com](http://www.augusta-auction.com)

\* Items with a (Dior) indicate that he designed for the House of Dior during these years after Christian Dior died

\*\*Ages listed "23-33" (etc.) correspond with items listed as designed in "1960s" (etc.) since this was the age range for Yves Saint Laurent during those years and exact date of clothing was not provided

# The Dangers of Letting Money Float

Jaewoo Jang

Stanford University

May 2015

## **Abstract**

The emergence of South Korea as one of four Asian tigers in the early 1990s was a facade of success based on an unsustainable and premature financial structure. The government's imposition of the financial liberalization policy induced an environment where domestic conglomerates could borrow loans and manipulate capital without extensive government scrutiny. This paper examines the ways in which the rapid liberalization of the financial market in Korea set the climate of collapse in the Korean economy in 1998. Then, this research observes how the large conglomerates took advantage of this sudden shift and eventually brought about the financial crisis of 1998.

## **1 Contextual Background**

Following the ceasefire of the Korean War in the mid-1960s, South Korea suffered high levels of unemployment and severe destitution (Yoo & Moon 269). However, with the advent of a more influential presidential leadership under Park and Jeon, the government of Korea sought to adopt a more export-oriented economic strategy to reduce its deficit terms of trade and generate national revenue by implementing the seven five-year economic plans in 1962 (Yoo et al., 272). The provision of this governmental policy include the deregulation of interest rates offered by domestic and foreign banks, the removal of restrictions on offering loans to domestic companies, consolidation of a more

managerial autonomy to wholesale banks, the reduction of entry barriers to financial activities and most importantly, the liberalization of cross-border capital account (Kwon, 331).

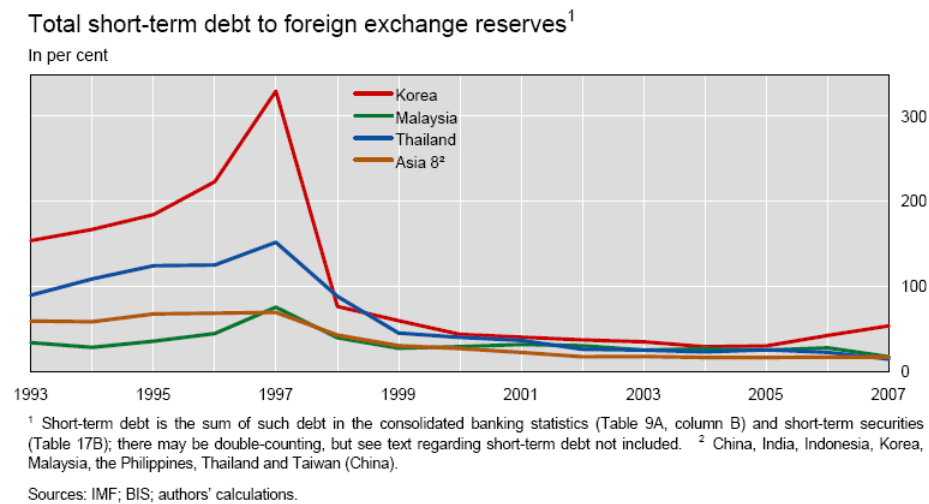
Many scholars assert that the hasty financial liberalization and dismantling of the traditional financial system in Korea had brought an imbalance in the country's current account and instigated a culture that lacks transparency within the financial market, eventually paving opportunities for overinvestment (Kim). But some emphasize that although financial liberalization in 1989 increased the firm's vulnerability in amassing more nonperforming debt, the corporate governance under the *chaebols* stands out as one of the primary factors in deteriorating the Korean economy by 1998, as *chaebols* themselves had handled the capital through unsuccessful ventures and investments (Yoo & Chul, 271). This paper shall assess the arguments that examine the two primary factors that plunged the South Korean economy into a financial crisis in 1998. Understanding these two arguments would allow economist and policy makers to determine the ways in which these main factors - financial liberalization and conglomerate business practices - have instigated the financial crisis in Korea.

To synthesize the myriad of views on this topic, this paper uses a set of sources that focus on either of the two disparate arguments. Amess, Demetriades, Kim and Kwon, who represent the views of IMF, World Bank or Korean Economic Research Institution respectively, approached this issue on a more macroeconomic perspective, viewing the issue at a more financial aspect. The title of Amess's work, for example, "Financial Liberalization and South Korean Financial Crisis" already lends itself into discussing financial liberalization as the prime focus. Since Amess and Demetriades both represent IMF and their works are highly focused on financial liberalization, their works provide us with the ways in which the international community perceived the causes of the Korean financial crisis. On the other hand, Powers, Chang, Noland, Yoo and Moon approached this issue at a more microeconomic perspective. Especially with Powers's research, her work was published in the Journal of East Asian Affairs, rendering her work to focus more on the internal affairs of Korea. Chang, Yoo and Moon, who are academic scholars in Korea's Business and Enterprise Association, have tailored their research on the business models and tactics which these conglomerates have used prior to the financial crisis. Due to their profession, Chang and Yoo, in particular, focus on

the business strategy that contributed to the collapse of Korea's financial structure. In a similar vein, the economic professor Noland shall provide us with the ways in which the Korea's business structure contributed to the collapse of Korea's economy. By blending both internal and external perspective to this issue, this paper addresses the question: How did *chaebols* and the liberalization of the financial market contribute to the South Korean financial crisis in 1998?

## **2 Hasty Imposition of the Financial Liberalization and Deregulation**

Since 1960s, the government has had a tight control on the banks and financial flows. But the government sought the need to liberalize the financial market to supply the capital needs of the *chaebols* in a more systematic and efficient manner (Kim, 2006). Though the Korean government in 1962 had imposed the seven "five-year economic plan" to liberalize the financial activity in the country, Kim Young Sam government hastened the "five-year economic plan" to "100-day economic plan" by removing government jurisdiction over private banks and establishing a number of 'merchant banks' or wholesale banks to cater short-term debts to the conglomerates in 1989 (Chang et al., 265). During the first half of the decade, the market has responded positively to this policy, marking a stable 2% increase per annum (Chang et al., 266). Liberalization of the financial market has far quickly attracted many foreign investors, generating an increase in capital inflow from \$20 billion dollars in 1990 to \$100 billion dollars by 1997 (Chang, et al. 738). Similarly, the policy induced growth, increasing the country's GDP per capita from \$67 in 1960s to \$8,500 worth of GNP in 1997 (Yoo et al., 271). Through financial liberalization, South Korea therefore demonstrated an optimistic projection of its economic discourse.



While Korea initially reaped economic growth through financial liberalization, Kwon proposes that this real GDP increase only created a faade of success, hiding the underlying flaw of the entire financial system in South Korea. He criticizes the government's lack of patience and assessment in evaluating the misleading World Bank statistical data of South Korea's overall debt/Gross National Product ratio by the eve of the crisis, which was at a low 22% as to the accepted benchmark of a low-risk debt/GNP ratio which is any figure below 48% (Kwon, 333). Offering an explanation to this misrepresentation of the data, Noland explains that the debt/Gross National Product ratio does not reflect the short-term debt ratio which has comprised more than 60.1% of Korea's total debt (See figure 1). Being unaware of this underlying high short-term debt ratio is problematic because the acquisition of large amount of short-term debt, which was financed excessively following the liberalization of the capital market, does not require the firms to submit their investment proposals to the Ministry of Finance and Economy (Kwon, 335). Following up on Kwon's and Noland's argument on high levels of debt, Chang and his team argue that liberalizing the financial market has not only paved way but also encouraged private firms to seek short-term loans to minimize their overhead costs of transactions and to take advantage of the deregulated market (Chang et.al, 736). With the absence of governmental regulation, total debt stock surged from \$7.27 billion to \$18.62 billion over the last 5 years and the real GDP growth that accompanied with those results have blinded the government from taking precautionary measures and assessing the mismatched

maturity of the loan structures (Chang et. al, 736-737). Thus, Chang concludes that, as private firms had undergone a perpetual cycle of satisfying debt-servicing charges with more short-term debt, the Korean economy has lost credit credibility by the time firms exhausted much of its capital in 1998 (Chang et. Al, 737).

### **3 Downfall of Korea's Credit-Rating**

The excessive borrowing of capital from the international market, which was permissible under the liberalized financial market, exposed the vulnerability of the Korean economy. In order to raise more capital of production for its conglomerates, the government aimed to ameliorate the nation's poor credit-rating by liberalizing its financial market (Chang et al., 738). Chang states that the influx of capital rose from \$78.4 billion dollars in 1990 to more than \$600 billion dollars by 1995, putting pressure on the domestic firms to enlarge and raise as much assets as possible (Chang et al., 738). Explaining the consequences of this policy, Kwon claims that this mounting pressure of debt has motivated the government to artificially appreciate the currency's value to generate more revenue when its products were exported (Kwon, 340). Thus, liberalized financial market accompanied by a high appreciated currency was a strategy to optimize export revenue (Kwon, 340).

However, as debt accumulated due to the firm's ability to unconditionally borrow capital without stringent government review, many Korean conglomerates were left insolvent, pulling down Korea's Standard and Poor's credit rating from an A+ to A- whereas the US treasury bill, for example, would receive an A+++ (Kwon, 351). Attesting to Kwon's observations of the decrease in Korea's credit-rating, Adelman and Nak both claim that this change had an unintended repercussion of the financial liberalization which has spurred the downfall of the Korean economy. Agreeing with Adelman and Nak's proposition, Amess and Demetriades also add that Moody's credit rating of Korea has likewise simultaneously dropped from A1 to A3 in 1997 (Amess & Demetriades, 226). Compiling all these statistical information together, Adelman and her colleague Nak make sense of these plunges in credit-rating as a pivotal event in casting a sense of insecurity in the international

market, making foreign investors withdraw money and create a net outflow of \$978 billion dollars by November 1997 (Adelman & Nak, 2). Such drastic withdrawal of capital, Adelman states, was also made possible due to the liberalization of the financial market as the government no longer imposed strict regulation in cross-border financial activities. This made the Korean won uncompetitive by depreciating its value from 87 in 1994 to 140 against the dollar (Adelman & Nak, 3). The price of Korean exports dropped as well as the firms' revenue and eventually, many firms were left bankrupt and requested a bailout from the government (Adelman & Nak, 3). Therefore, Adelman and Naks concur with Amess and Demetriades on the proposition that financial liberalization has exposed the fragility of Korea's economy to the international investors and let them withdraw the capital which they have initially invested.

### **3.1 *Chaebol's* Overdependence on the Government**

Though financial liberalization had provided the opportunities for exploitation of investment to take place, the conglomerate's overdependence of the government encouraged the excessive deficit and adoption of a poor investment strategy which led to the downfall of the Korean economy. The emergence of the *chaebols* has been a plan to gear the country into adopting a more export-oriented structure. Hence, in order to protect these exporting firms, the government cast numerous safety nets and set high barriers to entry as a mean to reduce competition (Powers, 106). Therefore, the proliferation of the *chaebols* are the repercussions of the highly regulated industrial market (Kim). Continuing onto Kwon's discussion on the origins of chaebol, Powers contends that the government's protection of the *chaebols* has spawned a culture of dependency. One of the most prevailing views about the *chaebols* was that they were considered to be the pillars of the Korean economy and that they were "too big to fail" forced the government to regard the *chaebols* with much respect (Powers, 106). This prevalent Moral-Hazard "too big to fail" argument paved way for a pathological corporate governance, rendering the assumption that *chaebols* as fragile as they must be protected at all cost. Kwon reaffirms Powers's observation by claiming that as part of the financial liberalization process, the Korean government imposed the "Foreign Capital Inducement Promotion Act" and "Law for Payment Guarantee of Foreign Borrowing" which ensured the eradication of foreign investment tax

on domestic firms and provided almost unconditional agreement to bailout the *chaebols* in times of bankruptcy (Kwon, 357).

Though Kwon emphasizes the accumulation of debt as the primary factor in the demise of the Korea economy, Yoo, Moon and to some extent Powers recognize that the primary factor which gave blind courage to the conglomerates to raise their debt-to-equity ratio to 338% was not merely due to the liberalization of financial market but also chaebol's bold over-reliance on government protection (Yoo & Moon 266). The government's approval to allow electronic manufacturer Samsung enter the highly-protected automobile industry in 1992 was one example of many risky ventures (Amess & Demetriades, 215). Chang also cites two examples that exposes the absurdity of the risky ventures: Hanbo entered the steel industry while the car manufacturing company Kia was granted permission to prospect mining sites around Korea in 1993. Hanbo, Samsung and Kia went bankrupt in 1998 (Chang, et al. 742), proving that much of its investments turned out to be too expansive and inefficient to compete against. Therefore, the faade of government protection has encouraged the increase in risky debt capital, which in turn imploded the nation's current account in 1998 (Yoo & Moon, 267).

### **3.2 *Chaebols's* Flawed Investment Strategy**

As *chaebols* worked under the impression that the government would always protect them, their investment strategies remained flawed and risky. Noland has championed the argument that the *chaebols's* primary goal to expand their business rather than maximizing profit, which placed businesses at a low profitability ratio of 2.8%, led them to bankruptcy and eventually the economic collapse of the country (Noland, 500). Powers evidences Noland's claim by discussing that *chaebols* blindly merged with many unrelated small-medium enterprises (SME) to diversify their investment portfolios (Powers, 112). The Halla Shipbuilding company for example owned clothing productions and pharmaceutical firms while Samsung electronics owned car manufacturing and insurance companies (Powers, 112). Though these vertical integrations may have contributed to minimizing transaction costs, Chang and his team clarify that the management of this wide array of firms is often ineffective (Chang et al., 743). Yoo, who is also part of KBEA but represents the financial



market side, likewise agrees with Chang's claim on *chaebols'* ineffective business strategy by asserting that outsourcing resources and components from other small-medium sized enterprise is less risky and more efficient (Yoo, 339). Therefore, the need to diversify their portfolios at the expense of minimizing profit dividend and pile up their debt displays signs of ostentatious ambitions (Yoo, 340). Both the financial department and the industrial department of KBEA) attribute the flaw in investment strategy as the main factor in the Korean financial crisis.

The *chaebols'* ambitions not only to diversify their investment but also to radically expand their homogenous businesses and increase their market share had placed the Korean economy at risk by 1998. The chemical and heavy industries have comprised many of the industrial sectors that the government had chosen to develop in 1960 (Chang et al., 735). These securities assets are homogenous and often brought back low returns (Change et al., 736). Yoo, Park and Chang claim that the supply of these export goods increased with the absence of any increased demand in the international market (Yoo et al., 265). Powers notices that the repercussion of having increased production with no increase in real demand was the decrease in export-revenue by a stiff 45% margin by the end of 1998 (Powers, 112). Therefore, Noland wraps up the discussion on the causes of Korean economy downturn in 1998 by noting that though financial liberalization had exposed the opportunities with which private firms can exploit the financial market, it was the actual mismanagement of these debt capitals that deprived the nation of its financial capital and credibility (Noland, 506). As *chaebols* had already expanded so much of their market exposure in the domestic market, the six chaebol bankruptcy by November 1997 has caused a negative rippling effect upon the economy greatly (Noland, 507).

## 4 Conclusion

The liberalization of the financial sectors, the mismanagement and the overconfidence of the chaebols have respectively played a significant role in bringing about the Korean financial crisis in 1998. This case study brings a range of sources dangers of hastily bringing about the liberalization of the economy which the western world or any other capitalist nation presume it to be as the most effi-

cient. One group of sources contend that President Kim's decision to liberalize the financial market and dismantle the managed economic system 10 years earlier to its intended date of implementation has brought about unforeseen consequences. Hastening the process raised unemployment rates, paralyzed the market structure and deteriorated the real GDP growth. While the other set of sources discusses the actual corporate malfeasance as the primary causal factor in Korea's financial crisis. Therefore, the financial crisis of 1998 demonstrates an example where hasty implementation of a new financial system could ultimately bring down a country's economy if the nation is unprepared to embrace the change.

## References

- [1] Amess, Kevin and Panicos Demetriades. "Financial Liberalisation and the South Korean Financial Crisis: An Analysis of Expert Opinion." *The World Economy* 33 (2).
- [2] Chang, Ha-Joon, Hong-Jae Park and Chul Gyue Yoo. "Interpreting the Korean Crisis: financial liberalization, industrial policy and corporate finance" *Cambridge Journal of Economics* 22. London: University of Cambridge, 1998. 735-746.
- [3] Kihwan, Kim. *The 1997-98 Korean Financial Crisis: Causes, Policy Response, and Lessons: International Monetary Fund & The Government of Singapore*. July 10, 2006.
- [4] Kwon, Yul O. "The Korean Financial Crisis: Diagnosis, remedies and prospects." *Journal of the Asia Pacific Economy* 3 (3). Routledge: Taylor & Francis Group, 1998. 331-357.
- [5] Noland, Marcus. "South Korea's Experience with International Capital Flows." *Capital Controls and Capital Flows in Emerging*. Ed. Sebastian Edwards. National Bureau of ' Economic Research: University of Chicago Press, 2007. 481-509.
- [6] Powers, Charlotte Marguerite. "The Changing Role of Chaebol." *Stanford Journal of East Asian Affairs* 10 (2). Ed, Nina Chung. Stanford University: 2012. 105-116. Print.

- [7] Yoo, Jang-Hee, and Chul Woo Moon. "Korean Financial Crisis During 1997-1998 Causes And Challenges." *Journal Of Asian Economics* 10.2 (1999): 263-281. Business Source Complete. Web. 22 Oct. 2014.

# Afterlife Beliefs, Work Ethic, and Economic Outcomes

Nicholas Cornell

Rice University

May 2015

## **Abstract**

This paper expands on the body of literature relating spiritual or religious beliefs to economic outcomes. It focuses on a hypothetical mechanism whereby cultural values are produced by religious beliefs and these cultural values, in turn, produce economic outcomes. World Values Survey data is leveraged to provide evidence for this mechanism using the proportion of a countrys population that believes in heaven or hell and a measure of the countrys work ethic.

## **1 Introduction**

McCleary and Barro's 2006 paper conducted a survey of quantitative treatments of religiosity and economic outcomes. They systematically looked at religiosity as a dependent and independent variable, determinants of various religious beliefs and participation, and how beliefs and attendance impact economic performance. When looking at economic growth, they find robust support for the impact of belief in hell, monthly attendance, and shares of various religions in a country on economic growth, namely a positive effect on belief in hell and a negative effect on attendance, though they emphasize that the relation of beliefs to attendance is important (McCleary & Barro, 2006, pp. 67-68). McCleary and Barro suggest that this supports a Weberian view of religion and economics; that is, religion produces beliefs in individuals that impact their attitudes and values which, in turn, impact economic outcomes (Religion and Economy, 2006, p. 68).

Taking this analysis further, McCleary and Barro try to determine the drivers of specific attitudes or traits by examining how work ethic, honesty, and thrift are impacted by a belief in hell, heaven, or an afterlife where all variables are expressed as a proportion of the population. They find the strongest result for work ethic regressed on belief in hell, reported in Equation 1 with standard errors in parentheses (McCleary & Barro, 2006, p. 70).

**Equation 1 – McCleary and Barro (2006, 70) regression for work ethic on belief in hell**

$$\begin{array}{lclclcl}
 \text{work ethic} = & 1.20 & + 0.200 & (\text{belief in} & -0.091 & (\text{log per capita} & +0.307 & (\text{ex-communist}) \\
 & & & \text{hell}) & \text{GDP}) & & & \\
 & (0.28) & (0.0098) & & (0.027) & & (0.052) & \\
 R^2=0.55 & & & & & & & 
 \end{array}$$

They find slightly weaker results when other beliefs are substituted in or when a different cultural trait is looked at. The positive and statistically significant (at the 5% level) coefficient on belief in hell indicates that an increased proportion of the population that believes in hell drives, or is at least correlated with, an increased proportion of the population that thinks hard work is important to teach children, a proxy for work ethic. McCleary and Barro interpret this as evidence for Webers suggestion that religion, working through beliefs, impacts work ethic (Religion and Economy, 2006, pp. 70-71). This can, intuitively, impact economic outcomes.

Taking a step back from their results, I found it curious that only belief in hell was worth including in the model. A belief in an afterlife, generally, can mean many different things, including belief in heaven and hell, just one of the two, or some other notion of what happens when one passes that cannot be readily surveyed. The fact that this variable was not indicative is not surprising as it is too high-level and vague to capture any meaningful, specific belief. On the other hand, specific beliefs in hell or heaven are related, but neither perfectly correlated nor mutually exclusive. That is, someone can believe in heaven and not hell, hell and not heaven, both, or neither. Since hell is generally seen as a punishment while heaven is seen as a reward, I would expect people to respond based on their perception of both incentives and not just one of them.

McCleary and Barro also leave open the question of whether work ethic, especially the portion of work ethic attributable to beliefs, impact economic outcomes (Religion and Economy, 2006).

This intuitive importance of the combination of belief in hell and belief in heaven as a determinant of work ethic can be confirmed through analysis of World Values Survey data. The same analyses also suggest a different direction for the effect of belief in hell than McCleary and Barro found. I find support for the hypothesis that work ethic, specifically the portion due to an individuals beliefs in heaven and hell, determines economic growth.

## 2 Review of Cultural Explanations for Economic Outcomes

Cultural explanations for economic outcomes date all the way back to Adam Smith and John Stuart Mill, who considered cultural explanations to sometimes be stronger than those based on individual rationality (Guiso, Sapienza, & Zingales, *Does Culture Affect Economic Outcomes?*, 2006, p. 26). Wesley, the founder of the Methodist Church, demonstrated this perspective by urging congregants to accrue wealth, save, and give (McCleary & Barro, 2006, p. 51).

