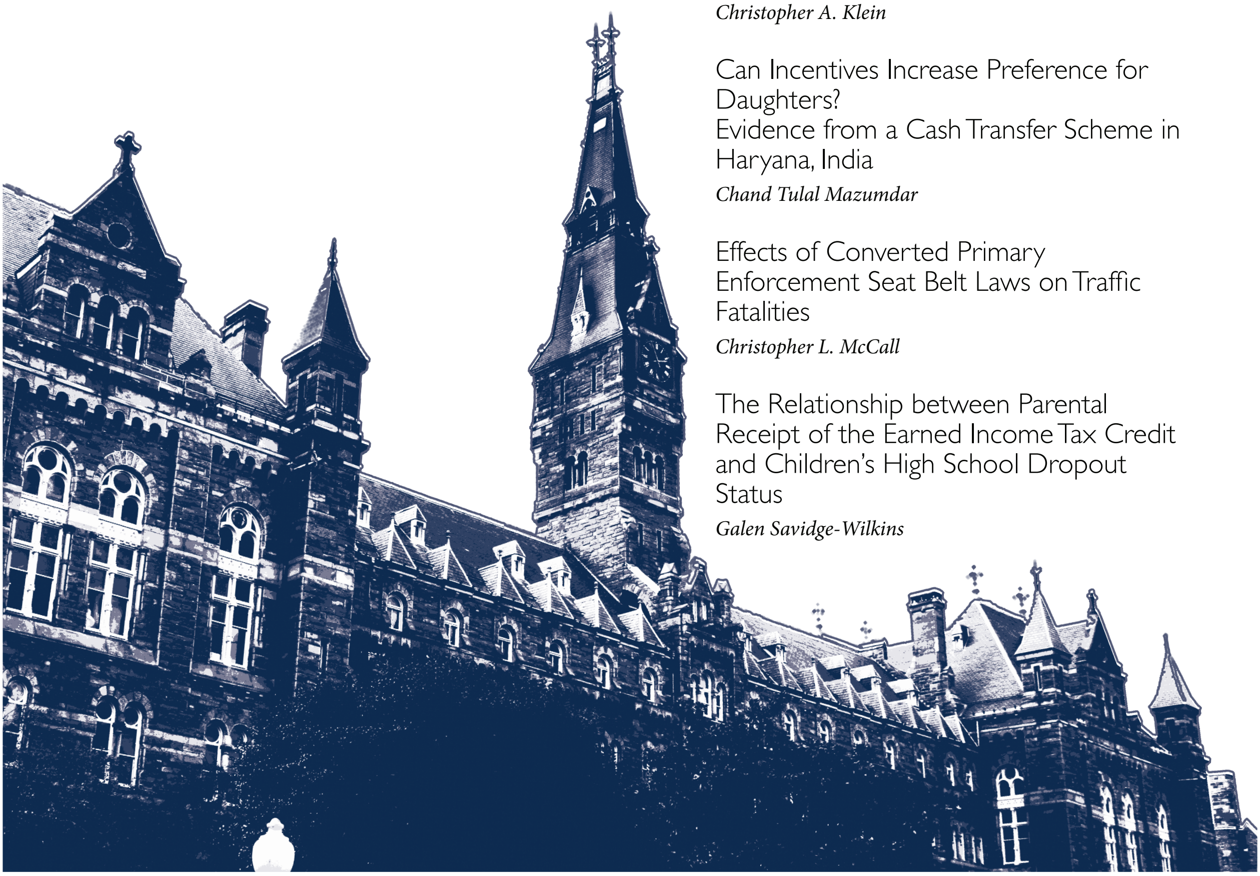
Volume 18 | Number 1 | Winter 2012–13

The Georgetown PUBLIC POLICY REVIEW

Graduate Thesis Edition



Graduate Thesis Edition

Editor's Remarks

Research

Special Thanks — We would like to express our infinite gratitude to David Boyer, Darlene Brown, Cristal Clark, Mary Denlinger, Carolyn Hill, Andreas Kern, David Konisky, Felecia Langford, Laura Masters, Lauren Mullins, Ed Montgomery, Alice Rivlin, Mark Rom, Barbara Schone, Dean Sirjue, Adam Thomas, Franck Wiebe, the GPPReview Alumni Board, the Graduate Student Organization, the Georgetown Public Policy Student Association, the Georgetown Public Policy Institute, and our faculty peer reviewers.

Interior Design & Layout — Nora Gregory and Danielle Parnass

Cover Design — www.acecreative.biz

Printer — Print 1 Printing & Copying

The Georgetown Public Policy Review publishes articles that contribute to the thoughtful discourse of public policy. The *Graduate Thesis Edition* is reserved for those thesis papers of Georgetown Public Policy Institute (GPPI) alumni that showcase superior policy analysis and particularly thoughtful writing. To be included in the *Graduate Thesis Edition*, each author must be nominated by a GPPI professor and take part in a comprehensive review process led by *Review* staff members and GPPI faculty. For more information about the *Graduate Thesis Edition*, visit www.gppreview.com.

INSTRUCTIONS FOR CONTRIBUTORS

The Georgetown Public Policy Review receives a large quantity of submissions for every issue. Articles submitted for the Spring Edition should be original and must not draw substantially from previously published works by the author. No simultaneous submissions please. Commentaries, opinion articles, and essays should be less than 3,000 words. Research articles should be between 3,000 and 5,000 words. All articles must utilize the accepted tools of policy analysis. We do consider longer pieces, but include only one such piece in a given issue. Adhere to the format in The Chicago Manual of Style.

Articles can be submitted by e-mail (printsubmissions@gppreview.com) or posted to our address. Submissions should include one cover document containing contact information (author's name, address, telephone, and e-mail address). Please indicate the primary contact if there are multiple authors. Do not show the author's name on any page of the article. Also include a one-paragraph abstract (approximately 150 words) and a brief biographical statement for each author (up to three lines). Posted copies should include one CD in Microsoft Word format with one printout. *The Review* editorial board reserves the right to edit or reject all submissions. It also fully reserves the right to publish all or part of an article from its print edition on its online edition.

Address all correspondence to:

The Georgetown Public Policy Review
Georgetown Public Policy Institute
37th and O Streets NW
100 Old North
Washington, DC 20057

E-mail: info@gppreview.com
Website: www.gppreview.com

SUBSCRIPTION INFORMATION

One issue of *The Review* costs $10.00 for an individual and $20.00 for an institution. For individuals and institutions outside the United States, rates are $17.00 and $27.00, respectively. This includes postage and packaging. To order or renew a subscription, please send a check payable to *The Georgetown Public Policy Review* to our mailing address above.

ADVERTISING

The Review accepts full- and half-page advertisements. E-mail our Executive Director of Finance and Operations at execfinancedirector@gppreview.com for details.

COPYRIGHT

THE GEORGETOWN PUBLIC POLICY REVIEW (ISSN 1083-7523) WINTER 2012-13 VOL. 18, NO. 1.

Editor's Remarks

This volume marks the second annual *Georgetown Public Policy Review Graduate Thesis Edition*, showcasing the sound research and policy analysis conducted by recent Georgetown Public Policy Institute (GPPI) graduates. As part of the GPPI capstone experience, students have the option of completing a quantitative thesis in partial fulfillment of the Master of Public Policy degree at Georgetown University. The *Graduate Thesis Edition* is reserved for those theses that showcase superior and relevant policy analysis and particularly thoughtful writing.

The *Graduate Thesis Edition* is peer reviewed by Georgetown faculty and publishes condensed versions of the authors' original theses. The result has been a powerful addition to *The Review*'s annual journals that draw on empirical analysis from the broader academic and policy communities.

In this issue, we begin with the work of Cynthia Brenner, who explores the role of gender on microfinance repayment in 110 countries. Using ordinary least squares and fixed-effects models, she finds strong evidence that, at a certain point, women pay back their loans at a greater rate than men do. This has major implications for microfinance institutions as they look to increase their financial sustainability.

On the domestic front, former *Review* editor Amanda Huffman explores the relationship between job mobility and wages. Her study suggests a previously unexplored link to tenure and experience, filling a gap in the job mobility literature. Utilizing a fixed-effects regression model, Huffman's analysis helps inform labor policies aimed at increasing wages of workers in the United States by determining when and how often workers should change jobs to maximize earnings potential.

Christopher Klein weighs in on the debate over feed-in tariff (FIT) policies in the European Union on whether encouraging renewable resource investment comes at the cost of higher electricity prices for consumers. His rigorous analysis employs a dynamic panel data model that captures both country- and year-specific effects for 20 European countries over time. Ultimately, he concludes that FITs have in fact had a relatively small effect on electricity prices, suggesting that these policies are both effective and economically viable.

We then turn to Chand Mazumdar's analysis of the gender imbalance in India and whether policies aimed at increasing household preferences for daughters have been effective. She uses a difference-in-differences model that compares two neighboring states to assess the impact of a cash transfer scheme on the likelihood of families having a second daughter. Her findings suggest that such policies may not yield their intended results very significantly.

Shifting back to the US, Christopher McCall delves into the issue of seat belt laws and whether states with tougher enforcement see fewer traffic fatalities. He expands upon research in the existing literature and, by using state-level fixed-effects models, finds a significant negative effect of upgrading to primary seat belt enforcement laws on car-related deaths. These results are even stronger in more populous states and in certain regions of the country, finding evidence to support more resources devoted to seat belt laws at the state level.

Our last study by Galen Savidge-Wilkins adds a unique analysis to the plethora of research on the Earned Income Tax Credit (EITC) in the US by looking at its specific effects on children's high school graduation rates. He employs a linear probability model to conclude that when parents receive the EITC at certain points in their children's lives, those children are more likely to graduate from high school. His findings suggest that policies aimed at family economic intervention may have positive implications for educational outcomes.

Each author carves out a unique analysis in his or her respective field of study, greatly contributing to academic and policy research. We cannot thank authors Brenner, Huffman, Klein, Mazumdar, McCall, and Savidge-Wilkins enough for working with us throughout the editorial process. We hope our readers similarly find this volume of the *Graduate Thesis Edition* rewarding and thought provoking, adding important nuances and layers to the dynamic policy conversation across the globe.

On behalf of *The Review*, I would also like to extend our gratitude to those members of the GPPI community who greatly enhanced our efforts this year, including Robert Bednarzik, our faculty advisor; Barbara Schone, MPP faculty director and thesis coordinator; and the nominating thesis advisors: Robert Bednarzik, Gillette Hall, Andreas Kern, J. Arnold Quinn, Omar Robles, Adam Thomas, Thomas Wei, and Andrew Wise.

I am most grateful to have worked with a remarkable group of peers to carry on *The Review*'s high-caliber contributions to public policy discourse. This publication is the result of the particular dedication of Lauren Shaw and each member of our exceptional print and copy editing teams.

Finally, I would like to extend special thanks to the Executive Team: Noora AlSindi, Lauren Shaw, Michelle Wein, Alex Engler, Josh Caplan, Kim Dancy, Tom Smith, and Sarah Orzell. Thank you for your unwavering support and inspiring leadership at every turn. I am looking forward to completing another outstanding year.

Danielle Parnass

Editor in Chief

Do Women Make the Difference?
The Effect of Gender on Microfinance Repayment Rates

By Cynthia M. Brenner

Cynthia Brenner is a graduate of the University of Notre Dame and the Georgetown Public Policy Institute. Gillette Hall, PhD served as her thesis advisor. Currently, Brenner is a Strategy and Operations Consultant with Deloitte Consulting's Federal Practice.

ABSTRACT

Since the financial crisis and the subsequent tightening of development assistance, foreign direct investment, and philanthropic donations, microfinance institutions (MFIs) have increasingly pushed financial independence as a means for ensuring their sustainability. Maximizing loan repayment rates is key to financial sustainability, reducing the cost of credit and dependence on subsidies. Many MFIs have adopted policies specifically targeting women in order to increase their impact on poverty reduction. The relationship between gender targeting and subsequent repayment rates has major implications for MFIs as they transition to financial independence. This study analyzes the effect of gender targeting on MFIs' financial sustainability by empirically examining the relationship between the proportion of female borrowers and repayment rates. Using the Microfinance Information Exchange's (MIX) global dataset covering 1,102 MFIs in 110 countries, the results indicate that female clients are associated with lower *portfolio-at-risks* and *write-off ratios* than their male counterparts. Furthermore, the results suggest there is a "tipping point" (30 percent female borrowers) above which women begin to repay at greater rates than men. Thus, this paper finds that the twin MFI goals of financial sustainability and targeting of loans to women are not contradictory; in fact, they are mutually reinforcing.

I. INTRODUCTION

As official development assistance[1] (ODA) continues to decrease, the sustainability of projects has become a key factor in all international development programs (OECD 2012). According to the International Finance Corporation (IFC 2011), sustainable development can be defined as "long-term business success contributing toward economic and social development, and to the overall stability of society." As donor budgets continue to tighten, achieving long-term program sustainability will require a shift from donor-funded subsidies to self-sufficient programs. Microfinance[2] has been touted throughout development circles as a powerful tool for navigating this transition. Meant to increase access to financial capital for the poor in the developing and developed world, microfinance refers to all financial products (loans, insurance, savings, and pensions) that are specifically geared toward poorer populations. Unlike many other antipoverty interventions, microfinance institutions (MFIs) can graduate from relying heavily on donor support to being completely financially independent. Yet, according to a recent United Nations (UN) study, only 10 percent of microlending organizations are self-sufficient as a majority of institutions still rely on outside support (Skoll Foundation 2012). To shift toward full sustainability, MFIs need to carefully determine factors that will increase their repayment rates and ensure their liquidity and sustainability.

Gender targeting has been presented as a parallel solution, supported by an extensive literature highlighting the role of women as drivers of economic development. Pitt and Khandker (1998) find that giving women access to financial capital has a greater impact on poverty reduction than male borrowing, thus increasing their development impact. The consequence of this gender targeting on repayment rates is a key consideration for MFIs seeking financial independence, but available evidence is mixed and far from rigorous. Some country studies find that having more women borrowers increases the repayment rate, while others find that a greater amount of male borrowers can strengthen the repayment rate. There is a need for further robust research on the default rate implications of MFIs' gender-targeting policies.