Other economists saw problems with this direction of causality. Namely, Marx suggested that the means of production influences the social and political structure of a society, in a word, its culture (Guiso, Sapienza, & Zingales, *Does Culture Affect Economic Outcomes?*, 2006, p. 26). Some, like Marx, argue that humans, through the way they organize, create religion (Guiso, Sapienza, & Zingales, *People's opium? Religion and economic attitudes*, 2003, p. 230). This explanation implies that religion conforms to the cultural context it is given and does not play a role in its creation. Given this view, religious explanations, or even cultural explanations, for economic outcomes are rooted in the fundamental organization of society, which already directly influences economics outcomes.

The basis for todays work on culture and economics, though, is Max Webers seminal work, *The Protestant Ethic and the Spirit of Capitalism*. Webers core thesis is the Protestant Reformation laid the groundwork for the rise of capitalism in Europe, meaning that religion lead directly to economic outcomes. A key teaching during the Protestant Reformation was that the accumulation of wealth was a respectable duty imposed on all followers (Guiso, Sapienza, & Zingales, *Does Culture Affect Economic Outcomes?*, 2006, p. 26). When a sufficient portion of the population holds this view, it

has the power to legitimize a change in economic order (Pryor, 2005, p. 451). Eisenstadt gave this phenomenon the term “transformative potential” to describe the capacity of religions to legitimize new institutions or activities (Guiso, Sapienza, & Zingales, *People’s opium? Religion and economic attitudes*, 2003, p. 229).

For quantitative analyses of cultural explanations for economic outcomes, a theorized mechanism is needed. Religion can produce some type of good, like religious beliefs. Alternatively, religion can have a network effect. That is, formal services or other organized events connect religious followers, and these connections lead to economic growth or other outcomes. For Weber, religion produces beliefs. These beliefs, including work ethic, trust, and thriftiness, are what matter in determining economic outcomes (McCleary & Barro, 2006, p. 51).

Following Weber, though, economists largely ignored cultural explanation for economic outcomes, especially religious explanations (Noland, 2005, p. 1216). As institutional analyses started to grow in the late twentieth century, culture received more attention (Guiso, Sapienza, & Zingales, *Does Culture Affect Economic Outcomes?*, 2006, p. 28). Landes, specifically, stresses the mechanism where culture influences beliefs and values, which influence economic outcomes (The Wealthy and Poverty of Nations, 1999). Values like hard work, tenacity, and honesty were linked to economic development. These cultural traits are notable in that fit well in a Weberian framework, a fact which Landes acknowledges (Landes, 1999, p. 516). McCleary and Barro analyze religiosity and economic growth and conclude that a combination of religiosity (beliefs relative to participation) and religious types (denominations) are important for explaining economic growth (Religion and Economy, 2006, p. 69).

This intellectual tradition of cultural explanations, focusing on cultural values, for economic outcomes has specific examples in the literature. Trust is a popular variable to examine, likely due in part to its easy connection to economic outcomes. Trust is a substitute for complete contracts. If two parties to an arrangement can trust the other to act in good faith, a complete contract is not needed. This lowers the cost of transactions and encourages the proliferation of economically beneficial arrangements. Guiso, Sapienza and Zingales find specific evidence of religions impact on trust; being raised religious increases a constructed trust index by 2% while regularly attending

service raises it by 20% (*Does Culture Affect Economic Outcomes?*, 2006, p. 30). In turn, several studies have found trust to be an important explanatory variable for economic growth (Knack & Keefer, *Does Social Capital Have an Economic Payoff? A Cross-Country Investigation*, 1997; Knack, Trust, associational life, and economic performance, 2001; Zak & Knack, 2001).

Another common cultural value that is studied is thriftiness. The hypothesized mechanism by which thriftiness impacts economic outcomes is through individual saving, which impacts investment and thus growth. Guiso, Sapienza and Zingales find a link between thriftiness as a cultural value and the national savings rate; a 10% increase in the proportion of the population that thinks thriftiness is an important value to teach children is related to a 1.3% increase in the national savings rate (*Does Culture Affect Economic Outcomes?*, 2006, p. 38). More importantly, they find evidence to suggest that the explanatory power of this cultural value is comparable to that of the life-cycle model. Relevant to the Weberian hypothesis, Guiso, Sapienza, and Zingales ascribe the cultural value of thriftiness to religion. They found significant, positive coefficients on Catholicism and Protestantism related to the likelihood of teaching thriftiness as an important value to children (Guiso, Sapienza, & Zingales, *Does Culture Affect Economic Outcomes?*, 2006, p. 38).

While thrift and trust have been studied extensively, work ethic has been looked at in less depth. Formerly, Protestant Work Ethic could be defined as “a dispositional variable characterized by a belief in the importance of hard work, rationality, and frugality which acts as a defense against sloth, sensuality, and religious doubt” (McMurray & Scott, 2013, p. 656). Today, a more generalized conception that severs the religious undertones can be used to define work ethic. When looking at worth ethic, literature has noted that the GNP of an individuals country of birth strongly indicates their work ethic, even after residing in another country for some time (McMurray & Scott, 2013, pp. 661-662). This could underscore a tie between economic performance and work values, but I find it more likely given the literature discussed above that GNP in an individuals birth country is a proxy for the culture the individual was socialized in and thus subscribes to. That is, there is a cultural explanation for the variation in work ethic across individuals. Examinations of a specific work ethic construct, Islamic Work Ethic (IWE), have found a positive relationship on the individual level between IWE and organizational commitment (Ali & Al-Owaihan, 2008, p. 15). This intuitively



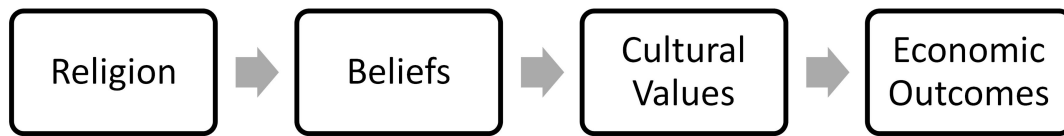
connects to economic outcomes, though an empirical link is not available. McCleary and Barro provide the strongest support that religious beliefs impact work ethic as a cultural value, but they do not tie this cultural value to economic outcomes (Religion and Economy, 2006, p. 70).

Yet other studies have criticized the tie between religion, work ethic as a cultural value, and economic outcomes. Blum and Dudley analyze Webers work and conclude that the work ethic taught by the Protestant Reformation was not responsible for the rise of capitalism. Instead, they argue for the importance of information networks that were enabled and spread by the Protestant Reformation (*Religion and economic growth: was Weber right?*, 2001, pp. 210, 229).

While the specifics may be unclear, the history of cultural explanations for economic outcomes is complemented by studies that conclude that religion or cultural explanations matter. Noland finds that the hypothesis that religious affiliation is uncorrelated with economic performance can often be rejected, although he does not find a clear result for any particular denomination driving this (Noland, 2005). Guiso, Sapienza, and Zingales, after a review of the literature, conclude that, on average, religion has a positive effect on the development of attitudes conducive to economic growth (*People's opium? Religion and economic attitudes*, 2003). Again, they find inconsistent evidence as to any denomination behind this relationship. On the question of whether religion drives economic performance, Weber may have been correct.

### 3 Hypothesized Mechanism

The reason studies find support for cultural, specifically religious, explanations of economic outcomes but not for any denominational effects may have to do with a misunderstanding of the mechanism by which religion affects economic growth. Following Weber, I argue that the chain of causality can be summarized as in Figure 1.



**Figure 1 – Theoretical Mechanism**

**Figure 1 - Theoretical Mechanism**

When studies look at religious denominations, they are skipping one or both of the intermediate steps by which religion impacts economic outcomes. There may also be more clear results if the definition of religion is expanded. In the United States, 18% of the population identifies as spiritual but not religious (Pew Research Center, 2012). That is, they do not identify with an established religious tradition or institution, but they may still believe in God, believe in an afterlife, and hold what have traditionally been thought of as religious beliefs. Scholars argue that religious institutions are in constant competition in today's society with non-religious institutions for the provision of social, cultural, or welfare goods that used to be solely provided by religious institutions (Hirschle, 2013, p. 414). In this framework, individuals attend and practice religion less. However, there need not be an immediate or long-term effect on religious beliefs (Hirschle, 2013, p. 421).

Given this perspective, I find it more valuable to look at spiritual beliefs. This raises the question of how these beliefs are produced as a Weberian argument, where religion produces beliefs, is unsubstantial. There may, in fact, be multiple pathways that produce spiritual beliefs, including casual interactions with established denominations or the way one is brought up, even if that never involved formal religious participation. Spiritual beliefs are beliefs in the supernatural that need not be related to any denomination; they include a belief in hell, heaven, or any other notion of an afterlife.

The theoretical connection between spiritual beliefs and cultural values is given by Azzi and Ehrenberg (Household Allocation of Time and Church Attendance, 1975). Azzi and Ehrenberg describe a rational choice model in which households maximize their utility between their time on

earth and their time after. In such a framework, seemingly irrational attitudes can be taken as perfectly rational. Consider the hypothetical case where an individual is told by a religious institution they have to take a mission trip, from which they derive no value, and pay the opportunity cost of schooling or lost wages that time represents. If the individual has no spiritual beliefs, they would never go on the mission trip. However, if the individual has beliefs that indicate a mission trip must be conducted for entrance into a good eternal state after death, then it becomes rational to undertake the mission trip as the opportunity cost is likely less than the benefit of a good eternal state. If the mission trip in this example is thought of as service, frugality, or hard work, then the connection, underpinned by rational choice, between spiritual beliefs and cultural attitudes becomes more apparent.

The connection between cultural values and economic outcomes is intuitive. Good cultural values, like thriftiness or hard work, when aggregated lead to observable economic outcomes like saving or productivity, respectively.

My data analysis focuses on the connection between spiritual beliefs and work ethic as a cultural value and the connection between work ethic as a cultural value and economic growth, tying both links in the chain to the entire mechanism whereby spiritual beliefs affect economic growth through work ethic.

## 4 Data Analysis

### 4.1 Data Sources

McCleary and Barro, for the regression presented in Equation 1, utilize the 1994-1999 and 1999-2004 waves of the World Values Survey. I began my analysis by trying to replicate their result, but on a wider range of data. My work used all waves of the World Values Survey readily available at the time of this analysis. These waves covered 1981-1984, 1989-1993, 1994-1999, 1999-2004, and 2005-2007. The data for this last wave was the first half of all the data that is now available for 2005-2009, meaning that the dataset for this time period is smaller than for other waves.

The World Values Survey samples representative samples from many countries worldwide. A

qualitative analysis of the surveys indicates that most responses were collected in the second year of the wave, though some responses may be from any time period within the years the wave covers. The dataset I used had 257,597 individual responses covering 1079 variables.

I looked at three major variables: belief in hell, belief in heaven, and a variable meant to reflect work ethic. Belief in hell was coded to be 1 if the respondent believed in hell and 0 otherwise. Belief in heaven was coded to be 1 if the respondent believed in heaven and 0 otherwise. Respondents who gave “dont know,” “no answer,” or “not applicable” to either of these survey questions were dropped from the analysis, as were those missing answers for these two variables. I then determined which individuals believed in heaven and not hell and coded a new variable as 1 if this was true. If this variable was 1, then belief in heaven was coded 0 for that observation to avoid double counting individuals.

To proxy work ethic in a country, I followed the lead of McCleary and Barro (Religion and Economy, 2006). Respondents were given a list of qualities that children can be encouraged to learn at home and asked, “Which, if any, do you consider to be especially important? Please choose up to five.” One of these qualities was “hard work.” If a respondent considered hard work to be a top five trait, out of 12 traits provided, the response was coded 1. If the respondent answered the question but did not select hard work as a top 5 trait, the response was coded 0. Henceforth, this will be referred to as work ethic, following the approach of McCleary and Barro in accepting this as a good proxy of an underlying cultural value (Religion and Economy, 2006, p. 70). As before, all other observations were dropped. After validating the data for meaningful answers to the questions of interest, I was left with 140,908 observations across 67 countries.

McCleary and Barro make their dataset available from Robert Barros website (Barro, 2003). This dataset was used primarily to determine which countries are or formerly were Communist. Any country that was Communist for a period since 1925 was marked as ex-communist for my analysis. Economic variables at the country level were acquired from The World Banks “World Development Indicators” database. GDP per capita data was gathered for the first two years of each wave, averaged together, and then logged for use in analysis. GDP per capita growth, as a percentage of constant 2005 US dollars (or real GDP per capita growth), was gathered for a

span of three years, starting with the last year of each survey wave, and averaged together. The intent was that this would be somewhat lagged from the bulk of the survey data, reducing concerns about endogeneity. Fertility data and gross capital formation data was also gathered from the same database for the first year of each survey wave. Life expectancy at birth was collected in the same way and inverted for use in my analysis, per McCleary and Barro (Religion and Economy, 2006).

## 4.2 Combinations of Beliefs and Work Ethic on the Individual Level

I began by looking at the data on the individual response level. I used a logistic regression with the work ethic indicator as the dependent variable. The results are reported in Equation 2.

### Equation 2 – Logistic regression for work ethic using individual responses

$$work\ ethic = 0.1793 + 0.4757\ (belief\ in\ hell) - 0.4045\ (belief\ in\ heaven)$$

$$(0.0099)^{***} \quad (0.0164)^{***} \quad (0.0177)^{***}$$

\* indicates significance at 10% level; \*\* significant at 5%; \*\*\* significant at 1%

Equation 2 indicates, with the significant coefficients on beliefs, that both belief in hell and belief in heaven, separately, are important in determining an individuals likelihood of holding work ethic as an important value. Since this is a logistic regression, the coefficients are additive and can be interpreted as log odds. This means that they can be converted to probabilities with a simple transformation. These probabilities describe the portion of the dataset that thinks work ethic is important, conditioned on each of the four combinations of beliefs. These proportions are summarized in Table 1 along with the proportion of the dataset that holds that combinations of beliefs.

**Table 1 - Combinations of beliefs and probability work ethic is important**

<b>Combination of beliefs</b>	<b>Probability of thinking work ethic is important</b>	<b>Proportion of dataset with combination of beliefs (rounded to nearest integer)</b>
Only hell	0.6581	<1%
Only heaven	0.4439	12%
Heaven and hell	0.5623	58%
Neither heaven nor hell	0.5447	29%

One observation from Table 1 is that there are very few, less than 1000 people in my sample, that believe in hell and not heaven. This is as small as to be treated as 0. That is, the beliefs and resulting values of this population likely have no bearing on macroeconomic outcomes and are generally too small, when spread across countries and time, to yield useful insights. The key insight of Table 1, though, is a nontrivial proportion of people believe only in heaven. This group is worth noting because they have radically different views on work ethic than other combinations of beliefs. There is little difference in the importance of work ethic between those that believe in heaven and hell and those that believe in neither while there is a precipitous drop from about 55% to about 44% in the proportion of people who think work ethic is important among the population that believes only in heaven.

This phenomenon makes sense when viewed through a rational choice framework. If individuals make choices that maximize their utility on earth and after earth, then those that believe they already have maximized their after-life utility, that is they think the only place to go is heaven, turn to maximizing their utility on earth. This likely involves less hard work than an individual who thinks hard work, while it may take away from utility in the present, will improve their chances of going to heaven. The flaw with this explanation is that it does not explain the slightly lower work ethic among those that believe in neither heaven nor hell. However, if we take this non-belief to suggest they are neither religious nor spiritual, then it may make sense for this group to work hard on earth. Work generally allows for consumption. If an individual is not spiritual or religious, it is

possible they may value consumption more than an individual who is, since religion can teach and incentive moderation. In order to obtain this consumption, they need to be productive and work. However, this discussion concerns itself with the sources of beliefs, a topic beyond this paper.

Given the insights of my logistic regression, I did not find it unimaginable that variables for specific combinations of beliefs could be more useful in explaining work ethic as a cultural value than just belief in hell.

### **4.3 Mimicking McCleary and Barro**

Putting this insight about the combination of beliefs aside for a moment, I was interested in whether the results McCleary and Barro arrived at for the association between belief in hell and work ethic could be repeated on a larger dataset. I followed their approach in aggregating the proportion of respondents that held a certain belief or felt work ethic was important and grouping the results by country and then by survey wave. I then ran a regression of work ethic on the same explanatory variables they used, though over more survey years. The result is presented in model 1 in Table 2.

**Table 2 Work ethic modeled by various beliefs**

	Reflecting Mc-Cleary and Barro (2006, 70)	Adding in belief in only heaven	Adding in survey waves as fixed effects	Removal of ex-communist
<b>belief in hell</b>	-0.0587 (0.0842)	-0.1358 (0.0828)	-0.1873 (0.0875)**	-0.1917 (0.0843)**
<b>belief in only heaven</b>		-0.6695 (0.1874)***	-0.6422 (0.1899)***	-0.6559 (0.1759)***
<b>ex-communist</b>	0.0649 (0.0430)	0.0098 (0.0436)	0.0085 (0.0428)	
<b>log of GDP per capita</b>	-0.0989 (0.0163)***	-0.0909 (0.0156)***	-0.0889 (0.0153)***	-0.08936 (0.0151)***
<b>wave 1989-1993</b>			0.2331 (0.0870)***	0.2343 (0.0864)***
<b>wave 1994-1999</b>			0.1991 (0.0748)***	0.1992 (0.0744)***
<b>wave 1999-2004</b>			0.2498 (0.0804)***	0.2495 (0.0800)***
<b>wave 2005-2007</b>			0.2170 (0.1110)***	0.2151
<b>constant</b>	1.3635 (0.1684)***	1.4382 (0.1610)***	1.2397 (0.01744)***	1.2504 (0.1649)***
$R^2$	0.3298	0.4025	0.4577	0.4575
<b>Adjusted <math>R^2</math></b>	0.3109	0.3797	0.4148	0.4203

\* indicates significance at 10% level; \*\* significant at 5%; \*\*\* significant at 1% Omitted wave is

1981-1984. All numbers are rounded to four decimal places. All had 110 observations. Observations are at



the country level. Work ethic is proportion of population that thinks hard work is an important value to teach children. Beliefs are as percentage of a countrys population.

Model 1 in Table 2 presents a favorable adjusted R2 and the model is significant at the 1% level. However, my results disagree with McCleary and Barros displayed in Equation 1. The reason for these substantially different results could be numerous. It is possible that McCleary and Barro used other control variables that I did not or they coded variables in a slightly different way. For example, my coefficient on ex-communist reaches fairly far back. So if a country had a communist regime in 1925, the indicator would still be true decades later. I might have also used a different temporal relationship between the survey waves and the GDP per capita data than McCleary and Barro. However, I have general confidence in my model specification because the coefficient on GDP per capita is nearly the same as McCleary and Barros: -0.0989 versus -0.091.

The most important disagreement is in the direction, and significance, of the coefficient on belief in hell. My results suggest that there is no discernible effect of the proportion of the population in a country that believes in hell and the proportion of the country that holds work ethic as an important cultural value. This fails to support the theoretical mechanism outlined above. However, the coefficient is insignificant and thus I find no need for further explanation as to its sign.

#### **4.4 Impact of Combination of Beliefs on Work Ethic**

My initial attempt to find support for the mechanism I outlined above should be viewed in the context of my findings about the importance of combinations of beliefs. That is, it is possible that the above specification for the model of afterlife beliefs impacting work ethic omits the beliefs, or combination of beliefs, that actually matter. Referring to Table 1, there are very few people that believe only in hell and not in heaven. Given the similarity, when considering work ethic, between those that believe in heaven and hell and those who believe in neither heaven nor hell, it is not surprising that the inclusion of just one belief yields insignificant results. Given my above findings, adding in a term for the proportion of the population that believes only in heaven though small should have a significant effect in determining the proportion of people that hold work ethic as an important cultural value. The results of this are given by model 2 in Table 2.

The results of this addition of a variable to capture those that believe only in heaven, and who I showed have radically different feelings about worth ethic, has the significant effect I expected. In model 2, any significance on the belief in hell term disappears while the belief only in heaven term becomes significant at the less than 1% level. The R2 is satisfactory for a model built on vague concepts like beliefs and values, which undoubtedly result in measurement error when attempted to be quantified in a survey across languages and cultural histories. The specific interpretation of these results is that the proportion of the population that holds work ethic as an important variable decreases by about 0.67% for every 1% increase in the proportion of the population that believes only in heaven.

Once again, the sign on belief in hell is negative, opposite the results of McCleary and Barro. However, I ignore that result for now as it is insignificant at any reasonable level of significance.

Model 2 of Table 2, through only two terms, captures the full range of opinions on an afterlife available in the World Values Survey. Since everyone that believes in hell believes in heaven with all but a few exceptions, the term for belief in hell generally captures all those that believe there is an afterlife described by a heaven or hell idea. However, since not everyone that believes in heaven believes in hell, this model also captures that subset of people. The results can be thought of as relative to a country of complete nonbelievers (when both the belief in hell and the belief only in heaven terms would be 0).

I wish to test the robustness of these results. This can be done by adding in dummies for the survey waves. The possible effects that these dummies control for are numerous. For one, though the same question was used in each wave, the accuracy of its translation between years may have differed, as might have the methodology for administering the survey. Alternatively, adding in fixed effects can capture temporal swings between periods that are otherwise hidden, like global economic outlook or the impact of recent scandals on various religions or individuals beliefs. The results of this are reported in model 3 in Table 2, with the coefficients on the survey wave dummies relative to the 1981-1984 wave. The first question when analyzing the results in model 3 of Table 2 is if dummies for the survey waves are needed. The result is clear that they are; the adjusted R2 increased from 0.3797 to 0.4148 with the addition of the fixed effects while the coefficients on the

dummies are significant at the 1% level, except in the case of 2005-2007 where the significance of the coefficient is hampered by the smaller sample available from that wave.

The second observation from model 3 in Table 2 is that the term for belief in only heaven is still significant and barely changes. This gives me confidence that my findings around the significance of belief in only heaven on work ethic are robust.

However, this regression also presents an intriguing result that I previously ignored: the sign of the coefficient on belief in hell is negative, and in model 3 it is significant at the 5% level. McCleary and Barro arrive at a positive coefficient and my analysis up until this point suggests that this should have a positive coefficient as well. One explanation is that this is an endogeneity problem. In order to make this argument, one has to believe that work ethic can influence beliefs. This is not unimaginable as it exhibits a desire of individuals to believe in a way consistent with their actions. If someone works very hard, he might reflect on this and decide he needs a reason or explanation for working very hard. If he formerly believed in hell, he may be less likely to believe in hell, or at least less likely to believe he might be going to hell, once he or recognizes how hard he is working. He thus might be properly classified as believing only in heaven after this reflection, but retains a token belief in hell that is reported in the World Values Survey. This endogeneity problem does not have any clear solution as no instruments are readily available to instrument for belief in hell.

In addition, I think individuals are less likely to change the core beliefs they might have grown up with than change their actions to be in line with their beliefs. Literature supports this by suggesting that individuals do not become more or less religious as they grow wealthier, but they just come to relate to religious needs in a different way (Hirschle, 2013, p. 421). If this is the case, a better explanation of this negative coefficient might be a view of these results as relative to nonbelievers. This interpretation says that as people work harder, they come to find meaning in their work and perhaps less in the incentive of hell or heaven as a stick or a carrot. This does not harmonize with my earlier findings about combinations of beliefs and views of work ethic on an individual level. I take solace in the fact that the significant coefficient on belief in hell is not a robust result, appearing in only two of the models in Table 2.

I also note from models 1, 2, and 3 in Table 2 that the coefficient on the ex-communist indicator

is insignificant. I remove it to see if the adjusted R2 increases from this better-specified model. The result is presented as model 4.

Removal of the non-significant term for ex-communist country improved the adjusted R2. Other conclusions remain the same, with a similar and significant coefficient on the proportion of the population that believes only in heaven and the proportion of the population that believes in hell. This latter result still lacks a clear interpretation.

Model 4 is the most powerful model I present in Table 2 for linking beliefs in heaven and hell to the cultural value of work ethic. With an R2 of 0.4575 and an adjusted R2 of 0.4203, the model is strong in explaining work ethic as a cultural value. An increase of 1% in the proportion of the population that believes only in heaven corresponds to a decrease of 0.656% in the proportion of the population that holds work ethic as an important value. Similarly, an increase of 1% in the proportion of the population that believes in hell corresponds to a decrease of 0.192% in the proportion of the population that feels work ethic is an important value. This significant latter observation does not fit with the theories I put forward above for explaining how beliefs, especially those in the afterlife, might impact work ethic. Intriguingly, it also is at odds with what is expected given the logistic regression results in Equation 2.

My conclusion from this final model is that afterlife beliefs are related to the importance of work ethic as a cultural value. I call into question the mechanism behind this relationship. I theorized that if an individual believes in hell and heaven, he or she should be more likely to work hard in an effort to earn a spot in heaven and avoid hell. However, my results suggest an increase in spiritual beliefs about the afterlife relate to less of an emphasis on work ethic. One possible explanation, not tested here, is that individuals are more concerned with salvation being related to formal religious services and thus feel that hard work is less important while time spent in religious contemplation or at religious services is more important. The direction of this causality can also be flipped, introducing a problem of endogeneity, as individuals who work harder shift their beliefs to line up with their actions, maintaining only a token belief in hell and being better classified as believing only in heaven. This endogeneity cannot be readily instrumented. Barring this endogeneity problem, my results suggest a link between beliefs and cultural values, specifically

combinations of possible beliefs in an afterlife and work ethic.

## 4.5 Work Ethic and Economic Performance

After establishing a link, albeit of questionable theoretical explanation, between beliefs and work ethic, I turn to the question of whether work ethic matters for economic outcomes. I focus specifically on economic growth in terms of per capita GDP.

Work ethic should impact economic growth. If work ethic is an important value for a culture, that is more people think it is an important value to pass on to children, that culture will work harder than one that does not share this value. That hard work will translate into productivity which will translate into economic growth.

To begin with, I identified five outliers in economic growth and removed them from the dataset. These outliers were three-year economic growth averages above 10.8% or below -5.1% and significantly skewed my work when left in.

With this cleansed dataset, I added in additional variables using guidance from McCleary and Barro (Religion and Economy, 2006, p. 67). They include variables such as the inverse of life expectancy, years of school attainment, openness and trade growth, indicators for democracy, the log of fertility, and the ratio of investment to GDP (McCleary & Barro, 2006, p. 66). Due to constraints on what was available from the World Bank, I use the inverse of life expectancy, the log of fertility, and the gross capital formation as a percent of GDP. I add in a variable for the proportion of the population that thinks hard work is an important cultural value. Model 1 in Table 3 presents the results.

**Table 3 Percentage growth in GDP per capita modeled by work ethic**

	<b>Proportion of population that thinks hard work is important to teach children used for work ethic</b>	<b>Belief in hell and only heaven used as instruments for work ethic</b>
<b>work ethic</b>	1.30613 (1.3296)	0.2300 (5.041)*
<b>log of GDP per capita</b>	-0.9894 (0.2605)***	-0.0782 (0.6007)
<b>log of fertility</b>	-3.0684 (0.8359)***	-0.2863 (1.9347)
<b>capital formation (investment)</b>	-0.0494 (0.0399)	-0.0311 (0.0491)
<b>inverse of life expectancy at birth</b>	-84.1969 (194.4104)	-562.5000 (194.4104)
<b>constant</b>	15.4434 (3.9782)***	7.4717 (6.5002)
<b>R2</b>	0.327	(Not applicable)
<b>Adjusted R2</b>	0.2912	(Not applicable)

\* indicates significance at 10% level; \*\* significant at 5%; \*\*\* significant at 1% All numbers are rounded to four decimal places. Both models had 100 observations.

Model 1 in Table 3 is significant at the 1% level with a respectable R2, indicating its explanatory power. The added term for hard work does not pick up a significant coefficient, though. This fails to support the theorized link between cultural values and economic outcomes, particularly work ethic and economic growth.

There may be a problem with the simple model specified in model 1, though. The two-way interaction between work ethic and economic growth is not hard to imagine. I explained how work ethic might drive economic growth, but it is possible that economic growth influences work ethic. As an economy grows, individual businesses see an increase in business. This increase in business prompts employees to work harder to keep up until new employees can be hired, a process that takes time. To remain consistent with their actions, employees might tend to think hard work is an important cultural value. Alternatively, if a countrys growth is due to factors not dependent on the labor force, like an increasing demand for a rare resource the country possesses, then economic growth will be disconnected from hard work. Individuals may feel a lack of need to work hard, thinking that the economy will grow regardless of their efforts. In terms of my regression, this describes an endogeneity problem.

I try to address this endogeneity problem with instrumental variables, namely the beliefs in an afterlife examined above. It is hard to make an argument for how beliefs might be influenced by economic growth, though if beliefs come from attendance at religious services and attendance at religious services declines due to economic growth as the secularization hypothesis suggests, there might be a two-way interaction. This two-way interaction would take place over many years, though probably decades. Different beliefs in an afterlife tend to be core to religions and thus the first thing religious individuals are taught, usually when they are growing up. Thus, they are likely among the last beliefs an individual would let go off as they drift away from religion. To imagine that these beliefs are quickly changed on a mass scale due to an already slow decline in attendance due to economic growth does not make sense, especially given that I look at economic growth data that lags very closely behind the survey data.

The second interpretation of an instrumental variables approach in this context is that it captures the portion of work ethic as a cultural value that is attributable to afterlife beliefs. This is an approach and interpretation that has precedence in literature (Guiso, Sapienza, & Zingales, *Does Culture Affect Economic Outcomes?*, 2006, p. 35). The results of my instrumental variables technique, using the proportion of the population that believes in hell and the proportion that believes only in heaven as instruments, are given by model 2 in Table 3.

The results of the instrumental variable technique in model 2 of Table 3 are fairly strong, taking away the significance of the coefficients on variables that are traditionally important in modeling economic growth. This model suggests that the only significant coefficient for explaining economic growth is the portion of work ethic that can be explained by beliefs in the afterlife. That is, an increase of 1% in the proportion of the population that thinks work ethic is an important value is attributable to a 0.0992 percentage point increase in GDP per capita growth. Since this was instrumented for by spiritual beliefs in the afterlife, this result refers specifically to the portion of work ethic attributable to such beliefs.

The diagnostics for the instrumental variables, presented in Table 4, suggest that I identified a problem of endogeneity and approached it properly. The models above discuss how work ethic can be modeled with afterlife beliefs; this is analogous to the weak instruments test in Table 4. The standard test statistic for the significance of an instrumental variables regression is the Wu-Hausman statistic. A Wu-Hausman test has as its null hypothesis that the original estimator, here the coefficient on work ethic in model 1 of Table 3, is consistent by comparing it to the output of an instrumental variables technique. If the null hypothesis is rejected, there is evidence it is inconsistent; said differently, there is evidence that an instrumental variables approach was needed. Here, I reject the null hypothesis at the 10% level of significance. This is confirmed by a Wald Test. Here, the null hypothesis is that the original variable, work ethic, is exogenous. I reject this at any meaningful level of significance and conclude that an instrumental variables approach was valid.

**Table 4 Instrumental Variables Diagnostics**

	Degrees of freedom 1	Degrees of freedom 2	Statistic	P-value
<b>Weak instruments</b>	3	93	101.996	$2 * 10^{-16}***$
<b>belief in only heaven</b>	2	92	2.881	0.0611*
<b>ex-communist</b>	5	94	7.04	$1.253 * 10^{-5}***$

\* indicates significance at 10% level; \*\* significant at 5%; \*\*\* significant at 1%

Given the strength of this result, it is important to check to see that the instruments are not correlated with the error terms from model 1 in Table 3, a requirement of instruments. I regressed the residuals from model 1 on the proportion of the population that believes in hell and only in



heaven. The results are given by Equation 3.

**Equation 3 – Test of instruments by regressing residuals of non-instrumented growth model on the instruments**

$$\begin{array}{rcccl} \text{residuals} = & 0.3083 & + 0.2744 & (\text{belief in hell}) & -3.6818 & (\text{belief in only heaven}) \\ & (0.6792) & & (0.8758) & & (2.0508)^* \end{array}$$

\* indicates significance at 10% level; \*\* significant at 5%; \*\*\* significant at 1%  
 $R^2=0.0439$      $F\text{-statistic: } 2.227 (2, 97 \text{ DF}), p\text{-value: } 0.1133$

**Equation 3 Test of instruments by regressing residuals of non-instrumented growth model on the instruments**

The results in Equation 3 give some cause for concern in that the proportion of the population that believes only in heaven has a significant coefficient at the 10% level. However, my instruments explain only about 4% of the variation in the residuals from model 1 in Table 3. Further, the regression model is not significant given a p-value of 0.1133; this test of model significance fails to reject the null hypothesis that all of the non-constant coefficients in the regression are zero. With this check of the instruments and the diagnostic tests discussed earlier, I conclude that the instrumental variables approach is applicable in instrumenting work ethic when trying to model economic growth.

With this validation of the instrumental variables technique, I return to the results of model 2 in Table 3. I am skeptical of this output given that the variables I used as controls have a vibrant history of use in modeling economic growth. My model was also certainly underspecified and should have included measures of educational achievement, democratic status, and openness. Regardless, my results suggest that work ethic, as explained by beliefs, is influential in determining economic growth.

## 4.6 Data Analysis Caveats

The above results are, ultimately, based on survey data. Any attempt to quantify abstract cultural phenomena is going to be met with difficulties. A cross-country survey compounds these difficulties by introducing problems of translation and understanding of religious or spiritual concepts. A

specific concern is the Westernized concept of heaven and hell as good and bad locations a soul ends up in for eternity. Eastern religions have concepts of an afterlife of sorts, but they may be sufficiently far from the Western concept that the survey questions I use inconsistently measure analogous beliefs.

I would also have liked to validate the above findings with results from similar World Values Survey questions or other sources. There are questions in the World Values Survey that could be used to validate the spiritual beliefs I considered and their impact on work ethic; for example, there is a question that asks peoples perception of hard work leading to a better quality of life. Another question asks if work is important in life. These same questions could also be used to validate the connection between work ethic as a function of beliefs and economic outcomes. There may also be other spiritual beliefs that relate to work ethic that can be pulled from the World Values Survey.

## 5 Conclusion

My hypothesized connection between spiritual beliefs and cultural values and between cultural values and economic outcomes has evidence from World Values Survey data focusing on belief in hell, belief only in heaven, and the importance of hard work as a quality to teach children.

Starting with an expansion on the work done by McCleary and Barro (Religion and Economy, 2006), I examined the importance of various afterlife beliefs in determining work ethic. I found that combinations of beliefs, namely a belief in heaven and not hell, provide a better explanation of work ethic on the country level. Specifically, a 1% increase in the proportion of the population that believes in hell is related to a 0.19% decrease in the proportion of the population that thinks work ethic is an important cultural value. For those that believe only in heaven, a 1% increase in believers corresponds to a 0.66% decrease in the proportion of people that think work ethic is an important cultural value.

I then turned my attention towards the connection between work ethics as a cultural value and economic growth as an economic outcome. Looking directly at work ethic has endogeneity problems, so I instrumented for it with belief in heaven and belief in hell per my earlier analysis.

This captures work ethic as a function of beliefs when modeling economic growth. This approach yielded a significant coefficient on work ethic that suggests a 1% increase in the proportion of the population that holds work ethic as an important value drives a 0.099% increase in real GDP per capita growth. Attempts to validate the instruments and the approach suggest that beliefs in hell and only in heaven were sufficient instruments and an instrumental variables technique was valuable.

The overarching result of this data analysis is support for the hypothesized mechanism that beliefs influence cultural values which influence economic outcomes. This mechanism means that beliefs influence economic outcomes, potentially in a very substantial way.

To relate my results to existing research and hypotheses, I find evidence for a somewhat Weberian view of economic outcomes. I say somewhat because the source of spiritual beliefs, like those in different versions of an afterlife, is not analyzed in my work. Rather, my analysis suggests that more time and attention should be put into looking at specific spiritual beliefs, their relation to each other, and how they impact cultural values. Further work should be done on various cultural values and how they impact economic outcomes, though some literature on this exists for cultural values like trust or thrift. Attempting to look at how religious denominations relate directly to economic performance is putting the cart before the horse.

## Acknowledgement

Thank you to Dr. Mahmoud Amin El-Gamal at Rice University for his input on this paper when it was being drafted for his research seminar.

## References

- [1] Ali, A., & Al-Owaihan, A. (2008). Islamic work ethic: a critical review. *Cross Cultural Management: An International Journal*, 15(1), 5-19.

- [2] Azzi, C., & Ehrenberg, R. (1975). Household Allocation of Time and Church Attendance. *Journal of Political Economy*, 83(1), 27-56.
- [3] Barro, R. (2003). *Religion Adherence Data*. Retrieved from <http://scholar.harvard.edu/barro/publications/religion-adherence-data>,
- [4] Blum, U., & Dudley, L. (2001). Religion and economic growth: was Weber right? *Journal of Evolutionary Economics*, 11, 207-240.
- [5] Guiso, L., Sapienza, P., & Zingales, L. (2003). *People's opium? Religion and economic attitudes*. *Journal of Monetary Economics*, 50, 225-282.
- [6] Guiso, L., Sapienza, P., & Zingales, L. (2006). *Does Culture Affect Economic Outcomes?* *Journal of Economic Perspectives*, 20(2), 23-48.
- [7] Hirschle, J. (2013). "Secularization of Consciousness" or Alternative Opportunities? The Impact of Economic Growth on Religious Belief and Practice in 13 European Countries. *Journal for the Scientific Study of Religion*, 52(2), 410-424.
- [8] Knack, S. (2001). *Trust, associational life, and economic performance*. Retrieved February 25, 2014, from Munich Personal RePEc Archive: <http://mpira.ub.uni-muenchen.de/27247/>
- [9] Knack, S., & Keefer, P. (1997). *Does Social Capital Have an Economic Payoff? A Cross-Country Investigation*. *The Quarterly Journal of Economics*, 112(4), 1251-1288.
- [10] Landes, D. (1999). *The Wealthy and Poverty of Nations*. New York: W.W. Norton.
- [11] McCleary, R., & Barro, R. (2006). Religion and Economy. *The Journal of Economic Perspectives*, 20(2), 49-72.
- [12] McMurray, A., & Scott, D. (2013). Work Value Ethic, GNP Per Capita and Country of Birth Relationships. *Journal of Business Ethics*, 116, 655-666.
- [13] Noland, M. (2005). Religion and Economic Performance. *World Development*, 33(8), 1215-1232.

- [14] Pew Research Center. (2012, October 9). *Religion and the Unaffiliated*. Retrieved April 30, 2014, from <http://www.pewforum.org/2012/10/09/nones-on-the-rise-religion/>
- [15] Pryor, F. (2005). National Values and Economic Growth. *American Journal of Economics and Sociology*, 64 (2), 451-483.
- [16] The World Bank. (2014). *World Development Indicators*. Retrieved from The World Bank: <http://data.worldbank.org/data-catalog/world-development-indicators>
- [17] World Values Survey Association. (2014). *Longitudinal Data*. Retrieved January 23, 2014, from World Values Survey: <http://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp>
- [18] Zak, P., & Knack, S. (2001). Trust and Growth. *The Economic Journal*, 111, 295-321.

# Evaluating the Impact of Microfinance on Health Decisions

Robert Fluegge

Harvard University

May 2015

## **Abstract**

This paper attempts to understand the impact that improved access to microfinance services has on health decisions for low-income communities in developing countries. Using data from a randomized trial in 104 slums of Hyderabad, India, regression-based investigation found no evidence that health decisions changed significantly for households in the communities after 15-18 months of improved access to microcredit. Specifically, there were no statistically significant differences between the treatment and control groups in the probability that a household purified their water or had health insurance of any kind. There was a statistically significant increase in the probability that a household uses a latrine in the treatment group, but the effect was quite small. Collectively, it appears that access to microfinance has little impact on health behaviors for members of low-income communities.

## **1 Introduction**

Microfinance, here defined as small-scale lending typically provided to low-income women in developing countries, has gained enormous public popularity in the past decade as a potential tool development tool. Starting largely with the award of the 2006 Nobel Peace Prize to Grameen Bank, a non-profit microfinance institution based in Bangladesh, microfinance has been touted as an avenue towards the end of poverty around the world. The idea, largely pushed by popular economics publications like Muhammad Yunus' "Creating A World Without Poverty" (Yunus & Weber, 2014),

has caught the eye of the public and policymakers as an alternative to traditional humanitarian aid.

But with all of the hype surrounding microfinance, there is a surprising paucity of economic literature devoted to rigorous analysis of the economic and social impacts of access to credit for impoverished communities in developing countries. In particular, there is almost no evidence on the effects that access to microfinance can have on the health behaviors of members of these communities. It has been well established that individual and public health are an important indicators of the overall development of a country, as can be seen in the Millennium Development Goals published by the United Nations ("United Nations Millennium Development Goals," 2014). This paper attempts to address the lack of evidence surrounding microfinance using data from a randomized trial in India conducted from 2005-2007. Specifically, the goal is to determine how the health behaviors of households in a low-income community change with access to a microfinance program using regression analysis. The evidence can hopefully separate myth from fact in order better inform NGOs, policymakers, and the general public about the real effects of access to credit for low-income communities.

## **2 Literature Review**

As previously stated, there is a surprising lack of evidence surrounding the health effects of access to microfinance. Of those academic publications that do exist, a majority focus on institutions that pair microfinance with other health services. For example, a study on microfinance paired with health education services found positive health changes across a number of indicators ("Leatherman & Dunford, 2014). However, these inquiries do little to isolate the effects of microfinance itself, which is critical to know because microfinance can be scaled far more easily and profitably if it is not paired with not-for-profit services. Another segment of the literature on the relationship between microfinance and health focuses on women's empowerment, particularly the reduction of domestic violence. Most of these publications are based on a study of microfinance and domestic partner violence in South Africa (Kim, 2014), which found that access to microcredit indeed can reduce

intimate partner violence significantly. While this is certainly an important result, this study fails to paint a picture of the effects of microfinance on more general health behaviors.

There are in fact a few studies on microfinance that mention the health effects of credit alone, but they do so in passing and without sufficient specificity. For example, the paper published by the designers of the Spandana trial in India in 2007, which is the data set that this paper uses, says simply that the authors did not find any health differences between treatment and control groups (Banerjee, 2014). Nowhere in the literature review undertaken for this paper is health behaviors or decision-making mentioned at all. That lack of analysis on the impact of microfinance on health behaviors and the superficial treatment of health in the microfinance literature in general forms the inspiration for this paper.

### **3 Conceptual Background**

The economic rationale behind a credit market is simple. Credit allows people or organizations to smooth their consumption over time rather than consuming large amounts at some points and consuming little at other times. This is generally viewed as maximizing consumer utility. For example, a person buying a house will take out a loan and pay it back in small increments over time so that their expenditures do not change drastically at the time of the purchase. At its essence microfinance allows low-income consumers to do the same, although many institutions provide the loans for the explicit purpose of investing capital in a self-run business. Investing in a self-run business will ideally increase the income of the consumer as it will make the business more profitable and allow the business to grow in ways that were not possible with its previous revenue stream.

The necessity of credit to grow a small business in a low-income area can be thought of as part of what economists term the poverty trap. The poverty trap usually describes the situation in which a low-income person does not have enough resources to invest. Investment in this context often describes investment into human capital, such as further education, which allows the individual to increase their income in the future. Without sufficient wealth or income to invest, a low-income person has no way to increase their human capital in order to generate more income, and thus



becomes trapped in their low-income state. In the context of a developing country like those that this paper investigates, the high prevalence of unofficial self-run businesses means that low-income individuals do not have the ability to invest either in their own human capital or in their business, and so their business is 'stuck' with extremely low profits and the individual remains with little income.