Identifying the causal mechanisms that create higher repayment rates allows for increased financial sustainability across MFIs and subsequently, the achievement of industry-wide growth and self-sufficiency. Unlike most of the previous research that is drawn from single-country studies, this study uses a large cross-country database from the Microfinance Information Exchange (MIX) covering 1,103 MFIs

[1] Official development assistance (ODA) refers to bilateral aid given from one country to another country, or from one country to a multilateral aid institution.

[2] This paper discusses microcredit, or microloans, but it will use microfinance institutions as the unit of analysis.

across 110 countries over 15 years to empirically examine the link between the percent of women borrowers and an institution's repayment rate. A more rigorous cross-country analysis of this relationship will contribute to the literature on gender and microfinance and shed light on the differences in the credit risk between men and women.

II. LITERATURE REVIEW

The bulk of the literature on microfinance examines the economic impact of increased access to financial services.[3] Less research is focused on how to increase the financial sustainability of the MFI and yields less rigorous evidence. Every MFI tries to maximize its repayment performance, regardless of whether it is a for-profit institution or more focused on economic development. The benefits of high repayment rates are considerable as they affect the borrower, the institution, and the investor. Marie Godquin (2004) discusses the importance of focusing on repayment rates in MFI operations as it reduces the financial cost of credit, extending access to more borrowers. Lower default rates reduce dependence on subsidies, lower the borrower's interest rates, and improve sustainability. Additionally, Godquin argues that higher repayment rates reflect the efficiency and effectiveness of the MFI.

[3] These studies use GDP, HDI, and level of poverty as dependent variables, thus analyzing the growth of these programs against these measures of economic and social development.

"... giving women access to financial capital has a greater impact on poverty reduction than male borrowing ..."

Past research confirms that investing in women can increase their community and family status, which is beneficial for the community at large. Mayoux (2003) examines the link between access to finance and development through women's empowerment. She identifies two underlying assumptions of investing in women. First, evidence suggests that microfinance will automatically lead to women's empowerment and therefore faster macroeconomic growth. The second assumption is that women's empowerment, household level poverty alleviation, and community development are inherently connected. Increased well-being and group formation will automatically enable women to empower themselves.
The Kiva Foundation concludes that financial services have improved the status of women within the family and the community: "Women have become more assertive and confident. In regions where women's mobility is strictly regulated, women have become more visible, are better able to negotiate in the public sphere… and play a stronger role in decision making."

While the strategy of targeting women in order to increase the development impact of microfinancing has been rigorously examined, studies on this policy's effect on the financial sustainability of the MFI are limited and have returned mixed results. Kappel, Krauss, and Lontzek (2010) use the repayment rate of an institution to measure financial sustainability for microfinance institutions, as this indicator typically plummets prior to the crises at an MFI. Because microcredit lending procedures do not include a physical capital requirement, MFIs cannot continue their lending operations if repayment rates are low.

In an effort to solve this problem, several country studies look at the effects of gender targeting on repayment rate. Roslan and Karim (2009) analyze the Malaysian microfinance market using a logit-probability approach and find that the proportion of female borrowers exhibits a statistically significant negative effect on the repayment rate. These findings suggest the policy of targeting women detracts from the financial sustainability of an MFI.

In a similar study from Ghana, Richman and Fred (2004) find that increasing the share of male borrowers increases the repayment rate. Unlike Roslan and Karim, Richman and Fred use a fixed-effect approach to account for the differences across countries and time. Both studies use portfolio-at-risk 30 and 90 days to measure repayment rate and find that the share of female borrowers has a negative impact on the repayment rate.

These country-specific case studies contrast with a 2009 cross-country analysis by Merslund, D'Espallier, and Guerin, who find a positive effect of percent of women borrowers on bank liquidity through the repayment rate. Using 350 MFIs from 70 countries, the team uses fixed-effects and random-effects models to examine the policy of targeting women, using a portfolio-at-risk measure to operationalize repayment rate. The authors conclude that women are better credit risks, especially for non-profits and NGOs that run programs in conjunction with lending procedures. As the first cross-country analysis, Merslund, D'Espallier, and Guerin (2009) provide a framework for which to study these effects but acknowledge the need for further research due to their incomplete data.

Unexplored by Merslund, D'Espallier, and Guerin (2009) is whether the magnitude of the gender effect varies with the proportion of women borrowing from an institution. Godquin (2004) argues that the composition of group dynamics and social support systems directly affect the repayment rate. Social ties and group homogeneity can indirectly heighten repayment performance by facilitating peer monitoring and peer pressure (Besley and Coates 1995). Thus, theory suggests that the effect of women borrowers would vary depending on the number of women participating in microlending. As more women begin to take out loans, the

"These results will contribute to a greater understanding of the relationship between gender targeting and repayment rates, testing the belief that women honor microcredit contracts better than men."

group dynamic would encourage better on-time repayment.

This paper aims to provide a more comprehensive study, including a sensitivity analysis of the relationship between gender and MFI repayment rates. This is important because of the unanalyzed belief that women honor their microcredit contracts more than men. Building on Merslund, D'Espallier, and Guerin (2009), this study uses a much larger dataset covering more institutions and more countries, accounting for regional variability in political economy and share of female borrowers in the hope of providing a more inclusive analysis. These results will contribute to a greater understanding of the relationship between gender targeting and repayment rates, testing the belief that women honor microcredit contracts better than men. A better understanding of these issues can help microfinance institutions strengthen repayment rates, creating positive benefits through lower interest rates and greater sustainability.

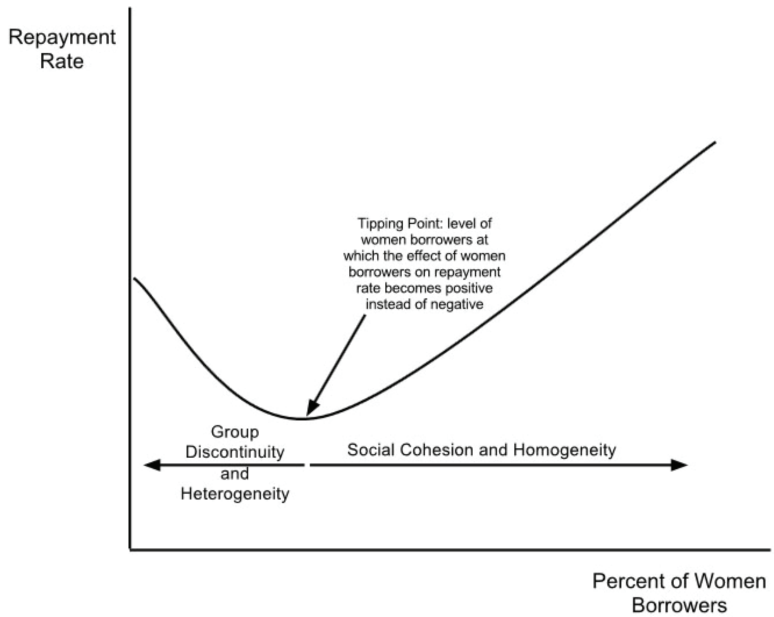
III. CONCEPTUAL FRAMEWORK & HYPOTHESIS

This paper asks what effect the number of women borrowers has on a bank's overall repayment rate, and whether the direction and magnitude of this effect varies depending on region and share of women borrowers. Microfinance advocacy networks and sponsors regularly assert that women are safer credit risks. The World Bank (2007) states that lending to women leads to higher repayment rates due to the conservative nature of women's investments and the lower risk of moral hazard. In their assessment of the different techniques in reducing repayment defaults, Armendariz and Murdoch (2005) consider targeting women to be a technique in its own right alongside group lending or dynamic incentives. Competing results from individual country case studies suggest that this effect varies depending on region.

Providing credit to women has been highly popularized by the microfinance industry. As women make up more than 70 percent of the world's poor, the demand for loans is higher for women than for men. Merslund, D'Espallier, and Guerin (2009) argue that women invest their income in order to nurture the well being of their families; therefore, one dollar loaned to a woman has greater development impact and multiplier effect than one dollar loaned to a man. They also argue that this results in greater ease of repayment for women. Karim (2011) finds evidence that this effect varies by region because in areas where the incentive to repay is greater a "political economy of shame" can develop in joint-liability and self-help groups, thus increasing a woman's repayment in order to avoid being disgraced. This would indicate that a woman's repayment rate might vary by region due to cultural norms.

Group dynamics and group homogeneity have been linked to higher rates of repayment, as joint liability strengthens peer monitoring

Figure 1. Conceptual Framework for Possible "Tipping Point"



and peer pressure. This suggests a non-linear relationship between loan repayment rates and the percent of women borrowers. Initially, the ease of repayment will be weaker with few women participating in microlending, as there is a lack of social ties and peer monitoring. As the number of women borrowers grows, social cohesion and homogeneity should increase the repayment performance of participants (Zeller 1998). Specifically, this paper theorizes that at low levels of women borrowing, female repayment performance will be worse than male, but as the group dynamics grow more influential and more women become borrowers, there is a "tipping point" at which the effect of female borrowing becomes increasingly positive. The theory is illustrated in Figure 1.

This paper hypothesizes that the ratio of female borrowers should ultimately have a negative effect on the portfolio-at-risk and write-off ratios in microfinance institutions. It further suggests that the direction and magnitude of this effect varies by region; in countries where the impetus and the cultural attitudes toward non-repayment are stronger, the effect of female borrowers on the repayment rate will be stronger. This effect is also allowed to vary according to the percentage of women borrowers in order to test the theory that as more women borrow, group dynamics contribute to increased repayment rates. Thus, this paper tests the following hypotheses through statistical analysis:

> H_1: An increase in the percent of women borrowers improves the repayment rate of a microfinance lending institution (MFI).

H_2: The effect of the percent of women borrowers on the MFI's repayment rate varies by region of the world.

H_3: The effect of women borrowers on an MFI's repayment rate varies over the range of the percentage of women borrowers in an MFI's portfolio.

IV. ECONOMETRIC MODEL

Repayment rate is operationalized through two different measures: portfolio-at-risk and write-offs. Through multiple estimation methods, this paper accounts for the methodological issues related to this type of assessment such as isolating the gender effect from other MFI or institutional effects that influence repayment, and accounting for the time-invariant nature of many covariates. Two different models are run in order to account for these methodological issues.

An ordinary least squares (OLS) model is run in order to assess the effect of female borrowers on the MFI's repayment rate while accounting for institutional and macroeconomic conditions as controls and contributors to the repayment rate. Both year and country controls are included in order to remove the bias on the coefficients and control for outside effects that may affect repayment rates. Developed by Plümper and Troeger (2007), a fixed-effects (FEVD) model is also run in order to account for time-invariant covariates and institutional fixed effects in the context of panel data. The OLS model is needed in order to let the region effects vary as these would drop out of a purely fixed model. Additionally, the OLS estimation allows

OLS Models

For Hypothesis 1:

$$PAR30 = \beta_0 + \beta_1 PercWomBor_i + \beta_2 MFIControls_i + \beta_3 CountryControls_i + \beta_4 YearControls_i + \mu_i$$

$$WriteOffRatio = \beta_0 + \beta_1 PercWomBor_i + \beta_2 MFIControls_i + \beta_3 CountryControls_i + \beta_4 YearControls_i + \mu_i$$

For Hypothesis 2:

$$PAR30 = \beta_0 + \beta_1 PercWomBor_i + \beta_2 PercWomenBor * Region_i + \beta_3 MFIControls_i + \beta_4 CountryControls + \beta_5 YearControls + \mu_i$$

$$WriteOffRatio = \beta_0 + \beta_1 PercWomBor_i + \beta_2 PercWomenBor * Region_i + \beta_3 MFIControls_i + \beta_4 CountryControls + \beta_5 YearControls + \mu_i$$

For Hypothesis 3:

$$PAR30 = \beta_0 + \beta_1 PercWomBor_i + \beta_2 PercWomenBor^2_i + \beta_3 MFIControls_i + \beta_4 CountryControls + \beta_5 YearControls + \mu_i$$

$$WriteOffRatio = \beta_0 + \beta_1 PercWomBor_i + \beta_2 PercWomenBor^2_i + \beta_3 MFIControls_i + \beta_4 CountryControls + \beta_5 YearControls + \mu_i$$

Fixed-Effects Models

For Hypothesis 1*:*

$$PAR30 = \beta_0 + \beta_1 PercWomBor_{it} + \beta_2 Z_{it} + \alpha_{it} + \mu_{it}$$

$$WriteOffRatio = \beta_0 + \beta_1 PercWomBor_{it} + \beta_2 Z_{it} + \alpha_{it} + \mu_{it}$$

For Hypothesis 3*:*

$$PAR30 = \beta_0 + \beta_1 PercWomBor_{it} + \beta_2 PercWomBor^2_{it} + \beta_3 Z_{it} + \alpha_{it} + \mu_{it}$$

$$WriteOffRatio = \beta_0 + \beta_1 PercWomBor_{it} + \beta_2 PercWomBor^2_{it} + \beta_3 Z_{it} + \alpha_{it} + \mu_{it}$$

Where Z is a vector of MFI time-variant controls.
Fixed effects (FEVD): Assume α_i is not independent of X_{it}, Z_i.

for an unbalanced panel, which boosts the sample size and strengthens the analysis.

The unit of analysis for all three estimation methods is the individual microfinance institution. In the OLS model, the analytical sample is limited to 2005 to 2009 due to missing data in earlier and more recent years.

V. DATA DESCRIPTION

This paper uses the Microfinance Information Exchange (MIX) Market data. It has been collected from more than 2,000 MFIs in 110 developing countries over 16 years and classified according to International Financial Reporting Standards (IFRS).[4] The financial, social, and operational information featured in MIX Market data is directly self-reported by individual institutions or affiliated network and/or gathered from the institutions' publications (i.e., annual report).

[4] This sample is limited for the analytical sample. Discussion of the creation of the analytical sample is later in the paper.