The other important economic theme of this paper is the link between wealth and health outcomes. It is clear from sources like the United Nations and the World Health Organization that country-level income is strongly correlated with a variety of public health indicators ("United Nations Millennium Development Goals," 2014). Furthermore, individuals with greater wealth tend to have better health outcomes across the world. However, the causal link is far less understood. Based on a review of the literature, it is generally assumed that causality extends in both directions. Greater wealth causes better health because wealth improves access to health services, and better health causes greater wealth because a healthy person is more productive and will often have a higher income as a result. In low-income settings like those discussed in this paper, the causality is likewise assumed to extend in both directions.

Microfinance might affect the health behaviors of those receiving the loan in two competing ways. First, since a microfinance institution is intended to increase the income of the individuals it serves, the changes in health behaviors depend on whether 'health' when viewed as a consumable good should be considered a normal good for this level of income. If a person in a low-income community tends to view health as a normal good, we should see an increase in behaviors associated with health consumption as income increases. Based on the fact that health generally increases with income, it is safe to assume that health is indeed a normal good. However, the analysis must also take into account the fact that loans allow people to smooth consumption even when buying a potentially expensive good. It might be the case that with a loan it is now possible for the person receiving the loan to buy a relatively expensive item, and that person buys the item by reducing the rest of their consumption over the period of time that they pay the loan back. This effect would tend to decrease the resources devoted to 'health' as a good because those resources would instead be devoted to repaying the loan. This effect will be particularly acute if the income of the individual

receiving the loan does not appreciably increase over the repayment period. Since these two effects act in opposite directions, it is difficult to tell from theory alone whether health behaviors will change with access to formal microfinance for low-income individuals in developing settings. This fact makes empirical analysis of the issue all the more important.

## 4 Methods

The data used in this regression analysis comes from a randomized controlled trial performed by researchers affiliated with the for-profit microlender Spandana in Hyderabad, the fifth largest city in India, from 2005-2007.<sup>1</sup> The researchers identified 104 "slums", also called "areas", that Spandana would be indifferent towards providing loans to. Each area contained between 46 and 555 households. The areas are characterized as "poor, but not the poorest of the poor" (Banerjee, 2014) and were non-adjacent to minimize spillover effects. The study purposely excluded areas with large densities of people that Spandana believed would be particularly profitable and areas that contained a large share of workers in migrant industries like construction because of difficulties tracking down migrant clients for repayment of the loan. Starting in 2005, Spandana began opening branches in 52 randomly selected areas out of the 104 total. In 2007, after between 15 and 18 months of the branch being opened, the researchers surveyed on average of 65 households per area for a total of about 6,850 households. The survey took between 1-2 hours to administer and contained questions on residents of the household, expenditures, business activities, and behaviors. Only people who had been living in the area since Spandana had introduced a branch in 2005 were surveyed. Because households in the treatment group were selected to be surveyed based on baseline data that Spandana had collected on the area in 2005, households in the treatment group were more likely to be Spandana clients. Therefore, the results from each area are weighted in the regression according to account for the increased probability of selection of Spandana clients. The data from each area are clustered to account for similar characteristics.

The structure of the loans was based on a group model developed at Grameen Bank. Spandana

---

<sup>1</sup>The methods outlined in the section are described fully in Banerjee, Duflo, Glennerster & Kinnan, but also briefly in references Banerjee, Karlan, & Zinman and "Primary Menu."

lent to women aged 18-55 who had lived in the area for at least a year. Each received about Rs. 10,000, which is equivalent to \$200 at the 2007 exchange rate or \$1000 adjusted for purchasing power parity. The interest rate was approximately 12%, to be paid back over 50 weeks. The recipients of the loans were required to form groups of 6-10 on their own, and each group was responsible for the loans all its member. This grouping tends to select for women who are more likely to repay their loans, or at least selects for women whose peers view them as more likely to repay their own loan. A collection of 25-45 groups formed a center, and each center met periodically to discuss and conduct repayment. The Spandana loan is slightly different from loans provided by most other microfinance institutions in that it does not require the women to own or start a business, although the explicit expectation of Spandana executives was that the loan recipients would indeed use the money for business purposes. From the point of view of the researchers who designed the study, this type of loan allows for analysis of the effects of access to credit alone without the interfering effects of additional services or requirements.

## 5 Relevant Prior Results

The authors of the original paper based on the study make several notes that are important for this paper's analysis of the data. First and most importantly, there were other lenders and microfinance organizations operating in both the treatment and control groups both at the start and throughout the trial. At the time of the survey, a full 89% of households had an outstanding loan of some kind. Although the usage of formal microfinance programs was only 18.4% in the control group (8.4 percentage points lower than the treatment group), analysis cannot hinge on access or lack of access to credit. Instead, regression demonstrates the "intent-to-treat" effect of putting a branch into the area and actively attempting to expand microcredit to the residents (Banerjee, 2014), which did in fact increase the ease and availability of microcredit products for the area. It is also possible that the characteristics of the people who already had microfinance may be different than the characteristics of those who utilize the service when access is easier, which tempers the analysis somewhat.

Second, it is important to keep in mind that this trial was carried during a period of extremely rapid economic growth in Hyderabad. From 2005-2008, real average household consumption increased by more than 100% (Banerjee, 2014). While level of growth might certainly be realistic for many of the communities in which institutions have an interest in establishing microfinance services, the trial will be less useful when trying to understand the effects of microfinance in a stagnant or low-growth setting.

Third, it is important when analyzing effects from this trial to note that the uptake of microcredit services was relatively small and in some cases took the place of other lenders. It appears that no more than 35% of this eligible population will actually utilize microfinance at a given time, which is a far cry from the 80% expected utilization and is lower than the uptake seen in the few other randomized trials of this type (Banerjee, Karlan, & Zinman, 2014). This means that any intent-to-treat effects will not show up as strongly in the data as it would if the uptake were higher. However, it is still true that microfinance utilization was 8.4 percentage points higher in the treatment group than in the control group (26.7% to 18.4%) so there is certainly a statistically significant intent-to-treat effect. It appears that often official microfinance institutions will take the place of existing community lenders rather than provide credit where there was none before. The number of consumers who borrow from informal sources decreased in the treatment group, and there was no statistically significant effect on overall amount borrowed (Banerjee, Duflo, Glennerster & Kinnan, 2014).

The original study also produced relevant results with respect to income and female empowerment that we might expect to prefigure our results. In the treatment group there was not a large increase in business profits, but profits did increase markedly for businesses that were already profitable, specifically those above the 95th percentile of profitability (Banerjee, Duflo, Glennerster & Kinnan, 2014). Furthermore, there were more businesses created but they tended to be smaller and less profitable than average. This may be due to the fact that the treatment likely had the biggest effect on the marginal clients: many of those who could put a loan to good use may have already gotten a loan from somewhere before treatment, which leaves only those who will take a loan if the cost is low but not if the cost is high (cost here being the economic cost including finding a

loan, traveling outside the neighborhood if necessary, etc). After treatment, businesses were slightly more likely to be female-run but no more people were employed by businesses than before (Banerjee, Duflo, Glennerster & Kinnan, 2014). Finally, labor supply in treatment groups increased, with the household head and their spouse contributing on average about 3 more hours per week to self-run businesses. With these results in mind, this paper will now examine the intent-to-treat effect on health behaviors.

## 6 Results

This analysis uses regression to determine the effects of improved access to microfinance services on health behaviors. There were 4 variables of interest: use of latrines, purification of water for adults, purification of water for children aged 0-2, and purchase of health insurance. The use of latrines and purification of water are general measures of health behavior that might prevent disease and are influenced by the Millennium Development Goals ("United Nations Millennium Development Goals," 2014). Purchase of health insurance includes both traditional health insurance as well as medical/accident insurance. Summary statistics for these dependent variables can be found in Table C. The regression used, which is the same format as the original paper from this data set (Banerjee, Duflo, Glennerster & Kinnan, 2014), was set up as follows:

$$y_{ia} = \alpha + \beta * Treat_{ia} + X'_{ia}\gamma + \epsilon_{ia} \quad (1)$$

Where  $y$  is an indicator variable for presence of behavior in household  $i$  and area  $a$ ,  $Treat$  is an indicator variable denoting whether or not the household was in a treatment area,  $\beta$  is the intent-to-treat effect,  $X'_{ia}\gamma$  is the control variables and their coefficients, and  $\epsilon_{ia}$  is the error term.

This analysis used several control variables, some at the area level and some at the household level. At the area level the regression controlled for number of households in the area, probability that an adult in the area was literate, probability that the head of a household was literate, and the number of resident run businesses. At the household level the regression controlled for number of people in the household and ownership of cell phones and a fridge, both of which are proxies for

the wealth of the household because no income data was collected. The summary statistics of all the control variables can be found in Table B.

The regressions found no statistically significant changes in either water purification variable or in the health insurance variable, but did find a statistically significant increase in the use of latrines among households in the treatment group. The effect on latrine usage was quite small: a resident of a treatment area was about 2% more likely to use a latrine than not use a latrine, with a 95% confidence interval of approximately (.008,.031). There were no statistically significant effects at all when the regressions were run without control variables. The R-squared values for the controlled regressions were quite low, with most in the range .02-.05. This was expected because health behaviors vary wildly from person to person with little explanation, and will be difficult to fully explain using only the variables available. However, this paper argues that even with relatively small R-squared values the regressions do explain enough of the variation to be useful. A full list of the results from the regressions can be found in Table A.

The control variables in the regression were chosen to account for various differences between both households and areas that might affect health behaviors. At the area level, the analysis controlled for number of households, proportion of adults who are literate, proportion of heads of households who are literate, and the number of businesses. A larger number of households might increase the likelihood that households split the cost of health goods like latrines, and therefore might be more likely to use a latrine. A higher proportion of literate adults and literate heads of households generally means a more educated community and leadership, which might be more knowledgeable about positive health behaviors. The number of businesses in a given area might indicate the amount of time or resources that residents put into their businesses, which might take the place of other health behaviors.

At the household level the regressions controlled for number of people in the household and ownership of cell phones and a fridge. The number of people in a given household might change the amount of resources that the household is able to devote to each person, and could potentially affect the investment into health-producing goods. Ownership of a cell phone or a fridge is used here as a proxy for the wealth of the household because neither income nor wealth data was collected.

## 7 Discussion

It appears based on regression analysis that access to microcredit does not significantly impact the health behaviors of households in a low-income community. This might have been expected from the relatively modest changes in business profitability observed in the previous study based on this data (Banerjee, Duflo, Glennerster & Kinnan, 2014), but is an important result in its own right. Since average business profits did not change appreciably for most businesses as a result of the trial, most residents did not see a large increase or decrease in income as a result of the loan. The relatively small change in income for most people means that few clients have extra resources to devote to health-related activities. Those that did profit from the loan tended to already control profitable businesses before the loan was provided and in all probability already had 'good' health behaviors.

The probability that a household uses a latrine was the only health behavior that increased in the treatment group compared to the control group. This is an interesting result that does not seem to have a clear explanation, but it is important to point out that this behavior is different from the other health behaviors in potentially significant ways. Out of the four dependent variables studied, use of a latrine might be the most instantly rewarding behavioral change because there is often a level of shame or repulsion involved with not using a latrine. And unlike the other indicators, use of a latrine may not require a significant financial investment because they can be built by community members using pooled resources. Ultimately, however, the change in probability of using a latrine, while statistically significant, is not particularly important in this case because the coefficient in the regression is so low. An increase of 2% in the number of residents who use a latrine is not cause for celebration and in reality will have little effect on the disease burden that the community faces.

All of these conclusions add further evidence to the notion that microfinance alone is not the 'cure' for poverty. Without a significant change in the health behaviors of a low-income community in a developing country, that community will struggle grow economically and socially as fast as it might otherwise. These results support the general consensus found in academic microfinance literature that microfinance does not have the incredible benefits its proponents once believed it to.

As such, it is limited as a development tool (Banerjee, Karlan, & Zinman, 2014).

It is important to note that while health behaviors did not appreciably increase, they also did not decrease. As discussed in the Conceptual Background section, the changes in consumption that accompany loans might very well reduce investment in health-producing activities. An individual receiving a loan might decide to reduce their consumption of, for example, water filters in order to pay back a loan that they took out from a microfinance institution. The regressions here show no statistically significant evidence of such an effect. One might argue, however, that while there was only one statistically significant result, each of the coefficients on the other 3 variables were negative. It is possible that with more power or in a different set of circumstances such effects might be significant, and it is certainly true that the topic merits further investigation.

There are several potential limitations to the analysis conducted based on these regressions, most of which have been alluded to already. First, the fact that there were a number of informal lenders and a few microfinance organizations operating in both the treatment and control groups throughout the process means that we cannot make a definite conclusion about the effects of absolute access to microfinance products. Instead, we can only look at the intent-to-treat effect in the population.

Second, the trial was conducted during a period of strong economic growth in Hyderabad, so the conclusions drawn from the data may not reflect the effects of improved access to microfinance in low-growth situations. It seems likely, however, that many applications of microfinance will be in similarly high-growth contexts, and this analysis is useful for those contexts. Furthermore, because the regression attempts to control for household wealth using indicators like ownership of a cell phone or fridge, the results may generalize to stagnant or low-growth situations as well as high growth ones. However, since the controls are imperfect measures of wealth, generalizations to low-growth settings should be made cautiously.

Third, as discussed in the Prior Relevant Results section uptake of microfinance products in the treatment areas was relatively low compared to similar randomized trials. This means that the power of the regressions is reduced.

Fourth, this trial only investigated the effects of pure access to credit, and does not draw conclusions on microfinance programs that are paired with other development initiatives like educational



programs or requirements to start a business. It seems likely based on some preliminary studies that microfinance does not interfere with the effects of humanitarian programs ("Leatherman & Dunford, 2014), but more research into specific health behaviors is needed in this area before definitive conclusions can be drawn. It is also important to realize that the timeframe analyzed was only 15-18 months after introduction of the branch, and it is possible that changes in health effects will take hold after a longer period of time. It seems likely that if microfinance does in fact increase income in the long term, development indicators like health behaviors will change as well. This paper's analysis is limited to the short term effects of improved microfinance access on health behaviors.

Finally, since every person in the sample population was living in the same city and was presumably part of the same 'culture,' there might be problems when generalizing these results to other countries with other cultures. Comparisons of results in the original paper published on this data to the results of other randomized trials across the developing world are similar, which lends some confidence to the idea that results on health behaviors is likely also consistent, but it is important to keep in mind that behavioral reactions to microfinance might differ across different cultures. As such, attempts to generalize these conclusions worldwide must be undertaken with a healthy caution.

## 8 Conclusions and Implications

As a whole, the evidence from this trial indicates that improved access to microfinance does not appreciably encourage or discourage health-related behaviors. The explicit aim of this paper is to provide private entities, policymakers, and the general public with a greater understanding of the true effects that access to microcredit has, and the information should inform both private decisions and policy debates. It appears that, consistent with analyses of other development indicators in microfinance trials, microfinance alone will not solve poverty or health problems. Microfinance as a development tool should be shown for what it is: an imperfect but probably positive part of the solution. It is also abundantly clear that the evidence on the health effects of microfinance

is simply too sparse to make any truly conclusive statements about causality. Further research must be undertaken, in different cultures and circumstances and for longer periods of time, before a verdict can be reached. This paper's most ardent recommendation is increased funding of trials like the one discussed here so that the question can be more fully explored. Until then, no definitive statements can be made.

Table A: Regression on Health Behaviors by Household								
Dependent Variables (binary, household):	Household has Health Insurance		Household uses a Latrine		Water Treated for Adults		Water Treated For Children Aged 0-2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Controls	No Controls	Controls	No Controls	Controls	No Controls	Controls	No Controls
Independent Variable:								
Treatment Group (binary, household)	-0.0267 (0.0564)	-0.00215 (0.0592)	0.0199*** (0.00588)	0.00920 (0.0223)	0.00507 (0.0233)	0.00592 (0.0279)	-0.00252 (0.00429)	-0.000364 (0.000690)
Controls:								
At least 1 child ages 0-2 (binary, household)							0.486*** (0.0312)	0.484*** (0.0318)
Treatment/Child interaction (binary, household)							-0.00729 (0.0399)	-0.00600 (0.0406)
Household owns a cell phone (binary, household)	0.0470*** (0.0120)		0.0463*** (0.00617)		0.0966*** (0.0138)		0.0233*** (0.00621)	
household owns a fridge (binary, household)	-0.0142 (0.0161)		0.0337*** (0.00765)		0.192*** (0.0164)		0.0334*** (0.00827)	
number of households (area)	-0.000245 (0.000232)		0.000115*** (2.21e-05)		2.02e-05 (8.69e-05)		-6.66e-05** (2.84e-05)	
Proportion of adults who are literate (area)	-0.194 (0.593)		0.278*** (0.0638)		0.0311 (0.238)		0.0821 (0.0830)	
Proportion of household heads who are literate (area)	-0.0699 (0.495)		0.0614 (0.0535)		0.434** (0.219)		0.0762 (0.0699)	
Number of businesses (area)	0.00372* (0.00201)		-0.00163*** (0.000193)		-0.000311 (0.000586)		4.08e-05 (0.000224)	
Residents in each household (household)	0.00138 (0.00276)		0.000829 (0.00136)		-0.00454 (0.00293)		0.000936 (0.00202)	
Constant	0.311 (0.227)	0.217*** (0.0450)	0.685*** (0.0240)	0.933*** (0.0145)	0.00358 (0.0962)	0.391*** (0.0228)	-0.113*** (0.0318)	0.000784 (0.000548)
Observations	6,790	6,791	6,787	6,788	6,789	6,790	6,790	6,791
R-squared	0.028	0.000	0.052	0.000	0.056	0.000	0.429	0.420
Number of Clusters	104	104	104	104	104	104	104	104
Notes								
(a) This table presents the results of the regressions on all four outcomes both with controls and without controls								
(b) Variables marked with a (binary) tag are binary variables, all dependent variables are binary								
(c) Variables marked with an (area) tag are averages over each area, variables marked with a (household) tag pertain to each individual household, all dependent variables are defined by household								
(d) Each regression uses the same controls except for the regression on whether the household treats the water that children ages 0-2 drink, which controls for whether or not there is a child in that age range and includes a an interaction term that is the product of the independent variable and the presence of a child in that age range								
(e) Robust standard errors in parentheses								
(f) *** p<0.01, ** p<0.05, * p<0.1								

**Table B: Control Variables Summary**

	Mean	Std. Dev.	Min	Max
<b>Area Variables</b>				
Number of Households	309.9	147.2	46	555
Proportion of Adults who are literate	0.65	0.09	0.4	1
Proportion of Heads of Households who are Literate	0.69	0.11	0.4	1
Number of Business in the Area	36.1	18.1	4	97
<b>Household Variables</b>				
Does the Household Treat Water for Adults (binary)	0.52	0.5	0	1
Does the Household Have a Child Aged 0-2 (binary)	0.21	0.41	0	1
Does the Household Own a Cell Phone (binary)	0.65	0.48	0	1
Does the Household Own a Fridge (binary)	0.18	0.38	0	1
House Many People Live in the Household	5.61	2.12	1	26

Notes (a) This table provides summary statistics for all of the control variables used, grouped by whether they are measured by area or by household (b) Variables noted with the tag (binary) are binary variables

**Table C: Dependent Variable Proportions by Treatment Group**

	Mean		Std. Dev	
<b>Dependent Variables:</b>	Treatment	Control	Treatment	Control
Latrine Usage	0.94	0.93	0.24	0.25
Treat Water for Adults	0.39	0.39	0.49	0.49
Filter Water for Children Ages 0-2	0.99	0.1	0.3	0.3
Have Insurance	0.2	0.22	0.4	0.41

Notes (a) All outcomes are treated as binary in the regression so numbers in the table refer to proportions of the population that exhibit this behavior

## References

- [1] Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman. "Six Randomized Evaluations of Microcredit: Introduction and Further Steps." *American Economic Journal*(2014): n. pag. Yale University. Web. 7 Dec. 2014.
- [2] Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." *Journal of Economic Literature* (2014): n. pag. *MIT Economics*. Web. 5 Dec. 2014.
- [3] "Primary Menu." *IFMR LEAD*. N.p., n.d. Web. 10 Dec. 2014
- [4] Yunus, Muhammad, and Karl Weber. *Creating a World without Poverty: Social Business and the Future of Capitalism*. New York: PublicAffairs, 2007. Print.
- [5] "United Nations Millennium Development Goals." *UN News Center*. UN, n.d. Web. 09 Dec. 2014.
- [6] Leatherman, Sheila, and Christopher Dunford. *Linking Health to Microfinance to Reduce Poverty*. Rep. 6th ed. Vol. 88. N.p.: n.p., n.d. *Bulletin of the World Health Organization - Linking Health to Microfinance to Reduce Poverty*. World Health Organization, June 2010. Web. 10 Dec. 2014.
- [7] Kim, Julia C. et al. "Understanding the Impact of a Microfinance-Based Intervention on Women's Empowerment." *American Journal of Public Health* 97.10 (2007): 1794-802. *American Journal of Public Health*. American Public Health Association. Web. 10 Dec. 2014.

# Probit Analysis of Non-Ferrous Metals as Leading Indicators: Testing Industrial Metal Prices in a Binary-Response Model of Pre-Crisis Thailand, 1997

Admund Tay

New York University

May 2015

## **Abstract**

There is a large body of scholarship in circulation that has examined the causal relationship between economic factors and output growth leading up to the Asian Financial Crisis of 1997, many of which proposing speculative attacks on the Thai economy as the primary cause. Yet despite its relevance to the study of the pre-crisis Thai economy, there are currently no studies examining the price movements of non-ferrous metals as indicators leading up to the event. This paper examines and compares the metal markets in the semiconductor-dominant Thai market in the mid-90s as an insight into the predictive value of metal prices over the short-, medium- and long-terms. Our study finds that, when assumptions are held, some non-ferrous metals appear to have great value as short- and long-term recession predictors.

## **1 Introduction**

In the run up to the crisis of 1997, Thailand's economy posted an average sub-10pct growth year-on-year. This was the period of the semiconductor boom, with large multinational companies funnelling large amounts of capital to newly-industrialized nations and their large stocks of cheap labor. One

of the beneficiaries of this was Thailand and its nascent semiconductor market. By the late 80s and early 90s, a good part of the Thai economy was propped up by large inflows of investment capital (Basu & Miroshnik, 2000) and along with that the taking off of its equity and real estate market. In quiet, speculative money grew and so did a bubble. Demand for non-ferrous metals used in the production of semiconductors also hit new highs. Yet beneath the veneer of the industrial activity, fragility was still in the bones of the Thai economy - characterized by a floundering export climate through the early and mid 90s and deep current account deficits that Thailand never seemed to pull out off. So when the semiconductor industry began to slump, the Thai economy was essentially left defenseless and vulnerable. In one fell swoop, the broad of speculators that gathered the speculative markets in Thailand, and certainly beyond, was taken out.

While the crisis itself was possibly unpreventable given the exogeneity of circumstances leading up to the event, there are merits in its study. The primary objective of this paper is to determine whether predictors exist, embedded in the pre-crisis metal markets, that could have helped provide a prescience to the circumstance and thereby mitigate the destructive fallout associated with the crisis.

## **2 Literature Review**

The body of literature currently in circulation that deals with modelling leading indicators for impending financial crises often deal with macro-financial data. One of the earlier attempts singularly studied the yield curve as a predictor of US recessions (Estrella & Mishkin, 1996) and provided a strong argument for the usage of the probit method as a useful tool to model event probabilities. Consequently, this method of analysis gained currency over the following years from modelling business cycles to dating (Bamara, 2006; Dopke, 1999; Katsuura & Layton, 2001). Similarly published at the end of the last century in the International Monetary Fund Staff Paper series was a paper studying a more comprehensive array of classical macro-economic indicators, like international reserves, domestic real interest rates, imports and exports (Kaminsky, Lizondo & Reinhart, 1998). This provided one of the first attempts at modelling a generalized warning system via a compos-

ite of indicators. Nonetheless, there has never been a paper that has covered an industry-specific approach to a recession.

### 3 Definitions and Methodology

Since a predictive model necessitates a demarcated timeframe, one of the dicier issues in this study is in regards to the definition of the start and end points of the crisis. There are, in theory, many unofficial definitions of a recession, with different econometric models pursuant to the authors choice of definition. In general, the accepted understanding of a recession is certainly any period of decline in economic activity (Claessens & Kose, 2009). Given the privy, the author of this paper has chosen to define the scope of a recession as at least two consecutive quarters of negative growth. And since this specifies a nominal, binary response, the analysis henceforth is therefore provided by the probit model. The primary merit being the provision of a clean and intuitive interface upon which conclusions may be drawn, the secondary being the long tradition of scholarship devoted to this process. In general, the non-heteroskedastic probit model assumes a given dependent variable characterized by a dichotomous outcome given by a vector  $I$  of  $n$  independent variables, transformed into a continuous variable and taking on a normal distribution:

$$P = pr(Y = 1|X) = \phi[\beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots \beta_n X_{ni}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{I=\sum \beta_i X_i} e^{-\frac{t^2}{2}} dt \quad (1)$$

, where phi provides the restriction on the model that results in a value between zero and 1, and where, by symmetry, when  $Y=0$ , the probability is  $1-P$ . The probability of a recession is thus given as the area under the normal distribution from negative infinity to estimated  $I$ .

The model, which assumes that variance in the error term is constant, ie homoskedastic, was chosen over the heteroskedastic probit, since: (i) each univariate probit regression deals with a single independent variable and how it relates to changes in the probability of our dependent

variable over time (as opposed to relating changes in the probability of our dependent variable with different variables cross-sectionally) , (ii) the heteroskedastic probit model has been found to be fragilely specified (Keele, 2006) and as a consequence is rarely used by researchers studying leading indicators. In sum, the lack of overt heteroskedastic tendencies in our data, and certainly of the form of research undertaken, do not necessitate the unorthodox usage of the modified probit.

## 4 Data

The data collected for the analysis is as follows: (i) From the second quarter of 1995 to the third quarter of 1997, the quarterly percentage change in GDP data of Thailand; (ii) From the second quarter of 1995 to the third quarter of 1997, the quarterly percentage price changes in the non-ferrous metals(excluding copper and nickel) spot market. From a static evaluation of each individual data point, we obtain a forecast some  $k$  quarters ahead (where  $k=2$  for short term,  $k=4$  for medium term and  $k=8$  for long term), and coded for 1=recession and 0=no recession if we witnessed one or the other  $k$  quarters from a given point. The choice of data collected for this study is intuitive; we avoided yearly and bi-yearly data due to its diminished precision, and monthly data due to the noise that short-term data often creates. Quarterly data thus ensures minimum noise for as much precision as we need.

A Variance Inflation Factor test was then conducted via least squares on all lagged periods with all variables (Tables 1.1-1.5) to ascertain variable suitability. Since VIF from the regressions of all independent variables on the remaining independent variables are less than 10, the variables selected do not violate the assumption of non-multicollinearity (When the R-squared between any given pair of independent variables is greater than 0.9, the VIF is greater than 10. This provides a common threshold of determining multicollinearity (Kopalle & Mela, 2002)).

What is endeavored to be achieved by this paper is now clear; a post-mortem analysis of the 1997 Asian Financial Crisis from the perspective of the industry (through metal prices as a proxy for industrial demand and supply). This comparison allows us to determine the first movers prior to the crisis and provide a useful reference for future pre-emptive hedges against crises of such nature.



Of course, it is noted that the non-ferrous metal market is not immune to its own speculation, however, with the removal of the more speculative non-ferrous metals - copper and nickel (Papp et al., 2008)- the price movements of the remaining metals are an acceptable proxy for industrial supply and demand. Further it is noted that price data, as opposed to inventory data, is often more accessible to the general public which would allow the results of this study to be referenced by the lay investor in other similar contexts.

## 5 Estimation

The analysis of estimated indicator performance is based on the z-statistic and the McFadden's R-squared, the latter being a pseudo R-squared adopted in logistical regressions, given by:

$$R_m^2 = 1 - \frac{\ln(L(M_{full}))}{\ln(L(M_{int}))} \quad (2)$$

Since the log-division component of the equation takes on a value between 0 and 1, the primary characteristic of this parameter is similar to its least squares counterpart, taking on values between zero and 1 as descriptors of how well a model specifies the investigated variable. However, unlike the OLS R-squared, the McFadden's R-squared accepts lower values as indicators of good fit, with anything between 0.2 and 0.4 considered as an *excellent* fit (McFadden, 1977). In our analysis, we first observe this pseudo R-squared value to determine if the independent variable is a good fit. If so, we then observe its marginal effect on the dependent variable.

In probit modelling, the sign of the coefficient only gives the direction of change and not the actual marginal effect on the dependent variable. To find the actual marginal effect, we assume the properties of the cumulative distribution function and derive the actual marginal effect via the first order condition of (1):

$$\frac{\partial P}{\partial X_i} = \frac{1}{2}[1 + \text{erf} \frac{X_i}{\sqrt{2}}]\beta X_i \quad (3)$$

which is the standard normal probability density function (of zero mean and unit variance) evaluated at a given value of an independent variable. Considering the form of our data set, we modified the coefficient of the independent variable from the preceding first order condition by a hundredth so that a one percent change in the independent variable corresponds directly with a one percent change in the dependent variable:

$$ME = \frac{1}{2}[1 + \text{erf} \frac{X_i}{\sqrt{2}}](\frac{1}{100})\beta X_i \quad (4)$$

Since there are various  $x$  values in a data set, one convention (which shall be applied here) is to take the mean of the marginal effects evaluated at each value of the given independent variable, summed then averaged. Given nine values for each variable in our data set:

$$AME = (\frac{1}{9}) \sum_{n=1}^9 [\frac{1}{2}[1 + \text{erf} \frac{X_{in}}{\sqrt{2}}](\frac{1}{100})\beta X_{in}] \quad (5)$$

This provides a closer approximation of the partial effects of an independent variable on the dependent variable vis-a-vis the marginal effect evaluated at the mean (Bartus, 2005). Therefore, we derive a modified average marginal effect (AME) of the independent variable which describes the direction and strength of an independent variable's relationship with the dichotomous dependent variable.

## 6 Results and Discussion

Salient findings are detailed in table 3, distilled from the probit regression results detailed in tables 2.1-2.13: Based on our findings, we observe that Tin is a fair short-term indicator, with a 1 percentage increase in the price of Tin resulting in a more than 2.4 percent decrease in the probability of a recession seen two quarters from that change, while Aluminium is an appropriate medium-term indicator, with a 1 percentage increase in the price of aluminium resulting in a more than 8 percent increase in the probability of a recession seen four quarters from that change.

We also observe that Lead and Zinc perform extremely well as long-term indicators, with a 1 percentage increase in the price of lead resulting in a 10.5 percent increase in the probability of a recession seen eight quarters from that change and a 1 percentage increase in the price of Zinc resulting in an approximately 8.9 percent decrease in the probability of a recession seen eight quarters from that change. We note that while there is, in theory, no upper or lower bound to the percentage changes in prices, the changes in probability of a recession are moving within the bounds of 0 and 1 given by the density function, therefore an eight to ten percent change should be interpreted significantly. Immediate conclusions, however, cannot be drawn from the above.

We assume that, *a priori*, all four metals each have exclusive (non-substitutable) functions in the semiconductor industry and low trading volatility, yet we observe contradictory movements in metal prices and their recession predictions: Aluminium and Lead being positive indicators and Tin and Zinc being negative. Taking the derivative of recession probability with regard to price change,  $p$ , the following holds:

$$\frac{\partial P}{\partial p(DD_{al})}, \frac{\partial P}{\partial p(DD_{lead})} > 0; \frac{\partial P}{\partial p(DD_{tin})}, \frac{\partial P}{\partial p(DD_{zinc})} < 0 \quad (6)$$

Where price is some positive function of demand. Given *a priori* assumptions, we further postulate that demand for all four metals and their recession predictions should move in tandem, *ceteris paribus*. Specifically, if price movements are antecedent to the recession, recession probability should be a negative function of price change, or simply:

$$P = -\alpha(p) \quad (7)$$

Some observations from our probit model however indicate otherwise.

Two (divergent) possibilities may be drawn: (i) price movements of the primary semiconductor metals in the industry is not evidenced to be indicative of the recession in 1997, but only appear to be leading indicators by incident; (ii) Both metals of the Tin-Zinc pair are leading indicators, with each having a substitutable function and thus a substitute metal within the pair. Assuming substitutability, we observe the following cross-price elasticity:

$$\frac{\partial DD_{m1}}{\partial p_{m2}} * \frac{p_{m2}}{DD_{m1}} > 0 \quad (8)$$

For equation (8) and thus conclusion (ii) to hold, metals, M1 and M2, when chosen, must be from a single pair of metals that moved in tandem with the change in recession probability (given by our probit), either Tin-Zinc or Aluminium-Lead. However, as established earlier, only Tin-Zinc moved in our postulated direction given by (7), therefore metals, M1 and M2, must individually be one from the tin-zinc pair. A theoretical consideration, however, is that, as substitutes, a runaway inflationary scenario exists, where an increase in the price of M1 leads to an increase in the demand for M2 and thus the price for M2 and thus the demand for M1, since it must hold that the cross price elasticities of M1 vis-a-vis M2 and M2 vis-a-vis M1 are simultaneously greater than zero. This is nonetheless mitigated by the inelastic and ponderous nature of physical stock contracts. Accordingly, from our results, zinc prices moved first as the leading indicator  $k=8$  quarters ahead of the recession, with tin prices lagging as the short term leading indicator  $k=2$  quarters ahead. Therefore, the condensed form of conclusions (i) and (ii) is that if Tin-Zinc substitutability exists, then both from the Tin-Zinc pair of metals are true leading indicators, otherwise there are none.

## 7 Conclusion

Based on our in-sample probit analysis, we may tentatively conclude that, where assumptions are held, tin and zinc do appear to have value as short- and long-term indicators respectively. This model may be extended upon further investigations made to determine the substitutability of tin and zinc in the semiconductor industry. Lastly, similar tests of substitutability may then be conducted on various input pairs for all out-of-sample, binary-specified, pre-crisis economies to investigate the extendibility of this model.

## 8 Appendix

**Table 1.1**

Variance Inflation Factors  
Date: 05/10/15 Time: 12:25  
Sample: 1995Q3 1997Q3  
Included observations: 9

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.000933	7.049139	NA
TIN	0.170870	2.222556	2.024090
LEAD	0.104771	4.667676	4.632852
ZINC	0.031309	2.312666	1.649993
SET	0.079650	9.371834	5.009581

**Table 1.3**

Variance Inflation Factors  
Date: 05/10/15 Time: 12:32  
Sample: 1995Q3 1997Q3  
Included observations: 9

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001224	4.341502	NA
ALUMINIUM	0.141944	1.797726	1.761152
TIN	0.285475	1.743927	1.588201
LEAD	0.206347	4.317509	4.285298
SET	0.135809	7.504821	4.011596

**Table 1.2**

Variance Inflation Factors  
Date: 05/10/15 Time: 12:32  
Sample: 1995Q3 1997Q3  
Included observations: 9

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.000317	4.844807	NA
ALUMINIUM	0.041817	2.279522	2.233146
ZINC	0.015410	2.300958	1.641640
LEAD	0.046491	4.186833	4.155596
SET	0.025402	6.041743	3.229528

**Table 1.4**

Variance Inflation Factors  
Date: 05/10/15 Time: 12:33  
Sample: 1995Q3 1997Q3  
Included observations: 9

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001980	6.270089	NA
ALUMINIUM	0.596150	6.739556	6.602443
TIN	1.080928	5.894205	5.367873
ZINC	0.258978	8.019609	5.721666
SET	0.082911	4.089707	2.186095

**Table 1.5**

Variance Inflation Factors  
Date: 05/10/15 Time: 12:34  
Sample: 1995Q3 1997Q3  
Included observations: 9

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.000820	2.329714	NA
ALUMINIUM	0.563737	5.714328	5.598072
ZINC	0.212015	5.886677	4.199906
TIN	0.734633	3.591797	3.271063
LEAD	0.103130	1.727039	1.714154

**Table 2.2**

Dependent Variable: KTHAI4  
Method: ML - Binary Probit (Quadratic hill climbing)  
Date: 05/10/15 Time: 12:37  
Sample: 1995Q3 1997Q3  
Included observations: 9  
Convergence achieved after 4 iterations  
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.361367	0.512257	0.705440	0.4805
ALUMINIUM	16.36667	11.76689	1.390909	0.1643
McFadden R-squared	0.242226	Mean dependent var	0.555556	
S.D. dependent var	0.527046	S.E. of regression	0.480900	
Akaike info criterion	1.485567	Sum squared resid	1.618857	
Schwarz criterion	1.529395	Log likelihood	-4.685053	
Hannan-Quinn criter.	1.390987	Deviance	9.370107	
Restr. deviance	12.36531	Restr. log likelihood	-6.182654	
LR statistic	2.995201	Avg. log likelihood	-0.520561	
Prob(LR statistic)	0.083512			
Obs with Dep=0	4	Total obs	9	
Obs with Dep=1	5			

**Table 2.1**

Dependent Variable: KTHAI2  
Method: ML - Binary Probit (Quadratic hill climbing)  
Date: 05/10/15 Time: 12:36  
Sample: 1995Q3 1997Q3  
Included observations: 9  
Convergence achieved after 3 iterations  
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.449662	0.441672	-1.018091	0.3086
ALUMINIUM	-1.849731	7.659237	-0.241503	0.8092
McFadden R-squared	0.005173	Mean dependent var	0.333333	
S.D. dependent var	0.500000	S.E. of regression	0.533188	
Akaike info criterion	1.710887	Sum squared resid	1.990027	
Schwarz criterion	1.754715	Log likelihood	-5.698991	
Hannan-Quinn criter.	1.616307	Deviance	11.39798	
Restr. deviance	11.45726	Restr. log likelihood	-5.728628	
LR statistic	0.059273	Avg. log likelihood	-0.633221	
Prob(LR statistic)	0.807649			
Obs with Dep=0	6	Total obs	9	
Obs with Dep=1	3			

**Table 2.3**

Dependent Variable: KTHAI2  
Method: ML - Binary Probit (Quadratic hill climbing)  
Date: 05/10/15 Time: 12:38  
Sample: 1995Q3 1997Q3  
Included observations: 9  
Convergence achieved after 4 iterations  
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.484841	0.476577	-1.017341	0.3090
LEAD	-8.012578	6.992303	-1.145914	0.2518
McFadden R-squared	0.134740	Mean dependent var	0.333333	
S.D. dependent var	0.500000	S.E. of regression	0.502035	
Akaike info criterion	1.545945	Sum squared resid	1.764276	
Schwarz criterion	1.589773	Log likelihood	-4.956753	
Hannan-Quinn criter.	1.451365	Deviance	9.913506	
Restr. deviance	11.45726	Restr. log likelihood	-5.728628	
LR statistic	1.543749	Avg. log likelihood	-0.550750	
Prob(LR statistic)	0.214061			
Obs with Dep=0	6	Total obs	9	
Obs with Dep=1	3			

**Table 2.4**

Dependent Variable: KTHAI4  
Method: ML - Binary Probit (Quadratic hill climbing)  
Date: 05/10/15 Time: 12:39  
Sample: 1995Q3 1997Q3  
Included observations: 9  
Convergence achieved after 3 iterations  
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.245561	0.466436	0.526463	0.5986
LEAD	-9.055339	6.472955	-1.398950	0.1618
McFadden R-squared	0.181036	Mean dependent var	0.555556	
S.D. dependent var	0.527046	S.E. of regression	0.499758	
Akaike info criterion	1.569638	Sum squared resid	1.748309	
Schwarz criterion	1.613465	Log likelihood	-5.063369	
Hannan-Quinn criter.	1.475058	Deviance	10.12674	
Restr. deviance	12.36531	Restr. log likelihood	-6.182654	
LR statistic	2.238570	Avg. log likelihood	-0.562597	
Prob(LR statistic)	0.134605			
Obs with Dep=0	4	Total obs	9	
Obs with Dep=1	5			

**Table 2.6**

Dependent Variable: KTHAI2  
Method: ML - Binary Probit (Quadratic hill climbing)  
Date: 05/10/15 Time: 12:41  
Sample: 1995Q3 1997Q3  
Included observations: 9  
Convergence achieved after 5 iterations  
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.154925	1.000050	-1.154867	0.2481
TIN	-29.53373	27.59637	-1.070204	0.2845
McFadden R-squared	0.203049	Mean dependent var	0.333333	
S.D. dependent var	0.500000	S.E. of regression	0.483610	
Akaike info criterion	1.458986	Sum squared resid	1.637154	
Schwarz criterion	1.502813	Log likelihood	-4.565435	
Hannan-Quinn criter.	1.364406	Deviance	9.130870	
Restr. deviance	11.45726	Restr. log likelihood	-5.728628	
LR statistic	2.326385	Avg. log likelihood	-0.507271	
Prob(LR statistic)	0.127197			
Obs with Dep=0	6	Total obs	9	
Obs with Dep=1	3			

**Table 2.5**

Dependent Variable: KTHAI8  
Method: ML - Binary Probit (Quadratic hill climbing)  
Date: 05/10/15 Time: 12:41  
Sample: 1995Q3 1997Q3  
Included observations: 9  
Convergence achieved after 5 iterations  
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.291496	0.585618	-0.497758	0.6187
LEAD	20.98002	12.63783	1.660097	0.0969
McFadden R-squared	0.517059	Mean dependent var	0.444444	
S.D. dependent var	0.527046	S.E. of regression	0.383789	
Akaike info criterion	1.107968	Sum squared resid	1.031058	
Schwarz criterion	1.151796	Log likelihood	-2.985857	
Hannan-Quinn criter.	1.013388	Deviance	5.971714	
Restr. deviance	12.36531	Restr. log likelihood	-6.182654	
LR statistic	6.393594	Avg. log likelihood	-0.331762	
Prob(LR statistic)	0.011453			
Obs with Dep=0	5	Total obs	9	
Obs with Dep=1	4			

**Table 2.7**

Dependent Variable: KTHAI4  
Method: ML - Binary Probit (Quadratic hill climbing)  
Date: 05/10/15 Time: 12:42  
Sample: 1995Q3 1997Q3  
Included observations: 9  
Convergence achieved after 3 iterations  
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.070441	0.506177	-0.139163	0.8893
TIN	-14.13409	13.23688	-1.067781	0.2856
McFadden R-squared	0.114303	Mean dependent var	0.555556	
S.D. dependent var	0.527046	S.E. of regression	0.520791	
Akaike info criterion	1.661324	Sum squared resid	1.898561	
Schwarz criterion	1.705151	Log likelihood	-5.475956	
Hannan-Quinn criter.	1.566744	Deviance	10.95191	
Restr. deviance	12.36531	Restr. log likelihood	-6.182654	
LR statistic	1.413396	Avg. log likelihood	-0.608440	
Prob(LR statistic)	0.234493			
Obs with Dep=0	4	Total obs	9	
Obs with Dep=1	5			