MIX analysts validate all data received and after doing a thorough accuracy check, MIX updates data and makes it publicly available. Data has been collected from 1995 to 2011, and there are 10,223 observations and 77 variables. This data is collected at the bank level, which creates an aggregation bias as it would be preferential to have individual data. Aggregation bias occurs if there is a loss of detail when the unit of analysis is aggregated around the institution instead of the individual. However, as this is a cross-country analysis, this data is a strong representation of what is occurring in the worldwide microcredit industry.

Portfolio-at-risk 30 and the *write-off ratio* are used to operationalize the repayment rate of the bank. *Portfolio-at-risk 30* is a variable that indicates the percent of loans that are overdue at least 30 days and takes a value ranging from 0 to 1. It is measured as portfolio-at-risk that is greater than 30 days late divided by the gross loan portfolio. The *write-off ratio* describes the percent of loans written off during the period, also measured as write-offs divided by the gross loan portfolio. A write-off is an accounting procedure that removes the outstanding balance of the loan from the Loan Portfolio and from the Impairment Loss Allowance when the loan is recognized as uncollectible.

These variables are measured against the primary independent variable and control variables. The primary independent variable is the percent of women borrowers measured as the number of active women borrowers divided by the number of active borrowers. The average share of women borrowers is 64 percent but this varies by type of institution. Table 1 shows a breakdown of the percent of women borrowers by type of institution. As can be seen in Table 1, every type of institution has more than 50 percent women borrowers but Non-Banking Financial Institutions and Non-Governmental Institutions have much higher rates. These institutions also comprise the majority of the sample, which is similar to the composition of the microfinance industry as a whole (Rhyne 2010).

In addition, the analysis includes an interaction between the percent of women borrowers and region of the world to determine whether this effect varies by country or region of the world.

As indicated by Merslund, D'Espallier, and Guerin (2009), numerous MFI control variables must be included in the analysis to properly isolate the effect of the primary variable of interest (percent female borrowers). Using past research, a number of key firm-level variables have been identified. These variables appear in the MIX data set as: *Target Market, Type of Institution, Regulated, Growth Rate of the Total Loan Portfolios, Staff Efficiency, Operational Self-Sufficiency, Cost per Borrower, Age of Institution, Average Loan Size, and Gross Loan Portfolio*. Additionally, the OLS models include a control for the country and year in which the MFI operates in order to account for varying macro-effects.

VI. RESULTS

Table 2 depicts the impact of gender on firms' *portfolio-at-risk* and the *write-offs* through an ordinary least squares (OLS) estimation method with robust standard errors to correct

Table 1. Breakdown of the Percent of Women Borrowers by Type of Institution

Descriptive Statistics of Type of Institution and Percent of Women Borrowers

Type of Institution	Bank	Credit-Union	Non-Banking Financial Institution	Non-Governmental Institution	Rural Bank	TOTAL
Percent of Women Borrowers	55.6%	52.3%	60.7%	75.9%	55.9%	64.3%
Number of Observations	312	679	1395	1578	286	4,260
Percent of Sample	7.3%	15.9%	32.7%	37.0%	6.7%	100.0%

Table 2. OLS Regression of Percent of Women Borrowers on Portfolio-at-Risk over 30 days and the Write-Off Ratio

Variables	(1) Portfolio-at-Risk	(2) Write-Off Ratio
Percent Women Borrowers	-0.0357***	-0.00568**
	(0.0074)	(0.00306)
MFI Age		
- mature age	0.0126***	0.0023*
	(.0037)	(0.0015)
- new age	-0.0228***	-0.0069***
	(0.0049)	(0.002)
Target Market		
- broad market	0.00567	-0.000505
	(0.0079)	(0.00317)
- highendmarket	-0.00673	-0.0044
	(0.00862)	(0.00347)
- lowendmarket	-0.00140	0.00238
	(0.00927)	(0.00372)
Type of Institution		
- Bank	0.0145**	0.00813**
	(0.00769)	(0.00314)
- NGO	0.00217	0.00299
	(0.00574)	(0.00238)
- NBFI	0.0015	0.00403
	(0.00553)	(0.00227)
RuralBank	0.039***	-0.00526
	(0.00884)	(0.00357)
Regulated	0.0115***	-0.00253
	(0.00419)	(0.00170)
Gross Loan Portfolio	-0.00457	-0.00210*
	(0.00286)	(0.00116)
Cost Per Loan	-2.83E-06	-1.78e-06
	(3.67e-06)	(1.49e-06)

Table 2 Continued

Variables	(1) Portfolio-at-Risk	(2) Write-Off Ratio
Operational Self-Sufficiency	-0.00569**	-0.00089
	(0.00252)	(0.00102)
log(Portfolio)	-0.0083***	0.000213
	(0.00168)	(0.000688)
log(Staff Efficiency)	-0.0102***	-0.00404***
	(0.00259)	(0.00106)
Provision for Impairment Loss	1.043***	0.492***
	(0.0437)	(0.0188)
Depositors per Staff Member	3.92e-05***	-1.25e-06
	(8.22e-06)	(3.42e-06)
Sustainability	-0.0233***	-0.00858***
	(0.0038)	(0.00154)
MFI Scale		
- large scale	-0.0019	-0.00457*
	(0.0067)	(0.0028)
- medium scale	-0.0058	-0.0019*
	(0.0044)	(0.0017)
Constant	0.246***	0.0681***
	-0.0199	-0.00817
Observations	4,326	4,260
R-squared	0.3211	0.329
Additional Controls		
Percent Women Borrowers missing values	Yes	Yes
Cost Per Loan missing values	Yes	Yes
Country Effects	Yes	Yes
Year Effects (2005-2009)	Yes	Yes

Standard errors in parentheses.

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

for potential heteroskedasticity. For gender, the proxy of *percent of women borrowers* is used to operationalize the influence of gender on the repayment rate. The different columns in Table 2 correspond to the different dependent variables. In both columns 1 and 2, the proportion of female borrowers is negatively related to the portfolio-at-risk and the write-off ratio. These estimated effects are highly statistically significant, though their effect is small. An increase of one percentage point of female borrowers is associated with a 0.036 percentage point decrease in the portfolio-at-risk, holding all else constant.

Looking at the other controls in Model 1, all of the significant coefficients take the predicted signs. *Rural bank, regulation, log(portfolio), staff efficiency, provision for loan loss impairment, depositors per staff member,* and *sustainability* are statistically significant. In particular, a lower portfolio-at-risk is associated with larger MFIs and larger portfolio growth rates, more highly regulated and sustainable MFIs, and MFIs with higher staff efficiency. Higher portfolio-at-risks are associated with MFIs that operate in rural areas, MFIs with a high depositor-to-staff ratio, and MFIs with a higher impairment loss-to-assets ratio.

For Model 2, the results are similar when the write-off ratio is the dependent variable. Higher write-off

Table 3. Regression on Portfolio-at-Risk 30 days and Write-Off Ratio from 2005-2009

Variables	(1) Portfolio-at-Risk	(2) Portfolio-at-Risk	(3) Write-Off Ratio	(4) Write-Off Ratio
Percent Women Borrowers	-0.0363***	-0.0399***	-0.00568*	-0.00829**
	(0.00745)	(0.00848)	(0.00306)	(0.00410)
Cost Per Loan	-3.37E-06	-2.28E-06	-1.81E-06	-2.87E-06*
	(3.67E-06)	(3.36E-06)	(1.49E-06)	(1.62E-06)
Percent Women Borrowers Missing Variable Indicator	0.0220***		0.000218	
	(0.00511)		(0.00208)	
Cost Per Loan Missing Variable Indicator	-0.00571		-0.00127	
	(0.00381)		(0.00156)	
Constant	0.227***	0.214***	0.0577***	0.0815***
	(0.0215)	(0.0239)	(0.00875)	(0.0116)
Observations	4,330	2,771	4,264	2,751
R-squared	0.31	0.376	0.326	0.313

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.10

ratios are associated with larger banks and higher impairment loss-to-assets ratio. Lower write-off ratios are associated with higher staff efficiency, greater sustainability, and a larger gross loan portfolio.

The OLS model's missing variable indicators have been used for the primary independent variable of interest—*percent of women borrowers*—and an important control variable—*cost per loan.* As can be seen in Table 3, Models 1 and 3 correspond with Table 2's Models 1 and 2. Models 2 and 4 from Table 3 do not control for missing variables. The table shows that the missing variable indicators do not dramatically change the statistical significance of the coefficients or their magnitude. They do, however, significantly boost the number of observations while controlling for potential bias created by the missing values.

As argued in the literature review section, the effect of women borrowers should be more prevalent under certain conditions, such as the theory put forward by Karim (2011) that a region's culture creates a greater stigma on non-repayment by women. Based on the interaction effects observed between the percent of women borrowers and six regions of the world, this analysis does not support Karim's theory. For both dependent variables of interest, the interaction effects are neither individually nor jointly statistically significant, as can be seen in Table 4. This paper does not provide evidence that the effect of women borrowers on the repayment rate varies by region, and therefore rejects Hypothesis 2.

The next models are analyzed in Table 5 through the fixed-effects estimation method. As can be seen in Table 5, most of the variables are statistically significant at conventional levels. This output not only tells us that the percent of women borrowers is associated with lower portfolios-at-risk and lower write-off ratios, but also suggests that more sustainable MFIs with larger portfolios and higher staff efficiency result in better repayment rates. On the other hand, higher costs per loan, higher impairment loss-to-assets ratio, and higher ratios of depositors-to-staff member ratios lower a bank's repayment rate.

In order to let the effect of women borrowers vary over the range of the percent of women borrowers, a quadratic term is added to the main regression in Table 6.[5] As can be seen, this functional form is significant at conventional levels, thus suggesting that the effect of women borrowers does vary over the range of percent of women borrowers in an MFI's portfolio. Specifically, when there is a smaller share of women borrowers compared to men (less than one-third female), the women borrowers increase the bank's *portfolio-at-risk* and *write-off-ratio*. At this level (33 percent of women borrowers for the write-off ratio and 35 percent of women borrowers for the portfolio at

[5] Table 6 shows only the main effect and quadratic term as added into the main regression found in Table 2. No other results differ significantly.

Table 4. Interaction Effects between the Percent of Women Borrowers and Region Indicator Variables

Variables	(1) Portfolio-at-Risk	(2) Portfolio-at-Risk	(3) Write-Off Ratio	(4) Write-Off Ratio
Percent Women Borrowers	-0.0357***	-0.118	-0.00568**	-0.00100
	(0.0074)	(0.194)	(0.00306)	(0.0787)
MENA*percent women borrowers		0.121		-0.00663
		(0.195)		(0.0791)
Europe*percent women borrowers		0.0971		0.00650
		(0.195)		(0.0792)
SouthAmerica*percent women borrowers		-0.0474*		-3.34e-05
		(0.0271)		(0.0110)
LAC*percent women borrowers		0.0718		0.00361
		(0.195)		(0.0791)
Asia*percent women borrowers		0.103		-0.00175
		(0.195)		(0.0789)
SubSahAfrica*percent women borrowers		0.0474		-0.0233
		(0.195)		(0.0789)
Constant	0.227***	0.201***	0.0577***	0.0585***
	(0.0215)	(0.0236)	(0.00875)	(0.00962)
Observations	4,326	4,326	4,260	4,260
R-squared	0.3211	0.314	0.329	0.328
Additional Controls				
Percent Women Borrowers missing values	Yes	Yes	Yes	Yes
Cost Per Loan missing values	Yes	Yes	Yes	Yes
Country Effects	Yes	Yes	Yes	Yes
Year Effects (2005-2009)	Yes	Yes	Yes	Yes

Standard errors in parentheses.

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

Table 5. Fixed-Effects Regressions on Portfolio-at-Risk and Write-Off Ratio

Variables	(1) Portfolio-at-Risk	(2) Write-Off Ratio
Percent Women Borrowers	-0.0346**	-0.0181**
	(0.0139)	(0.00861)
MFI Age		
- mature age	0.0044	0.0021
	(0.0052)	(0.0032)
- new age	-0.0056	-0.0113**
	(0.0073)	(0.0045)
Sustainability	-0.0135***	-0.00984***
	(0.00448)	(0.00277)
Cost Per Loan	2.97E-06**	1.65e-06*
	(1.46e-06)	(8.84e-07)
Gross Loan Portfolio	-5.81E-04**	-2.79E-04*
	(2.70e-05)	(1.63e-05)
Log(Staff Efficiency)	-0.0158***	-0.0325***
	-0.00478	-0.00304
Provision for Loan Loss Impairment	0.915***	0.213***
	(0.0447)	(0.0313)
Depositors per Staff Member	6.38E-04***	2.08E-04***
	(1.01e-05)	(6.24e-06)
Constant	0.158***	0.172***
	(0.0303)	(0.0194)
Observations	3,219	3,218
Number of MFI Names	1,102	1,103
R-squared	0.2323	0.113
Year Fixed Effects	Yes	Yes
MFI Fixed Effects	Yes	Yes

Standard errors in parentheses.

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

Table 6. Regression of Percent of Women Borrowers Quadratic on the Repayment Rate				
Variables	(1) Portfolio-at-Risk	(2) Portfolio-at-Risk	(3) Write-Off Ratio	(4) Write-Off Ratio
Percent Women Borrowers	-0.0357***	0.0552**	-0.00568**	.0085*
	(0.0074)	(0.0281)	(0.00306)	(0.00115)
Percent Women Borrowers Squared		-0.0781***		-0.0122**
		(0.0231)		(0.00095)
Constant	0.227***	0.226***	0.0577***	0.0651***
	(0.0215)	(0.021)	(0.00875)	(0.0085)
Observations	4,326	4,326	4,260	4,260
R-squared	0.3211	0.323	0.329	0.329
Additional Controls	Yes	Yes	Yes	Yes
Percent Women Borrowers missing values	Yes	Yes	Yes	Yes
Cost Per Loanmissing values	Yes	Yes	Yes	Yes
Country Effects	Yes	Yes	Yes	Yes
Year Effects (2005-2009)	Yes	Yes	Yes	Yes

Standard errors in parentheses.