**Table 2.8**

Dependent Variable: KTHAI8  
Method: ML - Binary Probit (Quadratic hill climbing)  
Date: 05/10/15 Time: 12:43  
Sample: 1995Q3 1997Q3  
Included observations: 9  
Convergence achieved after 3 iterations  
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.208191	0.454062	-0.458509	0.6466
TIN	-5.039239	11.47074	-0.439313	0.6604
McFadden R-squared	0.016434	Mean dependent var	0.444444	
S.D. dependent var	0.527046	S.E. of regression	0.558175	
Akaike info criterion	1.795788	Sum squared resid	2.180916	
Schwarz criterion	1.839616	Log likelihood	-6.081048	
Hannan-Quinn criter.	1.701208	Deviance	12.16210	
Restr. deviance	12.36531	Restr. log likelihood	-6.182654	
LR statistic	0.203213	Avg. log likelihood	-0.675672	
Prob(LR statistic)	0.652140			
Obs with Dep=0	5	Total obs	9	
Obs with Dep=1	4			

**Table 2.10**

Dependent Variable: KTHAI8  
Method: ML - Binary Probit (Quadratic hill climbing)  
Date: 05/10/15 Time: 12:44  
Sample: 1995Q3 1997Q3  
Included observations: 9  
Convergence achieved after 5 iterations  
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.373384	0.526360	0.709370	0.4781
ZINC	-16.99772	12.56831	-1.352426	0.1762
McFadden R-squared	0.343597	Mean dependent var	0.444444	
S.D. dependent var	0.527046	S.E. of regression	0.446190	
Akaike info criterion	1.346292	Sum squared resid	1.393602	
Schwarz criterion	1.390120	Log likelihood	-4.058314	
Hannan-Quinn criter.	1.251712	Deviance	8.116627	
Restr. deviance	12.36531	Restr. log likelihood	-6.182654	
LR statistic	4.248681	Avg. log likelihood	-0.450924	
Prob(LR statistic)	0.039281			
Obs with Dep=0	5	Total obs	9	
Obs with Dep=1	4			

**Table 2.9**

Dependent Variable: KTHAI2  
Method: ML - Binary Probit (Quadratic hill climbing)  
Date: 05/10/15 Time: 12:43  
Sample: 1995Q3 1997Q3  
Included observations: 9  
Convergence achieved after 3 iterations  
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.824082	0.554226	-1.486906	0.1370
ZINC	6.880069	5.605508	1.227377	0.2197
McFadden R-squared	0.145962	Mean dependent var	0.333333	
S.D. dependent var	0.500000	S.E. of regression	0.476748	
Akaike info criterion	1.531659	Sum squared resid	1.591022	
Schwarz criterion	1.575486	Log likelihood	-4.892463	
Hannan-Quinn criter.	1.437079	Deviance	9.784927	
Restr. deviance	11.45726	Restr. log likelihood	-5.728628	
LR statistic	1.672328	Avg. log likelihood	-0.543607	
Prob(LR statistic)	0.195947			
Obs with Dep=0	6	Total obs	9	
Obs with Dep=1	3			

**Table 2.11**

Dependent Variable: KTHAI2  
Method: ML - Binary Probit (Quadratic hill climbing)  
Date: 05/10/15 Time: 12:45  
Sample: 1995Q3 1997Q3  
Included observations: 9  
Convergence achieved after 4 iterations  
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.940924	0.703930	-1.336673	0.1813
SET	-5.357046	5.338137	-1.003542	0.3156
McFadden R-squared	0.096266	Mean dependent var	0.333333	
S.D. dependent var	0.500000	S.E. of regression	0.503882	
Akaike info criterion	1.594924	Sum squared resid	1.777276	
Schwarz criterion	1.638751	Log likelihood	-5.177156	
Hannan-Quinn criter.	1.500344	Deviance	10.35431	
Restr. deviance	11.45726	Restr. log likelihood	-5.728628	
LR statistic	1.102943	Avg. log likelihood	-0.575240	
Prob(LR statistic)	0.293621			
Obs with Dep=0	6	Total obs	9	
Obs with Dep=1	3			



Table 2.12

Dependent Variable: KTHAI4  
 Method: ML - Binary Probit (Quadratic hill climbing)  
 Date: 05/10/15 Time: 12:45  
 Sample: 1995Q3 1997Q3  
 Included observations: 9  
 Convergence achieved after 3 iterations  
 Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.102582	0.560933	-0.182877	0.8549
SET	-2.983791	4.686595	-0.636665	0.5243
McFadden R-squared	0.033044	Mean dependent var	0.555556	
S.D. dependent var	0.527046	S.E. of regression	0.551614	
Akaike info criterion	1.772967	Sum squared resid	2.129945	
Schwarz criterion	1.816795	Log likelihood	-5.978352	
Hannan-Quinn criter.	1.678387	Deviance	11.95670	
Restr. deviance	12.36531	Restr. log likelihood	-6.182654	
LR statistic	0.408604	Avg. log likelihood	-0.664261	
Prob(LR statistic)	0.522679			
Obs with Dep=0	4	Total obs	9	
Obs with Dep=1	5			

Table 2.13

Dependent Variable: KTHAI8  
 Method: ML - Binary Probit (Quadratic hill climbing)  
 Date: 05/10/15 Time: 12:47  
 Sample: 1995Q3 1997Q3  
 Included observations: 9  
 Convergence achieved after 2 iterations  
 Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.344964	0.603746	0.571373	0.5677
SET	5.956997	5.204483	1.144590	0.2524
McFadden R-squared	0.116776	Mean dependent var	0.444444	
S.D. dependent var	0.527046	S.E. of regression	0.523797	
Akaike info criterion	1.657926	Sum squared resid	1.920540	
Schwarz criterion	1.701754	Log likelihood	-5.460667	
Hannan-Quinn criter.	1.563346	Deviance	10.92133	
Restr. deviance	12.36531	Restr. log likelihood	-6.182654	
LR statistic	1.443975	Avg. log likelihood	-0.606741	
Prob(LR statistic)	0.229497			
Obs with Dep=0	5	Total obs	9	
Obs with Dep=1	4			

Table 3.	K=2	K=4	K=8
<i>Aluminium</i>			
McFadden's R-Squared	0.005173	0.242226	*
Average Marginal Effect	***	0.08128	*
z-statistic	-0.241503	1.390909	*
<i>Lead</i>			
McFadden's R-Squared	0.134740	0.181036	0.517059
Average Marginal Effect	***	***	0.10545
z-statistic	-1.145914	-1.398950	1.660097
<i>Tin</i>			
McFadden's R-Squared	0.203049	0.114303	0.016434
Average Marginal Effect	-0.02427	***	***
z-statistic	-1.070204	-1.067781	-0.439313
<i>Zinc</i>			
McFadden's R-Squared	0.145962	**	0.343597
Average Marginal Effect	***	**	-0.08856
z-statistic	1.227377	**	-1.352426

\* Perfectly predicts binary response failure

\*\* Perfectly predicts binary response success

\*\*\* Average Marginal Effect not needed since variable does not satisfy the criteria of a good fit based on McFadden's R-squared

## References

- [1] Bamara, H. (2006). The predictive power of yield curve in forecasting U.S. recessions. Simon Fraser University
- [2] Bartus, T. (2005). Estimation of marginal effects using `margeff`. The Stata Journal Volume 5 Number 3: pp. 309-329
- [3] Basu, D.R., & Miroshnik, V. (2000) Japanese foreign investments, 1970-1998: perspectives and analyses, pp 112
- [4] Claessens, S., & Kose, M.A. (2009) What is a recession. Finance and Development 2009; Vol. 46: 1
- [5] Dopke, J. (1999) Predicting Germany's recessions with leading indicators: evidence from probit models. Kiel Institute of World Economics
- [6] Estrella, A., & Mishkin, F.S. (1996) The yield curve as a predictor of US recessions. Current Issues in Economics and Finance 1996; Vol 2: 7
- [7] Kaminsky, G., & Lizondo, S., & Reinhart, C.M. (1998) Leading indicators of currency crises. IMF Staff Papers 1998; Vol 45: 1
- [8] Katsuura, M., & Layton, A.P. (2001) Comparison of regime switching, probit and logit models in dating and forecasting US business cycles. International Journal of Forecasting 2001; Vol. 17: pp 403-417
- [9] Keele, L. (2006) Ambivalent about ambivalence: A re-examination of heteroskedastic probit models. Retrieved from [http://www.academia.edu/157614/\\_Ambivalent\\_about\\_Ambivalence\\_A\\_Re-examination\\_of\\_Heteroskedastic\\_Probit\\_Models\\_](http://www.academia.edu/157614/_Ambivalent_about_Ambivalence_A_Re-examination_of_Heteroskedastic_Probit_Models_)
- [10] McFadden, D. (1977) Quantitative methods for analysing travel behaviour of individuals. Published in D. Hensher and P. Stopher (eds.), Behavioural Travel Modelling, Croom Helm, 1978:307

- [11] Papp J.F. et al. (2008) Factors that influence the price of Al, Cd, Co, Cu, Fe, Ni, Pb, Rare Earth Elements, and Zn. USGS Open-File Report 2008-1356; pp: 21

# The Reaction Function of The Federal Reserve Post 2008 Financial Crisis

Michel Cassard  
Princeton University

June 21, 2015

## **Abstract**

This paper uses vector autoregressive analysis to show a change to the Federal Reserve's reaction function post 2008. The model's prediction for the stance of Fed policy show that a Taylor Rule identification structure for the Federal Funds Rate no longer holds. Stress to financial markets during the crisis has meant that additional consideration of financial conditions is needed to accurately reflect Fed decision making since the Great Recession.

# 1 Introduction

The Federal Reserve has a dual mandate: maintaining stable inflation and maximum employment. Since this mandate was declared in 1977, the Fed has used its primary policy tool, the federal funds rate (FF), to try to best achieve these goals. In 2008, a collapse in the housing market triggered a crisis in the US financial crisis which spread to the rest of the world. The Fed reacted to the severity of the crisis by sharply cutting interest rates to zero, where they have remained since.

Economists analyzing the Fed's policy response have found that the Taylor Rule, which is based on inflation and unemployment deviation from their target, is a good predictor of monetary policy. However, when one uses that framework to analyze the response of the Fed to the 2008 crisis, one finds that the speed and severity of the monetary accommodation is not fully explained. This paper posits that the dislocation experienced by financial markets at the outset of the crisis caused financial conditions to become of central importance to the Fed insofar as they reflected the extent of impairment of transmission channels of monetary policy to the real economy. As such, this paper believes that in order to effectively understand Fed policy post 2008, a measure of financial conditions must be included in its reaction function.

This paper will use vector autoregressive analysis of the relationship between key economic variables to shed light on the Fed's reaction function pre and post Great Recession. The analysis will first examine a framework that solely considers the three variables of the Taylor Rule in predicting monetary policy. The paper finds that while this rule is effective for the pre-crisis period, it fails to predict Fed policy at the zero lower bound after 2008.

However, when financial variables are included, an unconstrained estimation finds that the FF should have gone negative after 2008 to provide adequate accommodation to the US economy. This result is not captured by a VAR model that does not include financial conditions. The difficulty of imposing negative rates led the Fed to turn to unconventional measures of easing such as forward guidance and quantitative easing. As the FF loses its information content on monetary policy at the zero lower bound, this paper compares its predictions to an unconstrained Shadow Interest Rate that includes measures of unconventional policy ([Wu and Xia, 2014](#)).

This paper establishes that when financial conditions are included in the VAR model, the prediction of the Fed’s monetary policy stance improves significantly. However, it also shows that between 2008 and 2012, the Fed should have engaged in even more quantitative easing than was the case and having done so, should have started tightening earlier by 2013. The results also suggest that to compensate for an overtly tight policy during the first part of the recession, the Fed has had to maintain QE for a longer period of time after 2013 to provide the same cumulative easing effect on the US economy for the entire 2008-2015 period.

The paper proceeds as follows: Section 2 discusses the relevant economic literature. Section 3 lays out the econometric framework of analysis. Section 4 describes the data. Section 5 presents the main results and Section 6 concludes.

## 2 Literature Review

### Vector Autoregression

The vector autoregression (VAR) framework was pioneered by Sims (1980) in response to fundamental flaws in previous techniques of econometric analysis. “A VAR is a  $n$ -equation,  $n$ -variable linear model in which each variable is in turn explained by its own lagged values, plus current and past values of the remaining  $n - 1$  variables.” (Stock and Watson, 2001) The technique allows for the model to capture rich relationships and dynamics between variables over a model comprised of many simultaneous univariate equations.

Stock and Watson (2001) “Vector Autoregressions” uses a small model VAR with three variables in order to describe the relationships between key macroeconomic indicators in the US. Stock and Watson’s paper is a survey article of different VAR forms: reduced, recursive and structural. They use inflation, the unemployment rate and the federal funds rate to assess the power of VARs in various applications of econometric analysis.

Stock and Watson use data from 1960-2001 to analyze their VAR’s power at the four key tasks of econometric analysis: data description, forecasting, structural inference and policy analysis. They find that the three variables all have significant predictive power over one another and that the

error in forecasts for one variable is largely explained by the other two: “For example, at the 12 quarter horizon, 75% of the error in the forecast of the Federal Funds rates is attributed to the inflation and unemployment shocks in the recursive VAR.” (Stock and Watson, 2001) For these reasons, they find that the recursive VAR is an effective tool for US monetary policy description.

Stock and Watson use the VAR for multistep ahead forecasting both out of sample (from 2000 on) and pseudo out-of-sample (1960-2001). For pseudo out-of-sample forecasts, they find that their three variable VAR improves on both a random walk and univariate autoregression in its predictions for the variable movements.

While the VAR gives strong results for these two tasks, Stock and Watson find it lacking in its use for structural inference and policy analysis. Structural inference and policy analysis require using a structural vector autoregression (SVAR) to identify how the variables are related to one another. These identifying assumptions are based on economic theory and “even modest changes in the assumed rule resulted in substantial changes in these impulse responses. In other words, the estimates of structural impulse responses hinge on detailed institutional knowledge of how the Fed set interest rates.” (Stock and Watson, 2001) Stock and Watson use a Taylor Rule as their identifying assumption, which will be discussed in greater detail later in this paper.

What Stock and Watson find is that the VAR “shocks” that would be used for structural inference largely just reflect omitted variables from the model. “Because of omitted variables, the VAR mistakenly viewed and labeled these increases in interest rates as monetary shocks, which led to biased impulse responses.” (Stock and Watson, 2001) In reality, identifying the monetary rule is near impossible and while Stock and Watson find that the inflation rate, unemployment rate and federal funds rate are the most important variables to consider, there was not a stable rule between them that the Fed followed. Frequent changes in the policy rule meant that the shocks were not real monetary shocks, but just the SVAR falling to the Lucas Critique (Lucas, 1976) of incorrect identifying assumptions.

## Taylor Rule

Stock and Watson find that inflation, unemployment and lagged values of the Fed Funds Rate are the most effective way to capture movements in the interest rate. The theory behind this comes from the dual mandate of the Fed: price stability and maximum employment. John Taylor created the Taylor Rule ([Taylor, 1993](#)) as a way of summarizing how the Fed responds to deviations from target output and inflation. In their model, Stock and Watson use Okun's Law ([Okun, 1963](#)) to replace the output gap with the employment gap. Stock and Watson write their Taylor Rule as:

$$R_t = r^* + 1.5(\bar{\pi}_t - \pi^*) - 1.25(\bar{u}_t - u^*) + \text{lagged values of } R, \pi, u + \epsilon_t$$

This equation forms the interest rate equation in the three variable SVAR.  $R_t$  is the interest rate,  $r^*$  is the neutral interest rate of the economy,  $\bar{u}_t$  is the average unemployment rate over the last year,  $u^*$  is the target employment rate or NAIRU,  $\bar{\pi}_t$  is the average inflation rate over the last year,  $\pi^*$  is the target inflation rate of 2%, and  $\epsilon_t$  is the error term.

However, the Taylor Rule does not account for interest rates that are constrained by the zero lower bound. After a period of time at the zero bound, the FF loses its information content about the true stance of monetary policy.

## Shadow Rate

Cynthia Wu and Fan Dora Xia at the Federal Reserve Bank of Atlanta have used Factor Augmented VAR (FAVAR) to construct a measure of monetary policy that is unconstrained by the zero lower bound. Their series continues from the FF when it hits the zero lower bound at the end of 2008 and allows the stance of monetary policy to be negative to reflect unconventional monetary policy measures such as forward guidance and quantitative easing.

[Wu and Xia \(2014\)](#) find that "that the shadow rate calculated by our model exhibits similar dynamic correlations with macro variables of interest in the period since July 2009 as the fed funds rate did in data prior to the Great Recession. This result gives us a tool for measuring the effects of monetary policy at the ZLB, and offers an important insight to the empirical macro literature



where people use the effective federal funds rate in vector autoregressive (VAR) models to study the relationship between monetary policy and the macroeconomy... The evident structural break in the effective fed funds rate [at the zero lower bound] prevents researchers from getting meaningful information out of a VAR during and even post the ZLB. In contrast, the continuation of our series allows researchers to update their favorite VAR using the shadow rate for the ZLB period.” (Wu and Xia, 2014)

Xia and Wu’s Shadow Rate looks at a number of different facets of the US economy in order to gauge their Shadow Rate prediction. They consider the FF as well as measures of forward guidance, the size of the Fed’s balance sheet and the length of time that interest rates have been at the zero lower bound. Using all of these, they use the Shadow Rate to “construct a new measure for the monetary policy stance when the effective federal funds rate is bounded below by zero, and employed this measure to study unconventional monetary policy’s impact on the real economy.” (Wu and Xia, 2014)

It will be important to assess the FF predictions of this paper’s different VAR specifications in relation to both the constrained FF and true measures of monetary policy that include unconventional policies. While Xia and Wu’s Shadow Rate is not the only one that has been created, it has been widely cited as effective as a measure of the monetary policy stance since 2009 and is consistently updated on the Federal Reserve of Atlanta’s website.

### 3 Methodology

This paper will use recursive vector autoregressions with different specifications to analyze the relationship between key macroeconomic variables in the United States. Using the findings, it will comment on the Federal Reserve’s monetary policy reaction function by seeing which set of variables best describe and predict movements in the Federal Funds Rate and the Shadow Interest Rate.

### Three Variable VAR

In Stock and Watson's structural vector autoregression, they use the original Taylor Rule as the identifying assumption for the interest rate rule. They let the rest of the coefficients be determined by a Choleski Decomposition. The SVAR equation therefore looks like:

$$\begin{pmatrix} In_t \\ Un_t \\ FF_t \end{pmatrix} = \begin{pmatrix} . & 0 & 0 \\ . & . & 0 \\ -1.5 & 1.25 & 1 \end{pmatrix} * \begin{pmatrix} In_{t-1} \\ Un_{t-1} \\ FF_{t-1} \end{pmatrix} + 3 \text{ additional lags} + \begin{pmatrix} u_t^{In} \\ u_t^{Un} \\ u_t^{FF} \end{pmatrix}$$

Where  $In_t$  is the inflation rate,  $Un_t$  is the unemployment rate and  $FF_t$  is the federal funds rate with their lags. A  $(.)$  means that the coefficient is estimated from the data using a Choleski Decomposition.  $u_t$  is the vector of error terms.

When this paper re-estimates this SVAR, first over the original period and then updating the sample to 2008, the parameters show instability over their sample and do not hold as a rule for Fed policy in setting rates. Using this SVAR finds large monetary policy shocks over the sample that in reality are probably errors in the identifying assumptions.

While the specific identifying assumptions and coefficients of this Taylor Rule do not hold, the economic theory that the Fed primarily bases its Fed Funds Rate on the inflation rate and unemployment rate is sound. Different iterations and coefficients for the rule have been considered, such as that by [Yellen \(2012\)](#). However, over such a long time period with a different macroeconomic paths and different monetary regimes, it would not make sense to consider one fixed interest rate rule for the entire period.

Furthermore, while the Fed considers a wide variety of factors in setting policy, one of the limitations of the VAR technique is that increasing the number of variables or lags significantly increases the number of parameters that need to be estimated. Therefore, it makes sense to continue to use the three variables of Stock and Watson's VAR as a benchmark for analysis of US monetary policy, but not to use their rigid identifying structure. This paper will focus on unstructured recursive VARs to allow the movements in the data over the sample period to identify the relationship between variables.

While theory will not form strict identifying coefficients for the variables, it will inform which variables will be chosen and the ordering of the recursive VAR. When considering Fed policy, the two most important variables to consider are inflation and unemployment. These are the two focuses of the Fed's dual mandate and theoretically inform all of their policy decisions. This three variable VAR should offer a strong description and prediction of their policy moves as the original benchmark VAR against which future specifications can be compared.

With this economic relationship in mind, the first VAR that will be used is a three variable VAR with inflation, unemployment and the federal funds rate as variables. While adding variables will always improve the ability of the model to fit the data, the cost of additional parameters is high. So, this paper focuses on a small model for the economy that captures key relationships and co-movements.

When using a recursive VAR, the variables have to be ordered according to how they causally affect one another in order to orthogonalize the error terms. "In the jargon of VARs, this is equivalent to estimating the reduced form and then computing the Choleski factorization of the reduced form VAR covariance matrix" (Lütkepohl, 2007). Figure 4 shows the relationship between inflation and unemployment, highlighting how lagged inflation causes unemployment but there is little correlation between lagged unemployment and inflation. As these two variables inform the Fed's policy setting, the VAR will be structured:

$$\begin{pmatrix} In_t \\ Un_t \\ FF_t \end{pmatrix} = \begin{pmatrix} . & 0 & 0 \\ . & . & 0 \\ . & . & . \end{pmatrix} * \begin{pmatrix} In_{t-1} \\ Un_{t-1} \\ FF_{t-1} \end{pmatrix} + \text{additional lags} + \begin{pmatrix} u_t^{In} \\ u_t^{Un} \\ u_t^{FF} \end{pmatrix}$$

Here, Inflation(In), Unemployment (Un) and the Fed Funds Rate (FF) are all allowed to dynamically explain moves one another. 3 or 4 lags will be considered and  $u$  measures the error term. Different tests must be conducted to assess the validity and strength of the VAR models being used. Granger causality statistics test whether the lagged value of one variable is useful when predicting another. This is important in determining whether there adding a variable improves the VAR's estimation by effectively predicting the other variables.