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

risk), there is a "tipping point" where the effect of more women borrowers increases or strengthens the banks overall repayment rate. This effect is consistent with the theory stated prior that with higher proportions of women borrowers, groups and support systems strengthen the ability of women to repay the loan as compared to men.

The results from Table 2 and Table 5 indicate a negative association between the number of female clients and repayment, confirming Hypothesis 1 that the proportion of female clients reduces the MFI's default rate. Hypothesis 3 is also confirmed by the results noted in Table 6, suggesting that the gender effect on repayment performance varies by the percent of women borrowers at an individual MFI. These effects hold for multiple measures of repayment (*par30* and *write-offs*) and for several estimation methods (OLS and FEVD).

VII. GENDER & LOAN SIZE

One can conclude from this study that women are repaying their microfinance loans at a greater rate than men are. Yet the reason behind these lower portfolios-at-risk and write-off ratios has not been determined. Examining the descriptive statistics of average loan sizes and the different quartiles of the share of women borrowers suggests that one reason for the higher repayment is the smaller loans being given at institutions that target women.

Table 7. Banks with Smaller Share of Women Borrowers Have Larger Average Loan Size Descriptive Statistics of Average Loan Size and Percent of Women Borrowers					
Quartile of Percent of Women Borrowers	First Quartile: 0% to 43%	Second Quartile: 43% to 65%	Third Quartile: 65% to 91%	Fourth Quartile: 91% to 100%	Total
Average Loan Size	$2,289.79	$1,810.88	$784.80	$194.37	$1,352.33
Number of MFIs	276	275	276	275	1,102
Percent of Sample	25%	25%	25%	25%	100%

Is higher repayment being driven by women, or are women choosing smaller loans and therefore greater ease of repayment? If women frequent institutions with smaller loans as a means to ensure on-time repayment or because of MFI preference, the endogeneity[6] would cause a bias in the coefficient on percent of women borrowers in the previous models. In order to correct for the endogeneity that may exist, an instrumental variable regression is used to parse out the gender effect from loan size on the repayment rate (Table 8). A variable was created that captures the deviation from the average loan size in order to measure the difference between what women are receiving as a loan and what size loan is being given to the average client.

[6] Endogenous variables are determined within the model, as compared to an exogenous variable determined outside the model. With an endogenous variable, causality would run both from the percent of women borrowers to the repayment rate and from the repayment rate to the percent of women borrowers.

Estimating through two stage least square[7] (2sls), a variable is an effective instrument if it passes both the relevance restriction and the exogenous restriction.[8] According to Stock and Watson (2011), an instrument is relevant when the variation in the instrument is related to variation in the independent variable of interest. This analysis finds both variables relevant based on their respective F-statistics in Table 8.[9]

[7] Stock and Watson (2011) describe this estimation method: "As the name suggests, the two stage least squares estimator is calculated in two stages. The first stage decomposes X into two components: a problematic component that may be correlated with the regression error and another problem-free component that is uncorrelated with the error. The second stage uses the problem-free component to estimate the coefficient."

[8] Computing the t-stat in the first stage can test the relevance condition. The validity/exogeneity condition, however, cannot be tested, because the condition involved the unobservable residual μ. Therefore this condition has to be taken on faith.

[9] This can be tested in the first stage of the process by looking at the F-stat for the overall regression being greater than 10. As seen in Table 8, the F-stat for the first stage is 312.85 for the *portfolio-at-risk* dependent variable and 296.03 for the *write-off ratio*.

Table 8. Regression Using Deviation from Mean Loan Size as an Instrumental Variable							
	(1)		(2)		(3)		(4)
Instrument	First Stage: PAR30	**Instrumented**	Second Stage: PAR30	**Instrument**	First Stage: WriteOff Ratio	**Instrumented**	Second Stage: WriteOff Ratio
Deviation from Mean Loan Size	-0.0000159*** (8.96E-07)			**Deviation from Mean Loan Size**	-0.0000156*** (9.08E-07)		
Constant	0.6328***	**Constant**	0.0352***	**Constant**	0.6347***	**Constant**	0.00166***
	(0.0030)		(0.0662)		(0.00321)		(0.0237)
		Percent Women Borrowers	0.04855 (0.1026)			**Percent Women Borrowers**	0.0254 (0.0368)
Observations	5,225		5,225		4,691		4,691
F-Statistic	312.85		.		296.03		.
R-squared	0.0563				0.0594		.

Standard errors in parentheses.

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

The instrument satisfies the exogenous condition if it captures movements in the independent variable of interest that are also exogenous while only affecting the dependent variable through the variable being instrumented. This condition cannot be empirically tested with only one instrument and one variable. Deviation from the average loan size could be argued as being exogenous in that higher- or lower-than-average loan amounts are driven partly by the women's preferences. This is exemplified by the greater number of women borrowing from institutions with lower average loan sizes or by women choosing to not take on as much debt as men. This variable is also a function of the borrower's demand for MFIs with certain average loan sizes and the supply of MFIs in an area. As this demand and supply cannot be easily measured, it biases the *percent of women borrowers* variable in the OLS and FEVD regressions. Furthermore, deviation from the average loan size does not directly affect the dependent variables representing repayment, as both dependent variables are functions of the total loan portfolio size, and thus the average loan of the bank is taken into account into their calculation. Therefore, the deviation variable captures the difference between the average loan sizes of the MFI compared to the worldwide average loan size. This should not directly affect an institution's repayment rate except through the gender of the borrower that chooses to frequent an MFI with either a lower or higher average loan size.

Table 8 depicts whether it is this choice, measured by the deviation from the mean loan size, which is driving the higher repayment rates. If the unit of analysis were the individual borrower or individual loan, this would not be an appropriate instrument as it would be biased by past repayment history and individual characteristics.

The results of the second stage, shown in columns 2 and 4, demonstrate that the percentage of women borrowers is no longer a significant predictor of repayment rates. This suggests that the positive effect of women borrowers found earlier can be partly explained by the loan size chosen by women compared to men. Thus, women are repaying at higher rates, but their ability to do so is driven partly by their preference of loan size or banking institution. This is a first step in the process of determining the characteristics of women that cause higher on-time repayment rates than men.

VIII. POLICY IMPLICATIONS & CONCLUSION

This paper uses MIX's global dataset covering 1,102 MFIs in 110 countries to test whether there is a gender effect on microfinance repayment. The findings indicate that MFIs with higher proportions of female borrowers have a lower *portfolio-at-risk* and *write-off ratio.* These results provide compelling, rigorous evidence that focusing on female clients enhances microfinance repayment rates and that women are generally a better credit risk.

By increasing repayment rates, an MFI not only benefits from lower risk and greater revenue, it also becomes a more sustainable institution and less dependent on donor contributions.

"These results provide compelling, rigorous evidence that focusing on female clients enhances microfinance repayment rates and that women are generally a better credit risk."

Borrowers also benefit from higher MFI repayment rates as interest rates decrease with lower default rates and investors will receive higher returns that make the investment more appealing. This is important because repayment has become increasingly important to the sustainability of microfinance as a development tool. Repayment is studied through two different measures: portfolio-at-risk and the write-off ratio, and gender is studied through the proportion of female clients.

Breaking down the results further, the findings indicate that at much lower levels of women borrowers, women are unable to pay their loans at higher rates than men. There is a "tipping point" at 33 percent and 35 percent where both the portfolio-at-risk and write-off ratio begin to benefit from women borrowers. The finding is supported by research from Zeller (1998) and Godquin (2004), who suggest that group dynamics and social homogeneity directly impact the repayment rate through peer monitoring, incentives, and peer pressure. These findings suggest that MFIs should move toward preferential treatment of women borrowers, thereby creating the group dynamic that benefits both borrowers and institutions.

"... MFIs' twin goals of financial sustainability and gender targeting are not contradictory; in fact, they are mutually reinforcing."

The results do not find evidence supporting the theory put forth by Karim (2011) that cultural norms inflict greater shame on women, thus increasing their repayment rates. The interaction terms reveal that there is no statistically significant variation of women's repayment by region. By extension, the findings do not support Karim's thesis that cultures with a higher impetus of shame will provide greater embarrassment for non-repayment and dishonor women who fall into debt.

In exploring the driving force behind women's higher repayment, this paper examines gender disparities in loan sizes. The analysis finds the difference in repayment rate between men and women to be negligible when controlling for women's preference for smaller loans or different types of institutions. Since women are given smaller loans on average, they are more likely to be able to repay them at on-time rates. Due to data limitations, this paper could not determine the reason for the smaller loans or whether the group dynamic would be able to compel greater repayment when women do take large loans. Further research is needed in this area to understand the effect of gender on the repayment rate at higher loan levels. In particular, studies should focus on high-value loans to determine whether women have lower default rates.

This paper shows that the difference in the loan size demanded by women compared to the worldwide average loan size could be one explanation for higher rates of on-time repayment. Potentially women are not receiving opportunities for larger loan sums. Or are banks that target women setting more lenient and favorable repayment schedules? Further research is needed on the driving force in order to understand the factors leading to higher repayment rates.

Overall, this paper finds compelling cross-country evidence of what policy makers and practitioners have long argued: women are better at repaying microfinance loans than men are. It is interesting to observe that despite a lower objective credit-worthiness due to greater lack of assets, women prove to be better borrowers and better credit risks. Therefore, MFIs' twin goals of financial sustainability and gender targeting are not contradictory; in fact, they are mutually reinforcing.

IX. REFERENCES

Godquin, Marie. 2004. Microfinance repayment performance in Bangladesh: How to improve the allocation of loans by MFIs. W*orld Development* 32: 1909-1926.

International Finance Corporation. 2011. Developing value: The business case for sustainability in emerging markets. Washington, DC: International Finance Corporation.

Kappel, Vivien, Annette Krauss, and Laura Lontzek. 2010. Over-indebtedness and microfinance–constructing an early warning index. Working Paper, Center for

Microfinance/University of Zurich.

Kiva Foundation. 2012. Microfinance can be a good tool for empowering women. *Kiva.org*, http://www.kiva.org/about/microfinance.

Mayoux, Linda. 2009. Gender and rural microfinance: Reaching and empowering women. Guide for Practicioners. Rome: International Fund for Agriculture Development, http://www.ifad.org/gender/pub/gender_finance.pdf.

Merslund, Roy, Bert D'Espallier and Isabelle Guerin. 2009. Women and Repayment in Microfinance, Working Paper RUME 2009-02, Provence University/Rume Project, Marseille, France.

OECD. 2012. Development: Aid to developing countries falls because of global recession. Working Paper, Organization for Economic Co-operation and Development, http://www.oecd.org/document/3/0,3746, en_21571361_44315115_50058883_1_1_1_1,00.html en_21571361_44315115_50058883_1_1_1_1,00.html.

Rhyne, Elizabeth. 2010. The state of global microfinance: How public and private funds can effectively promote financial inclusion. Hearing before the House Subcommittee on International Monetary Policy and Trade, 111th Congress, http://www.cmef.com/Document.Doc?id=773.

Roslan, A.H. and Mohd Zaini Abd Karim. 2009. Determinants of microcredit repayment in Malaysia: The case of Agrobank. H*umanity and Social Sciences Journal* 4: 45-52.

Skoll Foundation. 2012. Micro-lending. Social Edge, a program of the Skoll Foundation, http://www.socialedge.org/blogs/beyond-good-intentions/archive/2009/05/25/micro-lending.

Stock, J.H., and Watson, M.W. 2011. *Introduction to Econometrics*. Boston: Pearson Education.

Zeller, M. 1998. Determinant of repayment performance in credit groups: The role of program design, intragroup risk pooling, and social cohesion. *Economic Development and Cultural Change* 46: 599–621.

Does Changing Jobs Pay Off?
The Relationship between Job Mobility and Wages

By Amanda J. Huffman

Amanda Huffman completed her Master of Public Policy at the Georgetown Public Policy Institute in 2012. Adam Thomas, PhD, served as her advisor. Currently, she works as a Policy Analyst for Education First, a national, mission-driven strategy and policy consulting firm. Huffman also holds a BA in Religious Studies and Psychology from the University of Virginia.

ABSTRACT

Over the last three decades, wages have stagnated for most American workers, especially men. While demographic characteristics, education, and structural changes are common foci of wage analyses, more subtle factors may also play a role. One potential wage determinant is job mobility: the movement of an individual from job to job over the course of his or her career. The existing literature suggests that job mobility is associated with positive wage returns for workers early in their careers, but that the effect diminishes as workers gain experience and positive wage returns associated with job tenure grow stronger. Thus, the relationships between job mobility, tenure, and wages may depend upon work experience. This study uses a fixed-effects regression model and finds evidence of positive wage returns associated with high voluntary job mobility, which appear to diminish as workers gain experience. The study also finds that tenure is positively associated with higher wages for both low- and high-experience workers, not just for those workers with high work experience. These findings broadly indicate that some work patterns could result in higher average wages than others. A diverse portfolio of labor policies may therefore benefit workers who are just beginning their careers, whereas policies that foster longer tenure may create the greatest opportunity for wage growth among workers later in their careers.

I. INTRODUCTION

An increasing gap between productivity and compensation, along with persistent wage stagnation, are two US labor-market trends that mark the past three decades. Between 1979 and 2012, productivity in nonfarm business sectors grew by 86.2 percent, while real hourly compensation grew just 47.36 percent. When all industries and occupations are considered, an even more compelling fact emerges: wages for American men are no higher than they were in 1979 (BLS 2012).

In order to understand these trends, scholars often analyze individual characteristics like demographics and education, as well as structural changes in the labor market such as the decline in union participation. However, more subtle factors may play a role in wage levels over time. One potential wage determinant is job mobility: the movement of an individual from job to job throughout his career. It is possible that workers experience differences in wages according to whether or not, and when, they are highly mobile.

The influence of job mobility on wage levels has valuable implications for government policies that aim to assist workers. For example, job-search assistance programs that promote expanded access to general education and training may help workers change jobs multiple times throughout their careers. If job mobility is associated with wage gain, workers who take advantage of these programs could secure higher pay. Alternatively, policies that provide incentives for employers to train and invest in current employees may encourage workers to make fewer job changes. If job mobility is associated with wage loss, such policies could help workers avoid lower pay. In this study, I examine the relationship between job mobility and wages in order to understand which level of job mobility is associated with the highest wages according to how long workers have participated in the labor force.

II. BACKGROUND

Over the last several decades, trends in the US labor market suggest that job mobility may be on the rise. First, employment is shifting away from goods-producing industries toward service-oriented industries (Shin 2007). This pattern has eliminated a large number of jobs, especially for low-skilled workers, and has increased wage returns to jobs requiring skilled workers (Holzer et al. 2011). Second, the introduction of new technology has increased productivity, especially in the manufacturing sector, such that machines can now perform many jobs that once required additional human labor. The impact of this second trend reflects that of the first: job loss, especially for low-skilled workers. A third trend, increased competition owing to globalization, is also a cause of job loss, in this case to countries overseas where goods can be produced cheaper than in the US (Holzer et al. 2011).

A fourth major trend concerns how employers are restructuring their business models to accommodate the aforementioned labor market changes. Union participation is declining, which leaves workers with less collective bargaining power, decreased wages, and lower job stability (Holzer et al. 2011). As a result, employers have greater ability to terminate employees and workers have less incentive to stay in any one particular job. Additionally, to cut costs and remain flexible, businesses are employing more contract and temporary workers. Consequently, employer-employee relationships are much easier to dissolve, increasing the probability of job separation rates (Shin 2007; DOL 2011). Given these transformations in the US labor market, workers and policymakers alike may benefit from a better understanding as to what kind of job changes, as well as how often and when they are made, are most likely to create optimal wage opportunities for workers.

III. LITERATURE REVIEW

Job Mobility Measurement

In order to understand how the job mobility patterns of an increasingly mobile labor market may affect wage levels, two aspects of job mobility measurement call for clarification. First, studies employ two different measures to capture job mobility effects: 1) the short-term wage change that results from moving from one job to the next, and 2) the long-term wage change associated with the cumulative number of times an individual switches jobs throughout his career. Second,

"... workers and policymakers alike may benefit from a better understanding as to what kind of job changes ... are most likely to create optimal wage opportunities for workers."

most studies emphasize that job mobility effects depend upon whether job separations are made voluntarily or involuntary (Bartel and Borjas 1981; Topel and Ward 1992; Light and Ureta 1992; Keith and McWilliams 1995; Fuller 2008).

Voluntary Job Changes

Employee-initiated, voluntary job changes are associated, on average, with positive, short-term wage gains (Light 2005). Bartel and Borjas (1981) find that young and mature men who quit their jobs experience short-term wage increases of 11 cents and 3 cents an hour, respectively, relative to comparable men who did not quit their jobs. Further, Fuller (2008) finds that voluntary job mobility is associated with wages approximately 3 percent higher over the long term for job changes that occur within the first five years of potential work experience.

Involuntary Job Changes

In contrast, employer-initiated, involuntary job changes are associated, on average, with negative, short-term wage losses (Light 2005). This makes logical sense, as the types of workers who experience involuntary job separations may also be the types of workers, on average, for whom finding and keeping employment is

"This study fills a gap in cumulative job mobility literature by closely examining how the relationship between job mobility and wages is linked to tenure and experience."

more difficult. In addition, a worker who does not initiate a job separation is less likely to have been searching for an alternative, higher-wage job. Bartel and Borjas (1981) find that being laid off decreases the short-term wages of young and mature men by 2 cents and 19 cents an hour, respectively, and Fuller's (2008) cumulative study reveals that involuntary job mobility is generally associated with lower wages over the long term.

Having clarified job mobility measurement, I now turn to theoretical literature that investigates why job mobility may be an important factor for workers' wage trajectories. These studies contain two schools of thought, the first of which is search theory. Search theory suggests that "job-shopping" is associated with wage changes. The premise of this perspective is that workers are constantly searching for higher wage opportunities and, to the extent that their searches are successful in creating suitable employer-employee "matches," workers will experience positive, short-term wage returns from job mobility (Bartel and Borjas 1981; Antel 1986; Keith and McWilliams 1999). For example, Topel and Ward (1992) find that white men hold an average of seven jobs during their first 10 years in the labor market, and more than a third of wage growth during this period is accounted for by job mobility resulting from successful job searches.

Firm-specific human capital theory is the focus of a second school of thought in the literature. Studies in this camp examine the relationship between wages and the accumulation or loss of skills resulting from job changes. Theoretically, worker investments in firm-specific human capital result in lower job mobility and increased job tenure, which is the length of time spent with a particular employer (Antel 1986). For example, Fuller (2008) finds that cumulative overall job mobility is negatively correlated with wage growth, in part because highly mobile workers are not able to take advantage of positive job tenure effects.

In addition to the two schools of thought summarized above, job mobility literature suggests that the impact of job separations on wages depends upon when job changes occur. As cited earlier, Bartel and Borjas (1981) find that the consequences of quitting and being laid off are different for young and mature men. Larger wage gains are associated with voluntary job mobility (11 percent vs. 3 percent) and smaller wage losses are associated with involuntary job mobility (2 percent vs. 19 percent) for young workers compared to older workers. In addition to their short-term findings, Bartel and Borjas also conclude that increased job mobility later in life is associated with less long-term wage growth.

The studies summarized above inform this study in several ways. First, as the literature implies that the relationship between job mobility and wage change depends upon the reason for the job separation, I classify job separations into voluntary and involuntary categories. Nevertheless, I concentrate on the potential impact of voluntary job separations as voluntary job changes involve a choice by the worker as to what is best for his earnings trajectory. Second, studies that emphasize long-term, "cumulative," job-mobility wage effects are less common than those that focus on short-term, "job-to-job" mobility wage effects. As the cumulative approach offers a more complete picture of the relationship between job mobility and wages over the span of a worker's entire career, and to address a clear gap in the literature, I focus on long-term job mobility. Third, considering the prominent role that job tenure seems to play in determining the relationship between job mobility and wages, I follow the precedent of previous work and include tenure as a key independent variable in my analysis. Finally, as the benefits of job mobility appear to be realized in the early career years, with benefits to tenure becoming more important as individuals become older and more experienced, the length of time workers have participated in the labor force is a key consideration in this study (Bartel and Borjas 1981; Mincer 1986).

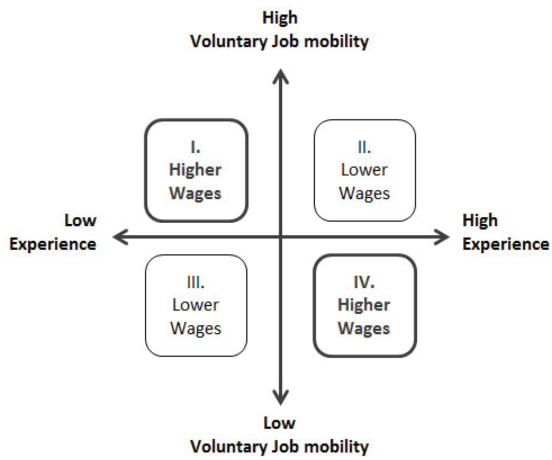
Despite the broad consistencies described above, academics have not reached a consensus as to which job mobility patterns are the most, or least, favorable for wage outcomes, leaving room for further research as to which level of job mobility is associated with the highest wages. This study fills a gap in cumulative job mobility literature by closely examining how the relationship between job mobility and wages is linked to tenure and experience. The conceptual model and hypotheses in the following section suggest the possibility for optimal levels of voluntary job mobility and tenure for workers at different points in their careers.

IV. CONCEPTUAL FRAMEWORK & HYPOTHESIS

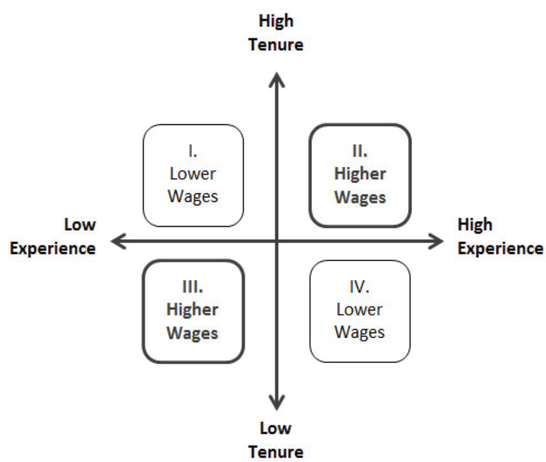
As indicated in the previous section, the literature suggests an inverse relationship between job mobility and tenure. This makes sense conceptually, as workers with high job mobility are likely to have lower tenure than their low job mobility counterparts. Conversely, workers with high tenure likely make fewer job changes than those with low tenure. If job mobility and tenure have an inverse relationship, in the extreme case, they may also be pulling wages in opposite directions. For example, workers who gain wages from frequently changing jobs may lose wages by staying with a single employer for a prolonged period of time. Furthermore, in theory, wage returns to different levels of job mobility and tenure are also related to experience, with returns to tenure growing stronger as workers move into the later stages of their careers (Bartel and Borjas 1981; Topel and Ward 1992; Fuller 2008).

Figure 1. Conceptual Model

A. Hypothesized Relationship between Voluntary Job Mobility and Wages by Experience



B. Hypothesized Relationship between Tenure and Wages by Experience



The conceptual model for this study, Figure 1, extends these theoretical relationships among wages, job mobility, tenure, and experience by suggesting that high voluntary job mobility and high tenure have opposite impacts on wages. Figure 1A depicts four quadrants predicting the relationship between low and high voluntary job mobility and wages according to experience level, while Figure 1B depicts four quadrants predicting the relationship between low and high tenure and wages according to experience level.

Based on this conceptual model, I make four predictions. As represented in quadrants I and IV of Figure 1A, I hypothesize that:

1) Among workers with low experience, wages are higher for those with high voluntary job mobility compared to those with low voluntary job mobility.

2) Among workers with high experience, wages are higher for those with low voluntary job mobility compared to those with high voluntary job mobility.

Conversely, as depicted in quadrants II and III of Figure 1B, I further hypothesize that:

3) Among workers with low experience, wages are higher for those with low tenure compared to those with high tenure.

4) Among workers with high experience, wages are higher for those with high tenure compared to those with low tenure.

The empirical analyses described in the next several sections test these hypotheses.

V. DATA & METHODS

Data Source

Analyses of cumulative job mobility require data that track respondents for a substantial portion of their careers. I use data from the 1979 panel of the National Longitudinal Survey of Youth (NLSY79). In the NLSY79, respondents were interviewed every year from 1979 to 1994 and then every other year after 1994. When first interviewed in 1979, respondents' age ranged from 14 to 22. In 2008, the most recent year for which NLSY data are publicly available, respondents' age ranged from 43 to 51. Among a host of other information, the survey contains a detailed work history for each respondent, including hourly wages for each job held, reasons for job changes, job tenure, and length of labor force participation.[1]

Analysis Plan

To isolate the relationship between job mobility and wages, this study uses person and year fixed effects to eliminate all potential omitted variable bias associated with time-invariant individual characteristics. Fixed effects controls for relevant, easy-to-measure factors such as gender, minority status, cognitive skills, and native-born status. Further, the specification also controls for person-specific characteristics, such as ability, that are more difficult to measure, are arguably fixed over time, and are potentially correlated with both wages and job mobility. All models also include time fixed effects, which control for omitted factors that vary over time but are common to all observations, such as the general state of the US economy and the unemployment rate.

The most basic empirical model used in this analysis specifies the log of real hourly wages as a function of job mobility, tenure, experience, and several time-variant control variables, where α_i and δ_t represent individual and year fixed effects, respectively, and ε_{it} represents an error term that varies within people, over time.

[1] For more detailed information on this study's data, variable descriptions, and data manipulation methods, see Huffman (2012).

Empirical Model

$$\textit{ln real hourly wages}_{it} = \beta_0 + \beta_1\textit{Job mobility} + \beta_2\textit{Tenure}_{it} + \beta_3\textit{Tenure}^2_{it} + \beta_4\textit{Experience}_{it} + \beta_5\textit{Experience}^2_{it} + \beta_6\textit{High school}_{it} + \beta_7\textit{Greater than high school}_{it} + \beta_8\textit{Age}_{it} + \beta_9\textit{Age}^2_{it} + \beta_{10}\textit{Percent time employed}_{it} + \beta_{11}\textit{Married}_{it} + \beta_{12}\textit{Occupation}_{it} + \beta_{13}\textit{Industry}_{it}\ \alpha_i + \delta_t + \varepsilon_{it}$$

I describe the key variables included in this fixed-effects model below.

Dependent Variable

The dependent variable for this analysis is the natural log of real hourly wages (in 2008 dollars). The real hourly wage measurement is the weighted average of hourly wages for up to five jobs held by a worker in any given year, multiplied by the Consumer Price Index (CPI) to account for inflation.

Key Independent Variables

Overall job mobility is a continuous measure equal to the number of job changes an individual has ever made up to the point of survey administration in any given year. While I concentrate on voluntary job mobility in most analyses, I use overall job mobility in descriptive statistics and in the base regressions depicted in Tables 3 and 4.

Voluntary job mobility and *involuntary job mobility* capture job changes that are either employee or employer initiated, respectively. For each job change, the NLSY79 records the reason that the respondent left his job, which allows for the characterization of each job separation as either voluntary or involuntary.

Tenure is a measure of how long, in years, an individual has been employed by the same employer. The literature suggests that wages increase as workers gain skills associated with a specific employer (Bartel and Borjas 1981). In addition, tenure has a conceptual correlation with job mobility; that is, the more tenure a worker accumulates, the lower his predicted overall job mobility. Research also suggests that tenure provides diminishing wage returns; thus, I include a quadratic term, tenure^2, in anticipation of a nonlinear relationship between tenure and wages.

My analysis also includes *high voluntary job mobility* and *high tenure* as dummy variables. These variables are set equal to one for person-years with voluntary job mobility and tenure at or above the median of their respective continuous variable counterparts (Fuller 2008).