It is necessary to test the stability of the VAR model in order to ensure that it does not have unit roots. Otherwise, the relationship between the variables would not be stable over time and the model would be ineffective.

Stock and Watson finds that four lags is appropriate for their model. However, as this model updates the data sample until 2015, it re-estimates the Bayesian Information Criteria (BIC) and Akaike Information Criteria (AIC) to check the required number of lags. The new model finds that the optimal number of lags is 3, so this paper will consider both VARs with 3 and 4 lags.

After assessing the descriptive properties of the three variable VAR until 2008, the paper will turn to its forecasting ability. Using the sample truncated at September 2008, to reflect when the FF and Shadow Rate first diverge as unconventional policy begins, the model will perform pseudo out-of-sample forecasts and compare its prediction of the Fed Funds Rate with the Federal Reserve's actual policy moves and the Shadow Rate. This is the fundamental question of this paper: whether the same model that effectively describes and predicts Fed policy until 2008 still holds during the period of the Great Recession. If the model loses its predictive power, it will be necessary to change or add variables to better reflect Fed policy considerations.

## **Four Variable VAR: Adding Financial Conditions**

The financial crisis began at the end of 2007 and spread from the subprime housing market through the entire financial sector. The Fed cut the Fed Funds Rate from 5.25% in July 2007 to the zero lower bound by the end of 2008 (Figure 1). This move was not predicted by just falling inflation or rising unemployment. It is likely that the Federal Reserve, which states that it looks at financial markets but does not set rates based on them, used its policy rate to provide additional monetary easing on account of the financial sector meltdown.

In order to address this departure of Fed policy from being predominately informed by inflation and unemployment, I will incorporate the conditions of financial markets into the VAR. There are numerous ways to do this using different variables or indexes. Therefore, a number of different VAR specifications will be considered.

$$\begin{pmatrix} In_t \\ Un_t \\ FCI_t \\ FF_t \end{pmatrix} = \begin{pmatrix} . & 0 & 0 & 0 \\ . & . & 0 & 0 \\ . & . & . & 0 \\ . & . & . & . \end{pmatrix} * \begin{pmatrix} In_{t-1} \\ Un_{t-1} \\ FCI_{t-1} \\ FF_{t-1} \end{pmatrix} + \text{additional lags} + \begin{pmatrix} u_t^{In} \\ u_t^{Un} \\ u_t^{FCI} \\ u_t^{FF} \end{pmatrix}$$

This equation represents the new, general 4 variable VAR which includes the variables of the Taylor Rule, Inflation (In), Unemployment (Un) and the Fed Funds Rate (FF) but also adds a fourth variable, a measure of financial conditions (FCI).

There are numerous measures of financial market conditions that can be included. Financial stress indicators vary significantly in complexity. Some are straightforward such as the VIX, which is a measure of the implied volatility of the S&P 500. However, the variability of the VIX is too great to meaningfully provide a measure of financial market conditions over the 2008-2015 period.

Bond spreads offer an effective manner of determining market stress. Using the spread between High Yield debt (BB rated) over T-bills, we can see how the market is pricing risk. During times of financial stress, these spreads will increase dramatically as can be seen in Figure 1.

Beyond these straight forward measures, there are also financial conditions indexes that have been created to more accurately reflect financial market stress. Two notable ones are the St Louis Fed Financial Stress Indicator (Figure 1) and the Chicago Fed's National Financial Conditions Index (Figure 1). These use a number of different financial market variables and weightings to construct an indicator for conditions in financial markets. The particular merit of these indexes is that they include not only a bond spreads, but a wide variety of other financial market indicators such as measures of confidence, liquidity and credit.

The St Louis Fed Financial Stress Indicator (SLFSI) "is constructed from seven interest rate series, six yield spreads and five other indicators" (Kliesen et al., 2010). A summary of the factors considered can be seen in Figure 3. This index adds significantly to just a measure of bond spreads as it contains information both about market risk perception and about liquidity risk. In it, zero is viewed as normal market functioning and values above are interpreted as above-average financial market stress.

The Chicago Fed National Financial Conditions Index (NFCI) ([Brave and Butters, 2012](#)), is calculated from an even wider range of financial market data. The three broad sub-areas considered in the index are risk, credit and leverage. This means that not only does the NFCI contain information about risk and liquidity as does the St Louis Fed FSI, but also about bank lending to consumers and mortgage markets. These would both be of key interest to the Fed when setting rates, who would be trying to maximize the efficacy of their policy stance. The bank lending and mortgage markets can be thought of as the multiplier effect of Fed rate changes, and so when in the Financial Crisis banks stopped lending and the mortgage market froze, FF moves would lose their potency as a policy tool. Thus, this index is likely the most reflective of Fed considerations of financial markets during the Great Recession and will offer the best prediction for FF moves.

For all of the new four variable VARs, granger causality statistics, lag tests and stationarity tests will be computed. Each model will be estimated over the pre crisis sample, 1997-2008 and the results compared to the original three variable VAR. Then, the VARs will be used for pseudo out-of-sample forecasting from September 2008 onwards and the results will be compared to one another, to the original VAR and to the FF and Shadow Rate.

Economic interpretation of the different policy predictions given by the VARs will shed light on whether the Fed's monetary policy reaction function has changed through the Great Recession and whether it was due to new consideration of financial market stress.

## 4 Data

The VARs used in this paper are small models that focus on key macroeconomic and financial variables. The three variable VAR uses the same data series as Stock and Watson, except taken on a monthly rather than quarterly basis to add more detail to the VAR estimation:

Monthly data on Inflation (Consumer Price Index for All Urban Consumers: All Items, Percent Change from Year Ago, Monthly, Seasonally Adjusted), the Unemployment Rate (Civilian Unemployment Rate, Percent, Monthly, Seasonally Adjusted) and the Fed Funds Rate (Effective Federal Funds Rate, Percent, Monthly, Seasonally Adjusted).

Monthly data for the Shadow Interest Rate ([Wu and Xia, 2014](#)) is found on the Federal Reserve Bank of Atlanta’s website.

The financial variables that are considered for the four variable VAR are the High Yield BB Spreads (BofA Merrill Lynch US High Yield BB Option-Adjusted Spread), The St Louis Fed Financial Stress Index and The Chicago Fed National Financial Conditions Index. All are taken on a monthly basis. The data is found on the Federal Reserve Economic Database (FRED).

The data series are each available for different sample lengths and so effective comparison limits the sample to Jan 1997 - Feb 2015.

## 5 Results

Multiple different VAR specifications are estimated in this paper. The results of pseudo out-of-sample forecasts for each from the period of 2008-2015 are shown in Figure 5. Before considering and comparing the forecasts, it is important to first address the framework of each of the VARs.

### Three Variable VAR Results

The three variable VAR using the original variables of [Stock and Watson \(2001\)](#) meets the stability criteria of having all eigenvalues less than one. Granger causality tests show that at the 10% level, all of the variables predictively cause the other two (Figure 7). The information criteria show that 3 lags are optimal under AIC and 2 are optimal under BIC. The estimation will be more robust with too many rather than too few lags, so the model uses the VAR with 3 lags.

The three variable VAR is estimated until September 2008 and then recursively predicts the FF until the start of 2015. Figure 5 shows that this model clearly misses the Fed’s policy of cutting rates to the zero lower bound and keeping them there until 2015. Instead, the model prescribes for rates to be lowered to 0.8% in October of 2009 and then to slowly to be increased to almost 1.5% by March of 2015. This policy prediction would provide a far less expansionary monetary policy than the Fed’s own policy, with a significant difference between the forecasted interest rate and both the FF and the Shadow Rate.

## Four Variable VAR Results

All four of the key VARs that included financial conditions variables met the stability conditions, each having all eigenvalues less than one. The AIC and BIC show that for the VARs including the BB Spread and St Louis Fed FSI, the optimal lag length is 6. For the VAR including the Chicago Fed FCI, the optimal lag length is 5.

In each of the four variable VAR specifications post 2008 crisis, the financial conditions indicator variable granger causes the Shadow Rate (as the Fed Funds Rate loses its information content) at the 1% level for the St Louis FSI (Figure 11) and Chicago FCI (Figure 13) and at the 5% level for the BB Spread (Figure 9). This notably contrasts to pre 2008, where the financial variables do not granger cause the FF at any significant level (Figures 8, 10, 12). This result highlights that the financial variables were not a significant part of the Fed's decision making process before the financial crisis, but were part of their reaction function following 2008.

While each of the predictions for the FF is different depending on the financial variable included, they all show a similar trend (Figure 6). All three of the VARs predict that the FF should have gone negative at the end of 2008 or start of 2009 to provide greater monetary accommodation to the US economy. The model also shows that after a period of negative interest rates, the FF should have been increased between 2010 and 2013 to above the zero bound towards settling a level more consistent with predictions for the US neutral interest rate, around 3.7% by Fed estimates.

Specifically, the four variable VAR including the BB spread shows interest rates falling to lows of -0.5% before eventually rising above the zero bound in April of 2010 and eventually settling at a rate of around 3.5% between 2013 and 2015 (Figure 6).

When we use a four variable VAR model that includes the St Louis FSI, the prediction is even more extreme. This model predicts that the interest rate would have fallen to -1.67% by November 2009, an even sharper and deeper cut to the FF than with the model using the BB spread. The model predicts that interest rates would rise to above zero in October of 2010 and then to also find a roughly stable level of around 3.7% over the 2013-2015 period (Figure 6).

Finally, the most extreme prediction come when the Chicago FCI is included in the model.



Interest rates are predicted to have fallen for longer to -2.6% in October 2010, when the prediction including the St Louis FSI was just crossing back above the zero bound. The Chicago FCI prediction only rises above zero in August of 2012 but quickly rises to 5% by March of 2015 (Figure 6).

This improvement in the predictions as the financial variable changes follows an intuitive explanation. The Chicago Fed FCI contains the broadest range of financial indicators that the Fed would be considering: not just risk as in the BB Spread variable, or liquidity and risk as in the St Louis FSI variable, but also lending conditions by banks and mortgage providers that would be key transmission mechanisms through which Fed policy works. With credit channels frozen in 2008, the Fed would be paying the closest attention to the Chicago FCI and the variables it is comprised of.

It is important to note that these predictions have sizable confidence bands for the predictions. In the case of the original Taylor Rule variable VAR, the confidence band includes negative interest rate values. While the 95% confidence interval is a large error margin, this uncertainty does highlight a key potential pitfall of the model. Within the 95% interval, the prediction for the FF with the BB and St Louis FSI could have never gone negative and only gone slightly negative with the Chicago Fed FCI.

## Comparison between Predictions

In order to effectively compare the predictions of the paper and assess which offers the best estimation of Fed policy during the Recession, both quantitative and qualitative analysis must be conducted. Looking at the predictions overlaid with the FF and Shadow Rate gives the most comparative image of the results (Figure 6). As Figure 6 shows, the three variable VAR prediction completely misses the zero lower bound as well as the unconventional policy measures of the Fed that drove the Shadow Rate negative. The prediction including the BB spread improves on this, but still does not predict the level of accommodation that the Fed provided. The estimate of the VAR including the St Louis Fed and Chicago Fed most closely mirror the movements of the Shadow Rate.

This is not to say that their predictions perfectly track the moves in the Shadow Rate or the FF. The intuition of the results is corroborated by the work of [Woodford \(2012\)](#) and [Reifschnei-](#)

der and Williams (2000). This paper’s FF predictions lie initially below both the actual FF and Shadow Rate, meaning that monetary policy in 2008-2013 was not as accommodative as the model would have prescribed. To compensate, expansionary monetary policy must be continued for longer than predicted to have the same overall monetary effect on the US economy. “When policy is constrained by the effective lower bound, policymakers can achieve superior economic outcomes by committing to keep the federal funds rate lower for longer than would be called for.” (Reifschneider and Williams, 2000) (Hakkio and Kahn, 2014)

In other words, the intuition for the Fed keeping rates at the zero lower bound and continuing QE for longer than the VAR model predicts is so that the cumulative deviation of the monetary policy stance below prediction offsets the cumulative deviation of policy above prediction. The significant improvement of the VAR prediction including the Chicago FCI is evident when assessing the mean error predictions of the different models from the actual FF. The model including the Chicago FCI finds the smallest overall prediction error by far (-0.39) over the 2008-2015 period, compared with triple the forecast error for the VAR comprised of only the Taylor Rule variables (-0.93) (Figure 2).

## 6 Conclusion

This paper shows that the Fed’s reaction function changed after the 2008 crisis. The inclusion of a variable for financial conditions into the VAR model in addition to the traditional inflation and unemployment significantly improves our prediction of Fed policy post 2008. We find that the Chicago FCI has the best predictive power because of its breadth and inclusion of measures of market risk, liquidity and credit availability. With this additional variable, our VAR prediction shows that the FF should have fallen to -2.6% by the end of 2010.

A zero bound interest rate meant that the Fed had to resort to unconventional policy to provide the required additional monetary stimulus to the economy. This paper also finds that since the Fed did not react with the speed and intensity prescribed by the model, they have had to keep

rates at the zero lower bound for longer. By doing so, they were able to provide the same total accommodation effect on the economy over the entire 2008-2015 period. The cumulative effect is now almost fully realized. Employment figures are improving and the financial system is healing, although inflation still has not picked up. One possible extension to this paper would be to use forward looking expectations for the variables in order to better reflect Fed decision making.

## Acknowledgements

I am grateful to Professor Iqbal Zaidi for his guidance and support as my advisor and to Olivier Darmouni for insightful discussion. I also thank Professor Mark Watson for his encouragement and mentorship and the readers of early drafts for their comments.

## References

- BRAVE, S. AND A. BUTTERS (2012): “Diagnosing the financial system: financial conditions and financial stress,” *International Journal of Central Banking*, 8, 191–239.
- HAKKIO, C. S. AND G. A. KAHN (2014): “Evaluating monetary policy at the zero lower bound,” *Economic Review*, 5–32.
- KLIESEN, K. L., D. C. SMITH, ET AL. (2010): “Measuring financial market stress,” *Economic Synopses*.
- LUCAS, R. E. (1976): “Econometric policy evaluation: A critique,” in *Carnegie-Rochester conference series on public policy*, Elsevier, vol. 1, 19–46.
- LÜTKEPOHL, H. (2007): *New introduction to multiple time series analysis*, Springer Science & Business Media.
- OKUN, A. M. (1963): “Potential GNP: its measurement and significance,” Tech. rep., American Statistical Association.
- REIFSCHNEIDER, D. AND J. C. WILLIAMS (2000): “Three lessons for monetary policy in a low-inflation era,” *Journal of Money, Credit and Banking*, 936–966.
- SIMS, C. A. (1980): “Macroeconomics and reality,” *Econometrica: Journal of the Econometric Society*, 1–48.
- STOCK AND WATSON (2001): “Vector Autoregression,” *JER*.
- TAYLOR, J. B. (1993): “Discretion versus policy rules in practice,” in *Carnegie-Rochester conference series on public policy*, Elsevier, vol. 39, 195–214.
- WOODFORD, M. (2012): “Methods of policy accommodation at the interest-rate lower bound,” in *The Changing Policy Landscape: 2012 Jackson Hole Symposium*, Federal Reserve Bank of Kansas City.

WU, J. C. AND F. D. XIA (2014): “Measuring the macroeconomic impact of monetary policy at the zero lower bound,” Tech. rep., National Bureau of Economic Research.

YELLEN, J. L. B. (2012): “The Economic Outlook and Monetary Policy,” Tech. rep., Board of Governors of the Federal Reserve System.

## Appendix

Table 1: Summary statistics

<b>Variable</b>	<b>Mean</b>	<b>Std. Dev.</b>
Fed Funds Rate	2.573	2.363
Shadow Rate	2.097	2.94
Inflation	2.286	1.213
Unemployment Rate	6.06	1.782
Inflation	2.286	1.213
High Yield BB Spread	3.846	1.988
Chicago Fed FCI	-0.352	0.594
St Louis Fed FCI	-0.036	1.069
N		219

Table 2: Mean Error Between VAR Predictions and the Fed Funds Rate

<b>Variable</b>	<b>Mean</b>	<b>Std. Dev.</b>
VAR with Taylor Variables Mean Error	-0.931	0.325
VAR with BB Spread Mean Error	-2.17	1.581
VAR with St Louis Fed FSI Mean Error	-2.056	2.343
VAR with Chicago Fed FCI Mean Error	-0.398	2.697
N		80