Experience measures the number of years an individual has ever worked as of a given year. As workers gain more experience in the labor market, it is expected that their wages will increase. Experience is also negatively correlated with overall job mobility; that is, workers generally change jobs more in the early stages of their careers compared to the later stages (Topel and Ward 1992). As is the case with tenure, I include a quadratic term, experience^2, to account for a well-documented nonlinear relationship between experience and wages. Finally, I divide my sample into low- and high-experience groups, with low experience including those person-years with less than 10 years of experience, and high experience including those person-years with at least 10 years of experience. I determine low- and high-experience groups at the 10-year mark to mirror existing studies that use 10 years as a benchmark for the end of the "early-career" period (Mincer 1986; Topel and Ward 1992).

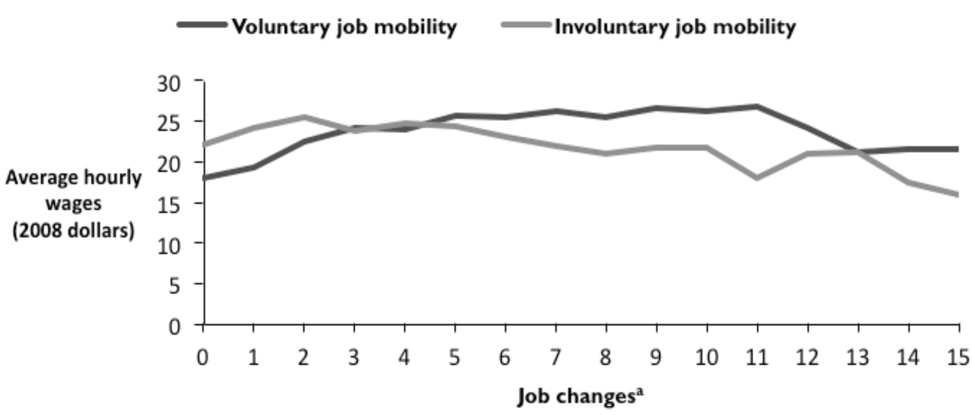
Control Variables
Whether workers are full-time, part-time, or marginally attached to the labor force is likely correlated to both wages and job mobility. Thus, I include *percent time employed* as a measure of labor force attachment equal to the number of hours or weeks spent working in a particular year. *Education, age, married, occupation,* and *industry* are additional control variables that account for worker skill level, age, marital status and type of work.

VI. RESULTS

Descriptive Results

Voluntary Job Mobility
Figure 2 displays average hourly wages according to workers' first 15 voluntary and involuntary job changes. The nonlinear trend line suggests that the relationship between voluntary job mobility and wage level may depend upon the number of voluntary job changes. Specifically, average hourly wages appear to increase for the first three to five changes, increase more gradually for the next five to six changes, and decrease after 11 job changes. On the other hand, involuntary job mobility only appears to be associated with wage increases for up to two involuntary job changes, at which point each additional involuntary job change is associated, on average, with lower wages.

Figure 2. Average Hourly Wages by Job Changes, Men in the NLSY79, 1979-2008



[a] The maximum number of voluntary job changes is 29 and the maximum number of involuntary job changes is 27. 97% and 99% of workers made 15 or fewer voluntary and involuntary job changes, respectively.

Table 1. Average Hourly Wages According to Voluntary Job Changes and Experience, Men in the NLSY79, 1979-2008

Low experience: < 10 years

	# voluntary job changes[a]	Observations	Sample frequency	Mean wage	Std. dev.
Low voluntary job mobility	0	1,242	5.27%	16.93	10.93
	1 to 3	5,380	22.84%	20.18	13.2
High voluntary job mobility	4 to 9	7,308	31.02%	23.59	14.65
	≥ 10	2,489	10.57%	21.65	13.7
Average		-	-	20.59	13.12
Total observations		16,419	69.70%		

High experience: ≥ 10 years

	# voluntary job changes[a]	Observations	Sample frequency	Mean wage	Std. dev.
Low voluntary job mobility	0	222	0.94%	23.48	11.88
	1 to 5	3,301	14.01%	27.53	15.04
High voluntary job mobility	6 to 9	2,196	9.32%	28.92	16.87
	≥ 10	1,420	6.03%	26.55	14.94
Average		-	-	26.62	14.68
Total observations		7,139	30.30%		

[a] The median for voluntary job mobility is 4 for the low-experience group and 6 for the high-experience group. Throughout this study, these medians define low and high voluntary job mobility groups within each experience group.

Voluntary Job Mobility and Experience

In order to further examine the relationship between number of voluntary job changes and wage levels, Table 1 displays average hourly wages for workers according to number of voluntary job changes and years of work experience. Specifically, Table 1 presents job changes for low- and high-experience groups, subdivided into two groups each for voluntary job changes below the median and above the median. In later analyses, I only break my sample into four voluntary-job-change categories. In Table 1, however, I create eight different categories in order to provide a more detailed comparison of average wages for workers with different levels of job mobility and experience.

Among workers with less than 10 years of work experience, those with between four and nine voluntary job changes experience the highest wages ($23.59). A two-sample mean comparison test reveals that even the smallest difference between the high and low voluntary job mobility groups ($21.65 and $20.18, respectively) is statistically significant ($p<0.001$). The finding of higher wages for low-experience workers with high voluntary job mobility provides evidence in support of Hypothesis 1.

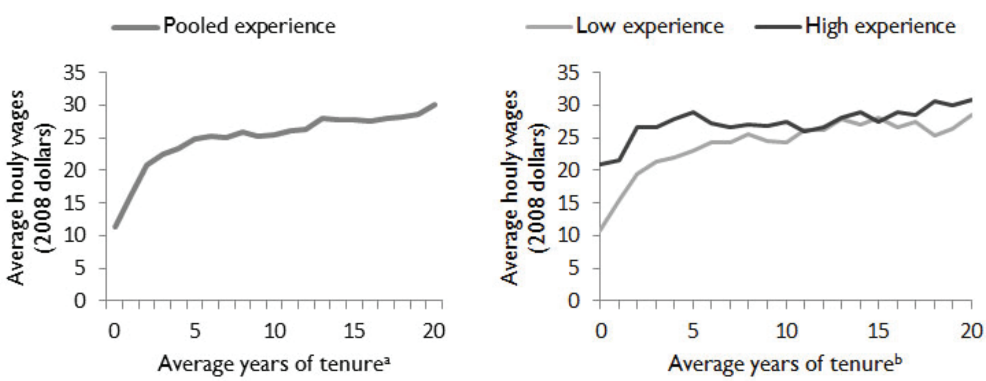
Table 1 also reveals that, for workers with at least 10 years of experience, workers in the high voluntary job mobility category with between six and nine voluntary job changes experience the highest wages ($28.92). Wages for high-experience workers in this middle category are significantly higher (p=0.001) than the wages of workers in either low voluntary job mobility group. These findings are somewhat at odds with Hypothesis 2. In contrast, high-experience workers with 10 or more voluntary job changes have significantly lower wages ($26.55, p=0.04), on average, than workers in the high-experience group with one to five voluntary job changes ($27.53).

The complex relationship between voluntary job mobility and wages for experienced workers indicates that support for this study's hypotheses may be limited by the definition of "high voluntary job mobility" as having made at least the median number of job changes, or by defining "high experience" as having at least 10 years of experience. However, in general, Table 1 provides suggestive evidence that the relationship between voluntary job mobility and wages may depend upon the number of voluntary job changes and when in workers' careers these changes occur.

Tenure and Experience

Having examined the relationship between voluntary job mobility and wages, I now examine the relationship between years of tenure and wages. As depicted by the trend line of the pooled-experience group in Figure 3, tenure, on average, appears to be positively associated with wages. Interestingly, tenure is also positively associated with wages for disaggregated low- and high-experience groups. This finding indicates a potential challenge to Hypothesis 3, which posits a negative relationship between tenure and wages for low-experience workers. Despite this result, Figure 3 does provide support for Hypothesis 4, which implies

Figure 3. Average Hourly Wages by Tenure, Men in the NLSY79, 1979 – 2008



[a] For the pooled-experience group, the maximum number of years of tenure is 33. 95% of workers averaged 20 or fewer years of tenure. 73% averaged 10 or fewer years, and 50% averaged 5 years or fewer. Although not shown here, wages steadily decline for the 5% of the sample who averaged more than 20 years of tenure.
[b] For the low-experience group, the maximum number of years of tenure is 28. 99% of workers averaged 20 or fewer years of tenure. For the high-experience group, the maximum number of years of tenure is 33. 87% of workers averaged 20 or fewer years of tenure.

a positive relationship between tenure and wages for high-experience workers.

Voluntary Job Mobility, Tenure, and Experience

For a combined examination of the relationship between wages, voluntary job mobility, tenure, and experience, Table 2 provides descriptive statistics for average hourly wages within low- and high-experience groups, which are further subdivided into both voluntary job mobility and tenure groups.

Two findings in the pooled-tenure group provide support for Hypotheses 1 and 2. Advancing Hypothesis 1, among low-experience workers, those with high voluntary job mobility have significantly higher wages ($23.10) than those with low voluntary job mobility ($19.56). Conversely, in partial contradiction to Hypothesis 2, among high-experience workers, those with high voluntary job mobility also have higher wages ($28.01) than those with low voluntary job mobility ($27.27). However, the fact that the high voluntary job-mobility wage premium for low-experience workers is five times greater than the same wage premium for high-experience workers suggests that high voluntary job mobility produces greater earnings opportunities for low-experience workers relative to high-experience workers. Ultimately, though, these relationships may mask a more interesting story revealed when different levels of tenure are considered.

Thus, I now turn to the disaggregated low- and high-tenure groups, beginning with the low-tenure group. Among workers with low experience and low tenure, those with high voluntary job mobility earn wages that are $6.35 higher than their low voluntary job mobility counterparts. In contrast, there is no significant difference between low and high voluntary job mobility groups for high-experience workers. In the high-tenure, low-experience group, workers with high voluntary job mobility earn $2.11 more an hour than their low voluntary job mobility counterparts. Interestingly, a similar wage premium is associated with high voluntary job mobility among high-experience workers ($2.03).

Finally, I consider findings across tenure groups. Among low-experience workers, the wage premium associated with high voluntary job mobility for low-tenure workers ($6.35) is greater than the premium associated with their high-tenure counterparts ($2.11). In contrast, among high-experience workers, the wage premium associated with high voluntary job mobility for low-tenure workers ($0.79) is less than that earned by their high-tenure counterparts ($2.03).

The disaggregated results expose two interesting wage relationships according to 1) experience within tenure groups and 2) tenure within experience groups. First, for workers in the low-tenure group, high voluntary job mobility is associated with a wage premium for low-experience workers only. For workers in the high-tenure group, the voluntary job mobility wage premium is approximately the same for low- and high-experience workers.

Table 2. Average Hourly Wages by Voluntary Job Mobility, Tenure, and Experience Groups, Men in the NLSY79, 1979 – 2008

		Average hourly wage (Std. dev.) n	Average hourly wage (Std. dev.) n
Pooled tenure (n = 23,558)			
		Experience	
		Low	High
Voluntary job mobility	Low	$19.56 (12.87) 6,622	$27.27 (14.88) 3,523
	High	$23.10 (14.44) 9,797	$28.01 (16.20) 3,616
	Mean difference	$3.54***	$0.74**
Low tenure (n =12,205)			
		Experience	
		Low	High
Voluntary job mobility	Low	$14.12 (9.23) 3,116	$26.26 (15.46) 1,237
	High	$20.47 (13.64) 5,449	$27.04 (16.48) 2,403
	Mean difference	$6.35***	$0.79
High tenure (n =11,353)			
		Experience	
		Low	High
Voluntary job mobility	Low	$23.99 (13.69) 3,506	$27.79 (14.56) 2,286
	High	$26.10 (14.75) 4,348	$29.82 (15.49) 1,213
	Mean difference	$2.11***	$2.03***

Note: As is the case with low and high voluntary job mobility, low tenure and high tenure are determined according to median tenure in low- and high-experience groups, 4.21 and 9.17 years, respectively. Mean differences are statistically significant at the *0.10, **0.05, or ***0.01 significance levels.

Second, high tenure is associated with a decrease in voluntary job mobility wage premiums for low-experience workers, but with an increase in these wage premiums for high-experience workers.

In sum, the data in Table 2 reinforce the notion that the strength of the relationship between high voluntary job mobility and wages diminishes with experience and varies according to length of tenure. The apparent complexity of these relationships among wages, voluntary job mobility, tenure, and experience provides impetus for multiple regression analyses that incorporate all of these factors into a single model.

Regression Results

To more precisely estimate the relationship between voluntary job mobility and wages, I run nine regression models that control for person and year fixed effects, education, age, percent time employed, marital status, industry, and occupation. The results of these regressions are presented in Tables 3 and 4.

Table 3 pools low- and high-experience observations and controls for experience, while Table 4 displays models with only low-experience observations and models with only high-experience observations. For each of the experience groups (pooled, low, and high), I specify three models. In terms of job mobility measures, the regression in the first column for each group includes only overall job mobility; the regression in the second column replaces overall job mobility with voluntary and involuntary job mobility; and the regression in the third column replaces the disaggregated job mobility variables with dummy variables for high voluntary job mobility and high involuntary job mobility. The regression in the third column of each group also replaces the continuous tenure variables, tenure and tenure^2, with a dummy variable for high tenure.

Job Change Timing

The coefficient on overall job mobility in the first regression of each experience group is positive and is not statistically significant in the pooled- and low-experience groups, but the coefficient is negative and statistically significant in the high-experience group (-0.040). This finding indicates that the relationship between job changes and wages cannot be estimated with precision in the early stage of workers' careers, but that the relationship becomes distinctly negative for experienced workers. Specifically, it is estimated that, on average, for each additional job change, wages for workers with at least 10 years of experience are expected to decrease by 4 percent.

Reason for Job Change

The relationship between job mobility and wages also differs according to whether the job changes are voluntary or involuntary, and the sign and precision of the variables' coefficients both change according to experience. In both the pooled- and low-experience groups, a voluntary job change is associated, on average,

Table 3. Fixed-Effects Regression of Voluntary Job Mobility on Log Hourly Wages

	Pooled Experience		
Independent variables	Model 1	Model 2	Model 3
Overall job mobility	0.003		
	(0.003)		
Voluntary job mobility		0.012***	
		(0.004)	
Involuntary job mobility		-0.014**	
		(0.006)	
High voluntary job mobility			0.083***
			(0.023)
High involuntary job mobility			-0.039
			(0.029)
Tenure	0.026***	0.026***	
	(0.002)	(0.002)	
$Tenure^2$	-0.001***	-0.001***	
	(0.000)	(0.000)	
High tenure			0.065***
			(0.009)
Experience	0.022	0.02	0.037**
	(0.015)	(0.015)	(0.016)
$Experience^2$	0.001*	0.001*	0.000
	(0.001)	(0.001)	(0.001)
HS education	-0.105***	-0.109***	-0.112***
	(0.030)	(0.031)	(0.031)
Greater than HS education[a]	0.027	0.019	0.013
	(0.053)	(0.053)	(0.053)
Age	0.104***	0.103***	0.109***
	(0.013)	(0.013)	(0.013)
Age^2	-0.002***	-0.002***	-0.002***
	(0.000)	(0.00)	(0.000)
Percent time employed	0.226***	0.228***	0.232***
	(0.040)	(0.040)	(0.041)
Married	0.107***	0.105***	0.108***
	(0.013)	(0.013)	(0.013)
Constant	0.728***	0.754***	0.682***
	(0.199)	(0.200)	(0.197)
Observations (person-year)	23,558		23,558
Adj. R-squared	0.693		0.691

Note: These regressions were estimated using 30 years (1979-2008) of panel data from the NLSY79. All models control for person and year fixed effects, as well as person-year industry and occupation. Robust standard errors are reported in parentheses under coefficients. Individual coefficients are statistically significant at the *0.10, **0.05, or ***0.01 significance levels.

[a] As it is widely accepted that having the equivalent of a high school education increases wages relative to high school dropouts, which is the omitted education category in all models, the negative coefficient on HS education is likely negatively biased owing to an omitted variable correlated with both wages and education. It is plausible that, among those who have the equivalent of a HS education but who do not go on to attend college, there are a considerable number who initially dropped out of school and later returned to high school or earned a GED. Compared to those HS dropouts who entered the workforce immediately and consistently earned wages, the late HS education equivalents, characterized as having a HS education in my sample, could have lower wages than dropouts due to some unobserved factor, e.g., motivation. In other words, a likely explanation for the negative coefficient on HS education is that the variation between late HS education equivalents and HS dropouts is driving the results, rather than the variation between regular HS graduates and dropouts.

Table 4. Fixed-Effects Regression of Voluntary Job Mobility on Log Hourly Wages, by Job Experience Category

Independent variables	Low Experience (< 10 years)			High experience (≥ 10 years)		
	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Overall job mobility	0.004			-0.040**		
	(0.004)			(0.017)		
Voluntary job mobility		0.015***			-0.026	
		(0.005)			(0.017)	
Involuntary job mobility		-0.013*			-0.092***	
		(0.007)			(0.031)	
High voluntary job mobility			0.077***			-0.034
			(0.026)			(0.044)
High involuntary job mobility			-0.05			-0.165*
			(0.032)			(0.088)
Tenure	0.046***	0.047***		0.013***	0.013***	
	(0.004)	(0.004)		(0.005)	(0.005)	
$Tenure^2$	-0.002***	-0.002***		-0.000***	-0.000***	
	(0.000)	(0.000)		(0.000)	(0.000)	
High tenure			0.098***			0.041**
			(0.013)			(0.018)
Experience						
$Experience^2$						
HS education[17]	-0.122***	-0.125***	-0.133***	-0.048	-0.045	-0.048
	(0.044)	(0.045)	(0.044)	(0.032)	(0.032)	(0.032)
Greater than HS education	0.020	0.012	-0.002	0.016	0.018	0.006
	(0.062)	(0.063)	(0.063)	(0.058)	(0.058)	(0.056)
Age	0.134***	0.131***	0.152***	0.142***	0.141***	0.151***
	(0.016)	(0.016)	(0.016)	(0.050)	(0.050)	(0.051)
Age^2	-0.002***	-0.002***	-0.002***	-0.001*	-0.001*	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Percent time employed	0.236***	0.237***	0.267***	0.033	0.038	0.061
	(0.047)	(0.047)	(0.048)	(0.044)	(0.044)	(0.043)
Married	0.110***	0.109***	0.111***	0.066*	0.065*	0.073**
	(0.017)	(0.017)	(0.017)	(0.036)	(0.036)	(0.036)
Constant	0.241	0.291	0.016	-1.231	-1.143	-1.582
	(0.223)	(0.224)	(0.221)	(1.654)	(1.652)	(1.735)
Observations (person-year)	16,419	16,419	16,419	7,139	7,139	7,139
Adj. R-squared	0.666	0.667	0.662	0.836	0.836	0.835

Note: These regressions were estimated using 30 years (1979-2008) of panel data from the NLSY79. All models control for person and year fixed effects, as well as person-year industry and occupation. Robust standard errors are reported in parentheses under coefficients. Individual coefficients are statistically significant at the *0.10, **0.05, or ***0.01 significance levels.

with an increase in wages between 1 and 2 percent, while an involuntary job change is associated, on average, with a wage decrease of approximately 1 percent. However, for high-experience workers, the coefficient on voluntary job mobility turns negative and is no longer significant (-0.026), while the coefficient on involuntary job mobility is highly significant and substantially greater in magnitude than its pooled- and low-experience counterparts (-0.092). This result suggests that the positive relationship between voluntary job mobility and wages decreases with experience, while the negative relationship between involuntary job mobility and wages becomes even more negative as workers advance in their careers.

Evaluating Four Hypotheses

The next four results of interest concern the coefficients on high voluntary job mobility and high tenure in the low- and high-experience groups, which I highlight to evaluate the four hypotheses set forth at the outset of this study. In support of Hypothesis 1, the coefficient on high voluntary job mobility in the low-experience group (0.077) is positive and highly significant, indicating that for workers with low experience, wages for workers with high voluntary job mobility are predicted to be approximately 8 percent higher than wages for workers with low voluntary job mobility.

In terms of Hypothesis 2, the coefficient on high voluntary job mobility for high-experience workers is negative, but this coefficient is imprecisely estimated (p=.0773). It is therefore difficult to draw any informative conclusions about this particular relationship using the results presented here.

The coefficient for high tenure in the low-experience group directly contradicts Hypothesis 3, that for workers with low experience, wages are higher for workers with low tenure compared to workers with high tenure. In fact, low-experience, low-tenure workers are expected to have, on average, wages that are 10 percent lower than low-experience, high-tenure workers.

Finally, the coefficient on high tenure in the high-experience group supports Hypothesis 4, indicating that high-experience workers with high tenure are expected to have wages that are approximately 4 percent higher than their counterparts with low tenure. In sum, not only do high-tenure workers realize wage premiums regardless of experience, but also, the relationship between high tenure and wages is greater in magnitude and precision for low-experience workers.

VII. DISCUSSION & CONCLUSION

Previous job mobility studies have largely focused on individual job changes, comparing wage changes of workers who switched jobs to those of workers who stayed in their current jobs. These “job-to-job” mobility studies find that, in general, voluntary job changes are positively associated with wages, whereas involuntary job changes are negatively associated with wages. A smaller number of studies have examined how job mobility may impact wage levels over time. These less common "cumulative" studies confirm that voluntary job mobility is associated with higher wage levels, yet they also indicate that any positive wage impact from voluntary job changes may diminish over time as wage returns to tenure become stronger. In this study, I hypothesize that voluntary job mobility and tenure, in the extreme, have the opposite effect on wages. Specifically, I evaluate four hypotheses:

> 1) Among workers with low experience, wages are higher for those with high voluntary job mobility compared to those with low voluntary job mobility.
>
> 2) Among workers with high experience, wages are higher for those with low voluntary job mobility compared to those with high voluntary job mobility.
>
> 3) Among workers with low experience, wages are higher for those with low tenure compared to those with high tenure.
>
> 4) Among workers with high experience, wages are higher for those with high tenure compared to those with low tenure.

I find that positive wage returns are associated with high voluntary job mobility that diminish over time. In particular, it is predicted that high voluntary job mobility is associated with wages that are approximately 8 percent higher for workers with less than 10 years of experience, but that high voluntary job mobility is not associated with either higher or lower wages for workers with at least a decade of work experience. I also find that high tenure is positively associated with higher wages for both low- and high-experience workers (10 percent and 4 percent, respectively), not just those workers with high work experience.

While my results are generally consistent with existing literature, one of the regression results diverts from previous research. Specifically, I find higher wage premiums for high-tenure workers with less, compared to more, experience. This may be because my conceptual model is too extreme in suggesting that high voluntary job mobility and high tenure are associated with opposite wage outcomes. My results indicate that voluntary job mobility and tenure are not so negatively correlated that 1) a high instance of one variable dictates a low instance of the other, and 2) for the same group of workers, a high instance of one variable being positively associated with wages indicates that a high instance of the other is negatively associated with wages. In fact, my results imply an overlap between categories that are conceptually mutually exclusive in my hypotheses. Future research that specifies a more subtle relationship between wages, voluntary job mobility, tenure, and experience may be enlightening.

Moreover, my regression results are based upon defining high voluntary job mobility and high tenure as simply being equal to or greater than the median number of voluntary job changes or years of tenure, respectively. Additional research that uses an alternate threshold or employs a more nuanced approach with more than two categories for each variable may be able to identify more precisely which mobility patterns are associated with the highest wages.

Finally, as the oldest workers in my sample were only 51, classifying all workers with at least a decade of work experience as "high experience" is not necessarily analogous to "late career." As the NLSY79 continues to be administered, additional research that divides workers into "early-," "mid-," and "late-" career workers based upon ten-year increments of experience should provide a more realistic picture of which job mobility patterns are associated with the highest wages.

I must also emphasize that, even after controlling for fixed effects and including several relevant control variables, unexplained variation in wages still exists. It is possible, then, that my results suffer from omitted variable bias. For example, some workers may have high instances of voluntary job changes because of health complications. This particular omission, which may be positively correlated with voluntary job mobility and negatively correlated with wages, would result in a negative bias in my job mobility coefficients.

In terms of policy implications, these findings broadly indicate that some job mobility patterns may result in

"... a diverse portfolio of labor policies stands to benefit workers who are just beginning their careers, whereas policies that foster increased tenure may create the greatest opportunity for wage growth among workers later in their careers."

higher average wages than others. In general, comparatively high numbers of voluntary job changes within the first 10 years of work experience are associated with higher average wages, lending credibility to policies—like job search assistance—that encourage workers to job-shop. Thus, policies that promote expanded access to general education and training opportunities may be a worthwhile investment to develop transferable skills in young workers.

However, the findings also indicate that tenure is positively correlated with wages for all workers, regardless of experience, providing support for policies that encourage employers to invest in their employees through firm-specific training. Perhaps the most appropriate conclusion is that a diverse portfolio of labor policies stands to benefit workers who are just beginning their careers, whereas policies that foster increased tenure may create the greatest opportunity for wage growth among workers later in their careers. However, since the results are relatively inconclusive as to their policy implications, they should be assessed with some caution.

In conclusion, heterogeneity in job mobility patterns may play a role in wage levels over time. Controlling for other factors, workers with the "best" mobility patterns will experience higher wages than their peers. More detailed research is needed to specify which cumulative mobility patterns maximize wages, and for whom. In particular, it may be that workers in specific industries or occupations are more likely than others to experience higher wages from voluntary job mobility or tenure. Additional research exploring which mobility patterns are associated with higher wages according to gender, race, skill level, and income level could lead to more targeted job search and training policies.

VIII. REFERENCES

Antel, J. J. 1986. Human capital investment specialization and the wage effects of voluntary labor mobility. *The Review of Economics and Statistics* 68, no. 3: 477-83. http://econpapers.repec.org/article/tprrestat/v_3a68_3ay_3a1986_3ai_3a3_3ap_3a477-83.htm.

Bartel, A. P., and G. J. Borjas. 1981. Wage growth and job turnover: An empirical analysis. *Studies in Labor Markets*. http://www.nber.org/papers/w0285.

Bureau of Labor Statistics. 2012. Major sector productivity and costs, http://www.data.bls.gov.

Bureau of Labor Statistics. 2012. Weekly and hourly earnings data from the Current Population Survey, http://www.bls.gov/data/#wages.

Department of Labor. 2011. Contingent workers special report. http://www.dol.gov/_sec/media/reports/dunlop/section5.htm.

Economic Policy Institute (EPI). 2012. The state of working America. November. http://stateofworkingamerica.org/files/book/Chapter4-Wages.pdf.

Fuller, S. 2008. Job mobility and wage trajectories for men and women in the United States. *American Sociological Review* 73, no. 1: 158. http://www.soc.washington.edu/users/brines/fuller.pdf.

Holzer, Harry J., Julia I. Lane, David B. Rosenblum, and Fredrik Andersson. 2011. Where are all the good jobs going? What national and local job quality and dynamics mean for U.S. workers. New York: Russell Sage Foundation.

Huffman, Amanda J. 2012. Does changing jobs pay off? The relationship between job mobility and wages. http://proquest.umi.com/pqdlink?did=2648112371&Fmt=7&clientId=79356&RQT=309&VName=PQD.

Keith, K., and A. McWilliams. 1999. The returns to mobility and job search by gender. Industrial and Labor Relations Review: 460-77. http://www.jstor.org/stable/2525145.

Keith, K., and A. McWilliams. 1995. The wage effects of cumulative job mobility. *Industrial and Labor Relations Review* 52, no.3: 121-37. http://www.jstor.org/stable/2524916.

Light, A. 2005. Job mobility and wage growth: Evidence from the NLSY79. *Monthly Labor Review* 128: 33. http://www.bls.gov/opub/mlr/2005/02/art5full.pdf.