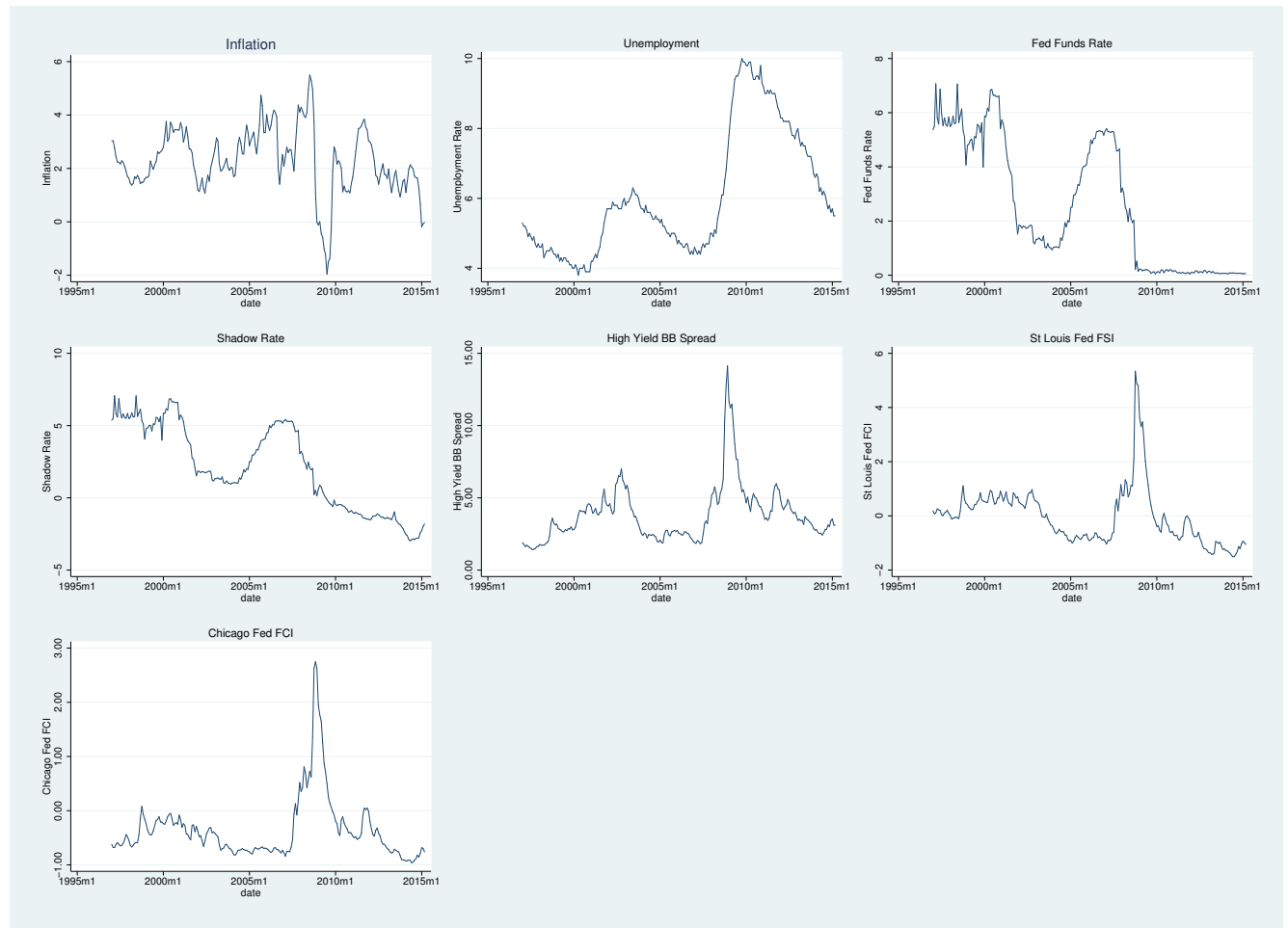


Figure 1: Data Summary



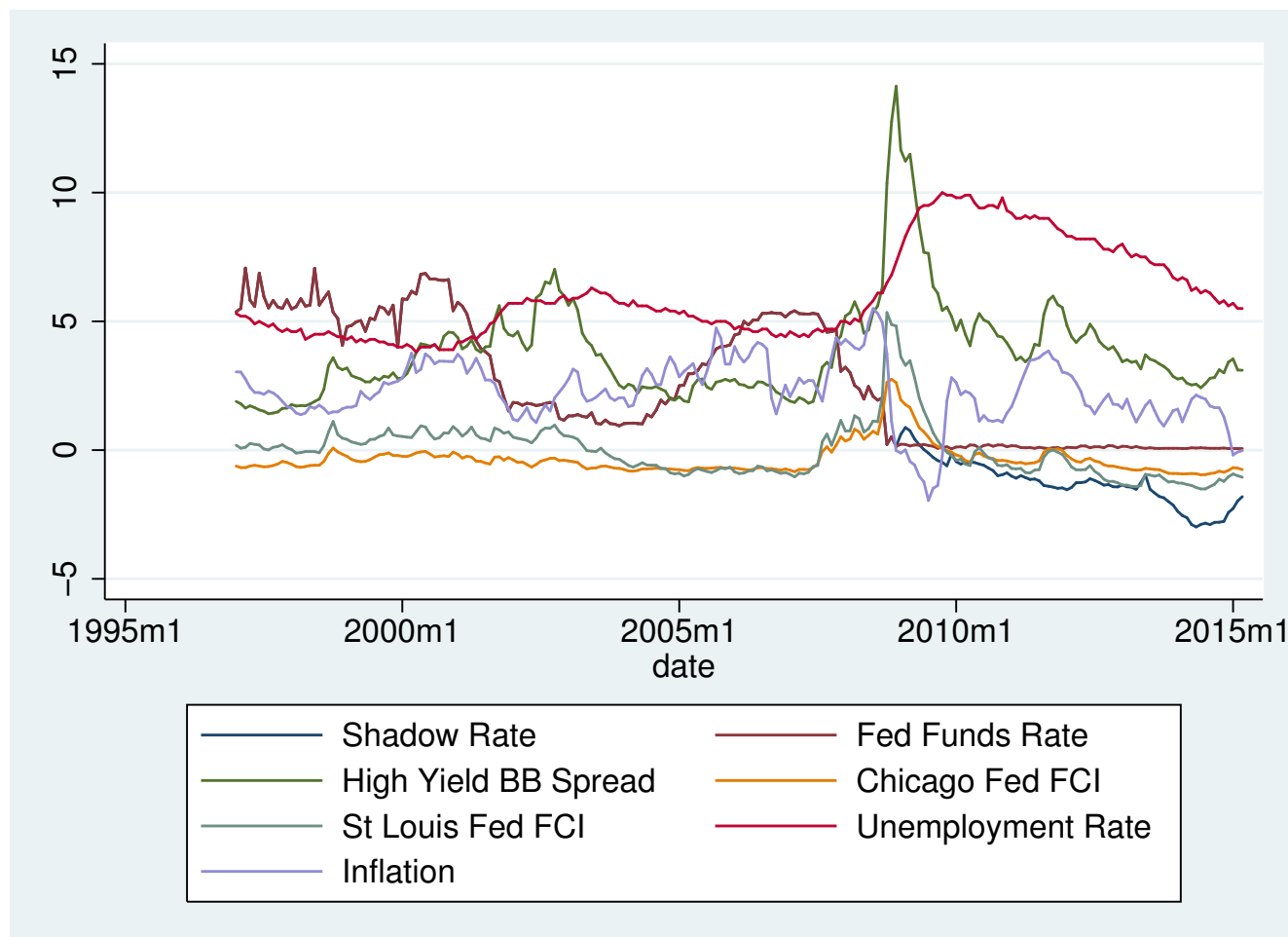


Figure 2: Data

### List of Data Series Used to Construct the St. Louis Fed Financial Stress Index (STLFSI)

STLFSI Key		
	Label in Chart	Series
Interest Rates	FedFund	Effective federal funds rate
	Treas2y	2-year Treasury
	Treas10y	10-year Treasury
	Treas30y	30-year Treasury
	BAA	Baa-rated corporate
	Mlynch_HighYld_MasterII	Merrill Lynch High-Yield Corporate Master II Index
	Mlynch_BBBAA	Merrill Lynch Asset-Backed Master BBB-rated
Yield Spreads	YieldCurve_10yr3mo	Yield curve: 10-year Treasury minus 3-month Treasury
	Corp_CRS	Corporate Baa-rated bond minus 10-year Treasury (corporate credit risk spread)
	HighYield_CRS	Merrill Lynch High-Yield Corporate Master II Index minus 10-year Treasury (high-yield credit risk spread)
	LiborOIS_3mo	3-month London Interbank Offering Rate–Overnight Index Swap spread (3-month LIBOR-OIS spread)
	TED	3-month Treasury-Eurodollar spread (TED spread)
	CPS_3mo	3-month commercial paper minus 3-month Treasury bill (commercial paper spread (3-month))
Other Indicators	EMBI	J.P. Morgan Emerging Markets Bond Index Plus
	VIX	Chicago Board Options Exchange Market Volatility Index (VIX)
	Mlynch_BMVI_1mo	Merrill Lynch Bond Market Volatility Index (1-month)
	BIR_10yr	10-year nominal Treasury yield minus 10-year Treasury Inflation Protected Security yield (breakeven inflation rate (10-year))
	SP500_FI	S&P 500 Financials Index

Note from: <http://research.stlouisfed.org/fred2/series/STLFSI>

As of 07/15/2010 the Vanguard Financial Exchange-Traded Fund series has been replaced with the S&P 500 Financials Index. This change was made to facilitate a more timely and automated updating of the FSI. Switching from the Vanguard series to the S&P series produced no meaningful change in the index.

Figure 3: St Louis Fed Financial Stress Index Components

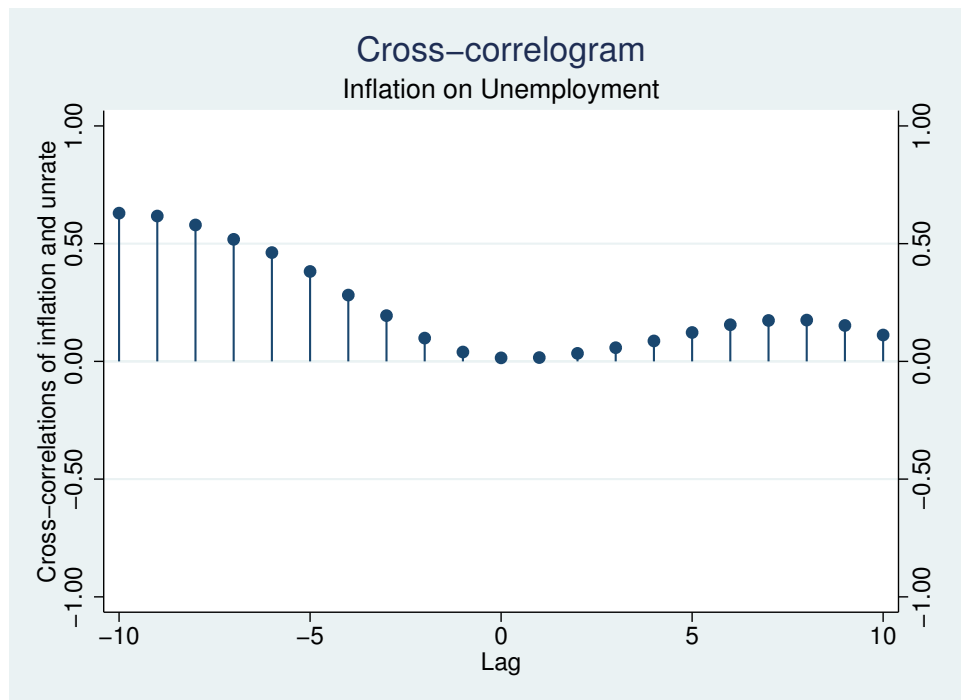


Figure 4: Cross Correlation of Inflation on Unemployment

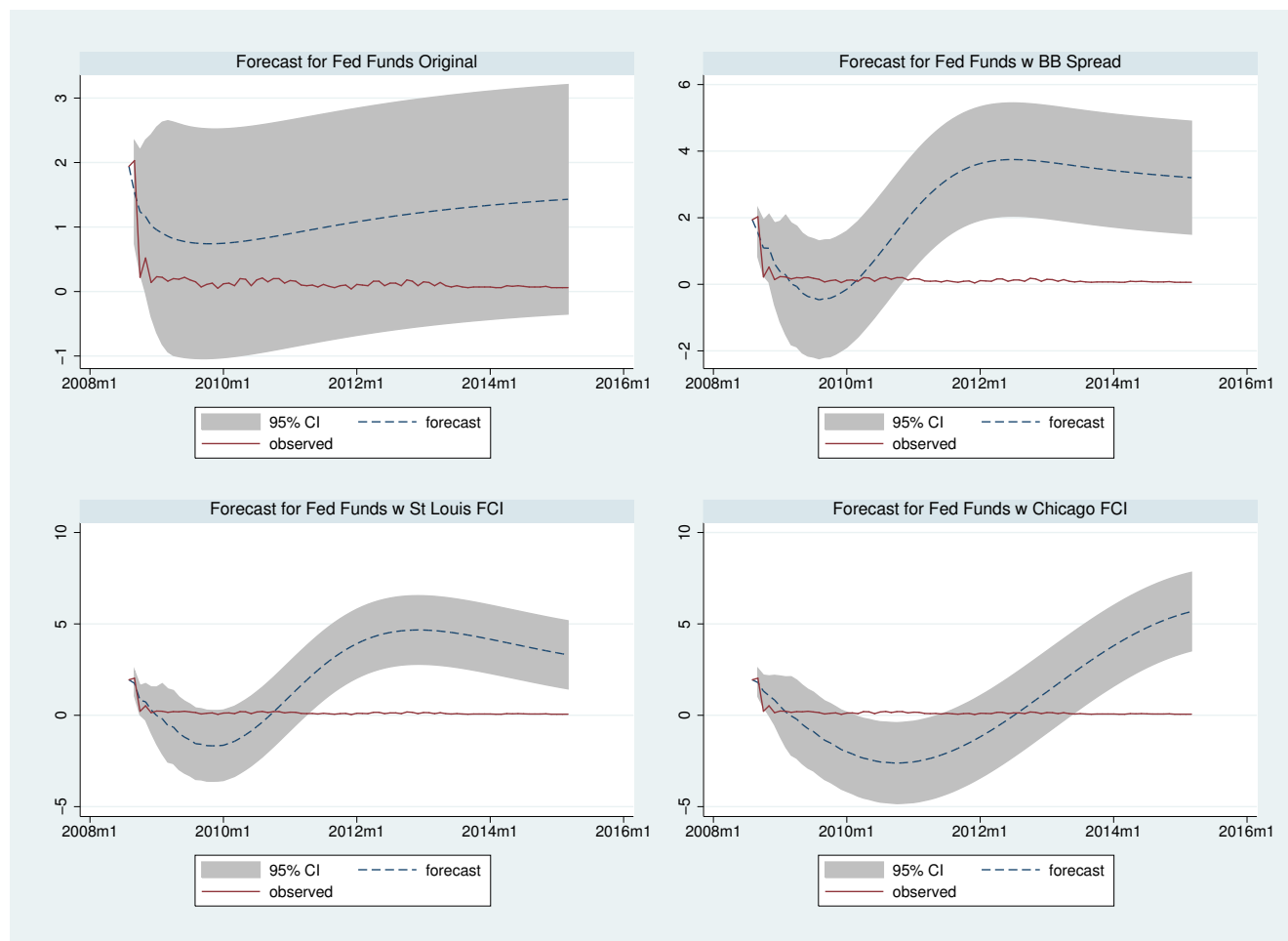


Figure 5: Fed Funds Rate Forecasts for Different VAR Specifications

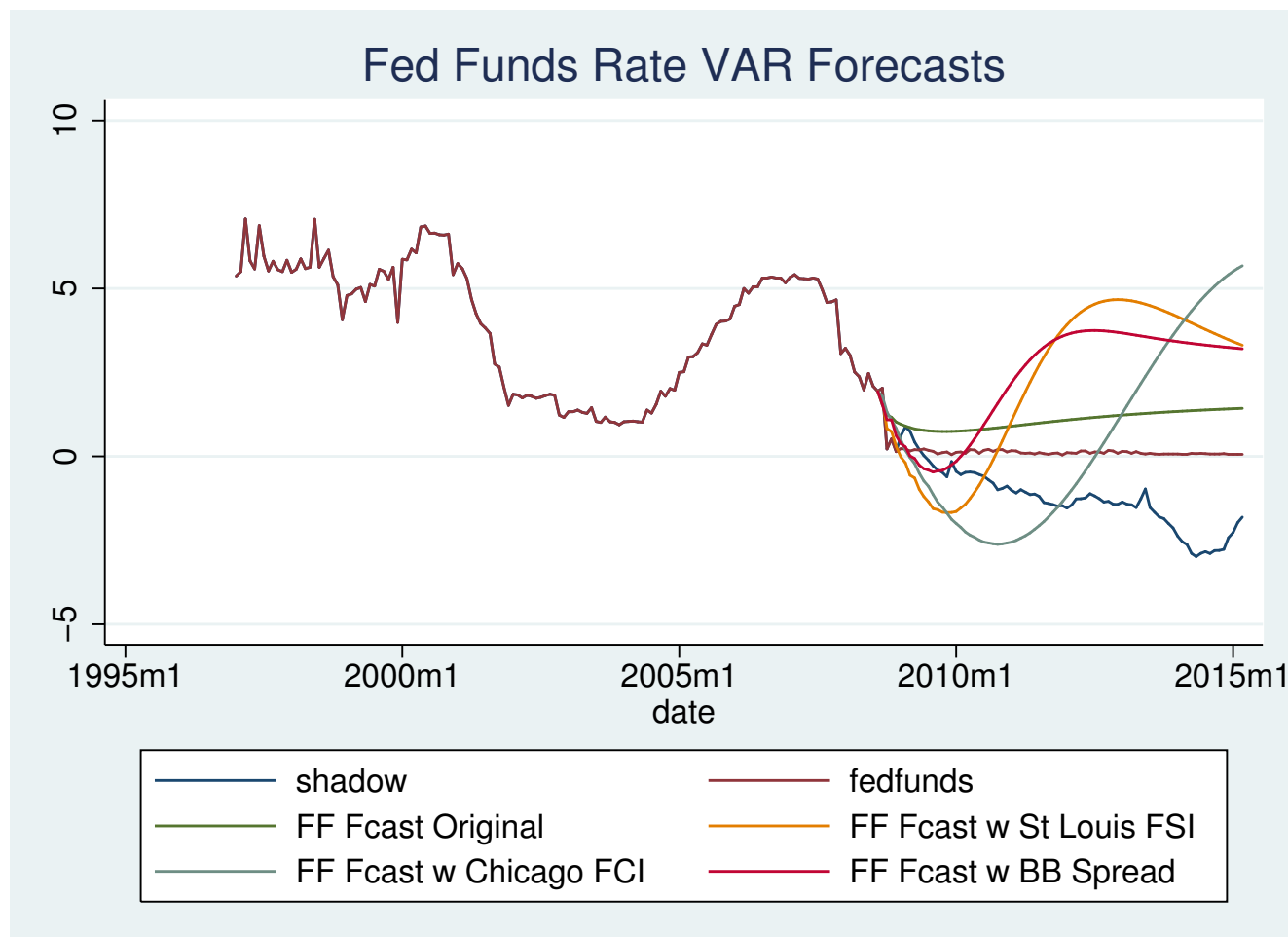


Figure 6: Fed Funds Rate Forecasts

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
inflation	unrate	<b>24.456</b>	<b>5</b>	<b>0.000</b>
inflation	fedfunds	<b>5.5179</b>	<b>5</b>	<b>0.356</b>
inflation	ALL	<b>30.264</b>	<b>10</b>	<b>0.001</b>
unrate	inflation	<b>20.796</b>	<b>5</b>	<b>0.001</b>
unrate	fedfunds	<b>18.913</b>	<b>5</b>	<b>0.002</b>
unrate	ALL	<b>40.24</b>	<b>10</b>	<b>0.000</b>
fedfunds	inflation	<b>8.1837</b>	<b>5</b>	<b>0.146</b>
fedfunds	unrate	<b>13.618</b>	<b>5</b>	<b>0.018</b>
fedfunds	ALL	<b>17.43</b>	<b>10</b>	<b>0.065</b>

Figure 7: Granger Causality Tests for Original Three Variable VAR Pre 2008

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
inflation	unrate	<b>9.6344</b>	<b>6</b>	<b>0.141</b>
inflation	bbspread	<b>5.708</b>	<b>6</b>	<b>0.457</b>
inflation	shadow	<b>6.5559</b>	<b>6</b>	<b>0.364</b>
inflation	ALL	<b>17.163</b>	<b>18</b>	<b>0.512</b>
unrate	inflation	<b>9.6844</b>	<b>6</b>	<b>0.139</b>
unrate	bbspread	<b>11.204</b>	<b>6</b>	<b>0.082</b>
unrate	shadow	<b>15.428</b>	<b>6</b>	<b>0.017</b>
unrate	ALL	<b>40.458</b>	<b>18</b>	<b>0.002</b>
bbspread	inflation	<b>12.838</b>	<b>6</b>	<b>0.046</b>
bbspread	unrate	<b>17.774</b>	<b>6</b>	<b>0.007</b>
bbspread	shadow	<b>1.5515</b>	<b>6</b>	<b>0.956</b>
bbspread	ALL	<b>32.66</b>	<b>18</b>	<b>0.018</b>
shadow	inflation	<b>3.1928</b>	<b>6</b>	<b>0.784</b>
shadow	unrate	<b>24.12</b>	<b>6</b>	<b>0.000</b>
shadow	bbspread	<b>10.857</b>	<b>6</b>	<b>0.093</b>
shadow	ALL	<b>43.284</b>	<b>18</b>	<b>0.001</b>

Figure 8: Granger Causality Tests for Four Variable VAR with BB Spread Pre 2008

Granger causality Wald tests					
Equation	Excluded	chi2	df	Prob > chi2	
inflation	unrate	23.358	6	0.001	
inflation	bbspread	10.019	6	0.124	
inflation	shadow	11.601	6	0.071	
inflation	ALL	49.797	18	0.000	
unrate	inflation	23.041	6	0.001	
unrate	bbspread	16.855	6	0.010	
unrate	shadow	19.069	6	0.004	
unrate	ALL	72.362	18	0.000	
bbspread	inflation	11.165	6	0.083	
bbspread	unrate	22.222	6	0.001	
bbspread	shadow	12.802	6	0.046	
bbspread	ALL	82.091	18	0.000	
shadow	inflation	3.3405	6	0.765	
shadow	unrate	12.613	6	0.050	
shadow	bbspread	7.804	6	0.253	
shadow	ALL	28.39	18	0.056	

Figure 9: Granger Causality Tests for Four Variable VAR with BB Spread Post 2008

Granger causality Wald tests					
Equation	Excluded	chi2	df	Prob > chi2	
inflation	unrate	14.924	6	0.021	
inflation	slfci	13.866	6	0.031	
inflation	fedfunds	12.968	6	0.044	
inflation	ALL	25.99	18	0.100	
unrate	inflation	13.25	6	0.039	
unrate	slfci	17.297	6	0.008	
unrate	fedfunds	20.93	6	0.002	
unrate	ALL	47.779	18	0.000	
slfci	inflation	12.648	6	0.049	
slfci	unrate	22.268	6	0.001	
slfci	fedfunds	9.2783	6	0.159	
slfci	ALL	42.353	18	0.001	
fedfunds	inflation	3.8974	6	0.691	
fedfunds	unrate	28.669	6	0.000	
fedfunds	slfci	24.158	6	0.000	
fedfunds	ALL	59.562	18	0.000	

Figure 10: Granger Causality Tests for Four Variable VAR with St Louis FSI Pre 2008

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
inflation	unrate	15.492	6	0.017
inflation	slfci	16.04	6	0.014
inflation	shadow	11.64	6	0.071
inflation	ALL	58.479	18	0.000
unrate	inflation	25.114	6	0.000
unrate	slfci	22.413	6	0.001
unrate	shadow	10.472	6	0.106
unrate	ALL	81.105	18	0.000
slfci	inflation	10.249	6	0.115
slfci	unrate	21.506	6	0.001
slfci	shadow	49.752	6	0.000
slfci	ALL	122.92	18	0.000
shadow	inflation	4.5695	6	0.600
shadow	unrate	10.455	6	0.107
shadow	slfci	12.647	6	0.049
shadow	ALL	34.369	18	0.011

Figure 11: Granger Causality Tests for Four Variable VAR with St Louis FSI Post 2008

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
inflation	unrate	7.1889	3	0.066
inflation	cfci	8.3066	3	0.040
inflation	fedfunds	5.8794	3	0.118
inflation	ALL	15.61	9	0.075
unrate	inflation	1.2654	3	0.737
unrate	cfci	8.0184	3	0.046
unrate	fedfunds	14.877	3	0.002
unrate	ALL	33.648	9	0.000
cfci	inflation	1.1317	3	0.769
cfci	unrate	5.7943	3	0.122
cfci	fedfunds	5.3547	3	0.148
cfci	ALL	11.897	9	0.219
fedfunds	inflation	1.8843	3	0.597
fedfunds	unrate	24.073	3	0.000
fedfunds	cfci	9.8233	3	0.020
fedfunds	ALL	40.918	9	0.000

Figure 12: Granger Causality Tests for Four Variable VAR with Chicago FCI Pre 2008



Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
inflation	unrate	<b>16.999</b>	<b>6</b>	<b>0.009</b>
inflation	cfci	<b>15.248</b>	<b>6</b>	<b>0.018</b>
inflation	shadow	<b>12.664</b>	<b>6</b>	<b>0.049</b>
inflation	ALL	<b>57.336</b>	<b>18</b>	<b>0.000</b>
unrate	inflation	<b>28.26</b>	<b>6</b>	<b>0.000</b>
unrate	cfci	<b>23.387</b>	<b>6</b>	<b>0.001</b>
unrate	shadow	<b>7.654</b>	<b>6</b>	<b>0.265</b>
unrate	ALL	<b>82.638</b>	<b>18</b>	<b>0.000</b>
cfci	inflation	<b>17.528</b>	<b>6</b>	<b>0.008</b>
cfci	unrate	<b>32.884</b>	<b>6</b>	<b>0.000</b>
cfci	shadow	<b>72.305</b>	<b>6</b>	<b>0.000</b>
cfci	ALL	<b>175.09</b>	<b>18</b>	<b>0.000</b>
shadow	inflation	<b>4.7651</b>	<b>6</b>	<b>0.574</b>
shadow	unrate	<b>10.222</b>	<b>6</b>	<b>0.116</b>
shadow	cfci	<b>19.54</b>	<b>6</b>	<b>0.003</b>
shadow	ALL	<b>42.877</b>	<b>18</b>	<b>0.001</b>

Figure 13: Granger Causality Tests for Four Variable VAR with Chicago FCI Post 2008

Comparative Advantage is grateful for the support of the Stanford Department of Economics and the Stanford Economics Association. The website that houses the Stanford Economics Association and Comparative Advantage can be found at <http://economics.stanford.edu/node/3015>.