Light, A. and M. Ureta. 1992. Panel estimates of male and female job turnover behavior: Can female nonquitters be identified? *Journal of Labor Economics* 10, no.2: 156-81. http://links.jstor.org/sici?sici=0734-306X%28199204%2910%3A2%3C156%3APEOMAF%3E2.0.CO%3B2-9&origin=repec.

Mincer, J. 1986. Wage changes in job changes. *Research in Labor Economics* 8a. http://www.nber.org/papers/w1907.

Shin, T. J. 2007. The impact of structural dynamics on job mobility rates in the United States. *Social Science Research* 36, no. 4: 1301-27. http://escholarship.org/uc/item/6112w1nw.pdf.

Topel, Robert H. and M. P. Ward. 1992. Job mobility and the careers of young men. *Quarterly Journal of Economics* 107, no. 2: 439-79. http://www.jstor.org/stable/2118478.

Renewable Energy at What Cost?

Assessing the Effect of Feed-In Tariff Policies on Consumer Electricity Prices in the European Union

By Christopher A. Klein, MA (Hons)

Christopher A. Klein is a 2012 graduate from the Georgetown Public Policy Institute's Master of Public Policy program, where he was awarded the "2012 Thesis Prize" for his research on European renewable energy support policies. J. Arnold Quinn, PhD, served as his thesis advisor. Klein holds an MA (Hons) in Economics and International Relations from the University of St. Andrews.

ABSTRACT

In the last two decades, feed-in tariffs (FIT) have emerged as the dominant policy instrument for supporting electricity from renewable sources in the European Union. This paper examines the effect of such feed-in tariffs on consumer prices for electricity. While a multitude of studies examine the effects of FIT policies on electricity prices within individual countries or across countries using complicated ex-post computer simulations, there are a dearth of rigorous ex-post, cross-country econometric analyses. Using 1992-2009 panel data across 20 European countries and a dynamic panel data model estimation, this paper analyzes the effect of FIT policies for electricity generated from wind and solar photovoltaic (PV) on electricity prices at the household consumer level. The analysis finds a mild association of the support level for wind energy with higher retail prices, but no price increase for solar PV support. This finding points toward the existence of a "merit-order effect" and, in particular, a strong "time-of-day" effect, where solar PV is able to replace more costly natural gas and petroleum generation because it is generated during times of peak demand, whereas electricity from wind is mostly generated at night when demand is low. However, the shares of solar PV electricity generated under the FIT are still very low; as the share of electricity generation that is covered by the FIT rises, adverse price effects may become more apparent. This paper also finds that feed-in tariffs for wind only increase retail prices in the presence of retail regulation, indicating that regulatory bodies may allow utility companies to charge higher prices in the presence of FIT payments, whereas utility

companies that are subject to retail competition are not able to pass on their additional costs to customers. In addition, the paper further finds that larger shares of electricity generated from hydro and nuclear power decrease retail rates, suggesting that, due to their similar cost profile, the same could be true for wind and solar PV in the long term, once a fleet of generation capacity from wind and solar PV is established and the initial capital costs are recovered.

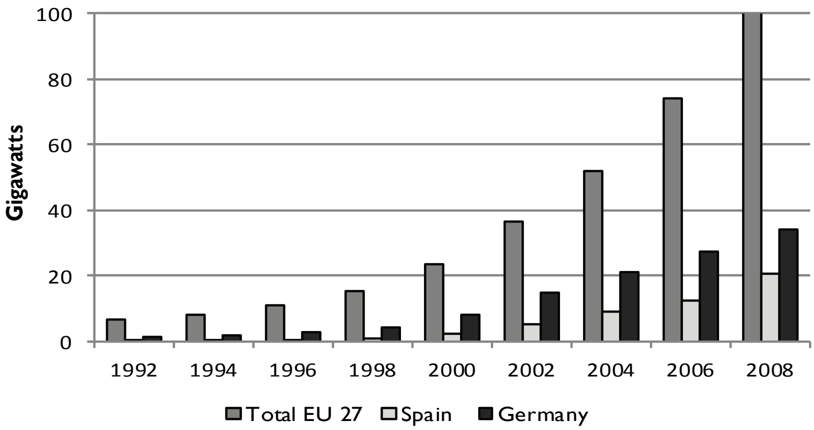
I. INTRODUCTION

The current consensus among European policymakers is that well-designed feed-in tariff (FIT) schemes are the most effective way of achieving the development of electricity-generating capacity from renewable sources (RES-E). This differentiates Europe from the United States, where Renewable Portfolio Standards have been the dominant RES-E support instrument. The European view is driven by the success of FIT policies in deploying renewable energy in Germany, Denmark, and Spain, among other countries, and is supported by a large body of academic evidence (EC 2008; OPTRES 2007; Lipp 2007; Butler and Neuhoff 2008; Lesser and Xu 2008; Alagappan, Orans, and Woo 2011). While unable to prescribe policy instruments to member states, the European Commission concluded in a 2008 communication to the European Parliament (EP) and European Council (EC) that FITs "achieve greater renewable energy penetration, and do so at lower costs for consumers" (EC 2008).

The success of FIT policies can largely be traced to the high level of security that FIT schemes provide for investors in RES-E generation. They are energy-supply policies that (1) impose obligations on utilities and grid operators to purchase the full output generated by qualifying renewable energy generators, (2) guarantee an above-market payment per unit output ($/kWh) for the full output of the system, (3) limit these special payments to a specified time period, and (4) differentiate payments between projects based on technology type, project size, the quality of the resource, or other project-specific variables. FITs thus mitigate future electricity market volatility, making it very likely that investors will recover large up-front capital investments (Butler and Neuhoff 2008; Lesser and Xu 2008). Since 1990, 24 EU member states have introduced FIT policies.[1] Over this time period, RES-E generation capacity in the EU-27 countries has developed rapidly, especially in those countries with established FIT programs (See Figure 1).

[1] Only Sweden, Poland, and Romania continue to rely solely on minimum quota regulations to achieve their EU targets.

Figure 1. Total Non-Hydroelectric RES-E Electricity Generation Capacity in the EU 27, Germany, and Spain



Although many scholars have conceded that FIT policies are the most effective way to promote RES-E (EC 2008; RES-financing 2011; OPTRES 2007), the increase in the share of RES-E in the electricity production mix has spurred debate about the cost of such policies. As FIT payments are normally above the spot-market price for electricity, policy makers are concerned with the effects FIT schemes may have on electricity prices and, subsequently, the competitiveness of European economies, inflation levels, and social welfare. Rising electricity costs were less of a concern when renewable energy targets were relatively low; however, as EU targets for RES-E generation have grown so has apprehension about rising electricity prices. Some countries have already responded to these concerns. Germany recently reduced its FIT payments for solar photovoltaic (PV) by 15 percent for the year 2012, while Spain retroactively cut its solar PV FIT program by up to €3 billion (BMU 2011; Mallet 2010).

There has not been a cross-country, ex-post econometric study of the negative effects of FIT regimes on consumer prices. In light of this dearth, this paper aims to illuminate the current debate on adverse economic effects of renewable energy promotion—in particular, feed-in tariffs—in EU member states. It provides a rigorous ex-post econometric analysis of the effect of technology-specific FIT legislations for wind and solar PV on electricity prices at the household consumer level across 20 European countries between 1992 and 2009.

"... policymakers are concerned with the effects FIT schemes may have on electricity prices and, subsequently, the competitiveness of European economies, inflation levels, and social welfare."

II. CONCEPTUAL FRAMEWORK & REVIEW OF THE LITERATURE

Like other network utilities, the electricity sector is characterized by extremely high up-front (i.e. fixed) capital costs, with considerably lower variable costs (Newberry 1999; Stoft 2002). Fixed costs reflect the capital necessary to build and maintain physical infrastructure such as generation facilities or high-voltage transmission networks. The structural costs for building generation facilities vary largely between countries. In particular, fixed costs depend on the differential between a country's average and peak electricity demand; the more extreme a country's electricity demand, the higher its fixed costs will be. Other factors affecting electricity costs include the cost of labor, the ability to obtain necessary permits, the cost and availability of financing, and geographic factors such as population density and distance between generation and demand centers. Unless they are subsidized or cross-subsidized from other sectors, household electricity prices generally allow investors in generation capacity to recover these capital costs (Newberry 1999).

In addition to the fixed costs, utility companies are faced with variable input costs for generating and supplying electricity. First and foremost among those are the variable costs of generating electricity, which companies pay in the form of wholesale prices, supply contracts, or as direct inputs in cases where they still have generation capacity of their own. These costs include the input costs for fossil fuels as well as labor costs required for operating power plants (Newberry 1999). Typically, the power portfolio is made up of different power suppliers, generating electricity through different technologies. The use of each generator depends on the amount of power it can supply at a certain price, and generators compete on the basis of the marginal (variable) cost of the plant. This is called "merit order," where different plants are ranked from low marginal cost (e.g. hydro) to high marginal cost (e.g. natural gas) (Fox-Penner 1997; Newberry 1999). In liberalized power markets, the wholesale price for electricity is thus determined by the marginal technology used. As most European electricity markets are only partially restructured, it is assumed that their operational decisions are determined by producers' variable costs as they would be in a liberalized wholesale market.

Within this framework, FITs increase utility companies' variable costs as they oblige companies to take off electricity generated by renewable producers at a pre-determined price (the fixed tariff) that is typically above the average spot-market price. This is in contrast to the logic of a wholesale market price determined by variable costs, as renewable generators are characterized by virtually zero variable cost. Therefore, it can be assumed that the obligation to pay feed-in tariffs to renewable generators raises companies' generation costs in the short-term, which they then likely pass on to consumers.

Numerous ex-ante studies estimate the additional costs of generating large shares of electricity from renewable resources. A 2011 study estimates the total cost of RES support in 2009 at the EU level to amount to approximately €35 billion (RES-Financing 2011). Additionally, for Germany alone, Frondel et al. (2008) calculate the discounted total net cost of subsidizing electricity production from wind and solar PV to be €20.5 billion and €53.3 billion respectively, for generators installed between 2000 and 2010. According to their account, German households paid a price mark-up due to the subsidization of green electricity of about 1.5 cent per kWh in 2008, amounting to about 7.5 percent of average household electricity prices.

In 2007, OPTRES predicted that a steady rise in average EU consumer prices for electricity was necessary to finance RES-E deployment over the next 10 years, foreseeing an increase from 2.1 €/MWh in 2005 to a rate between 5.0 €/MWh and 7.7 €/MWh for the period 2005 to 2020. Using a quantitative electricity market model that accounts for factors such as oligopolistic behavior, emission trading, and restricted cross-border transmission capacities, Traber and Kemfert (2009) also find an upward price effect of the German FIT.

Among the relatively few ex-post studies that have analyzed the price effects of FIT policies, Gual and del Rio (2007) assess the effect of the

"... it can be assumed that the obligation to pay feed-in tariffs to renewable generators raises companies' generation costs in the short-term, which they then likely pass on to consumers."

Spanish FIT between 1999 and 2003 in terms of additional costs paid by the consumer for renewable compared to conventional electricity (i.e. the share of RES-E promotion of the electricity bill). Their study finds that the additional cost for the consumer increased annually by 23 percent during the period considered.

However, some properties of RES-E generation could also potentially counteract the upward-price effect associated with FITs. The marginal cost of most renewable electricity generation is zero or close to zero. Once a plant has been put in place, the generator produces electricity at almost no extra cost. As utilities are mandated to take off this electricity and pay the generator a fixed price, this electricity is practically free (in the sense of already paid for). The marginal technology determining the wholesale price therefore depends on the level of "residual demand," defined as the electricity demanded minus the feed-in of electricity from RES-E. If the residual demand is low, the marginal power plant is less expensive than if the residual demand is high. Dependent on the price elasticity of power demand, RES-E generation pushes more expensive marginal plants (e.g., natural gas, petroleum, etc.) out of the market, which not only displaces the generation costs of these generators but also reduces inframarginal rents earned by all non-marginal sellers in the spot market. High feed-in of RES-E thus shifts the supply curve for electricity to the right, resulting in lower wholesale electricity prices. This is what Ragawitz, Sensfuß, and Barbose (2008) term the "merit-order effect."

In particular, there could be a substantial "time-of-day" effect that is related to the merit order. The merit-order effect could be particularly strong during peak time if RES-E was able to replace extremely expensive "super-peaker" plants, whereas it could be almost negligible during times of low demand. In extreme cases, the merit-order-related price savings across the entire electricity market could outweigh the costs of paying renewable generators above-market rates, depending on the magnitude of the tariff and the price reduction.

Some empirical analysis has confirmed that more RES-E supply can decrease spot-market prices in practice. Weigt (2009) finds that wind generation had a downward impact on both spot-market prices and generation costs in Germany for the period of 2006 to 2008. During the observation period, the study estimates a total savings of €4.1 billion due to wind power fed into the grid. Traber and Kemfort (2011), using a mixed complementary program computational model, also find that higher wind supply reduces German market prices by more than 5 percent. Their model estimates that the reduction in the spot-market price for electricity is 0.37 Eurocents per kWh.

Similar results have also been found in Spain. Gelabert, Labandeira, and Linares (2011), using a multivariate regression model of hourly electricity prices for 2005 to 2009, find that a marginal increase of 1 GWh of electricity from renewable sources is associated with a reduction of almost 4 percent (€2 per MWh) in wholesale electricity prices. Likewise, Jonsson, Pinson, and Madsen (2010) use a non-parametric least squares model regressing hourly area spot-prices on wind power forecasts for January 2006 to October 2007 to show that positive wind forecasts result in lower spot-market prices for the DK-1 price area of the Nordpool electricity area.

Comparing potential cost savings from higher RES-E feed-in to direct costs of the FIT, Bode (2006) shows that power costs might decrease due to FIT schemes such as the German EEG under certain conditions. Similarly, Rathmann (2007) shows that the support for renewable energy created by the German feed-in tariff can result in lower electricity prices. Ragawitz, Sensfuß, and Barbose (2008), offering a detailed analysis of the price effects of renewable electricity generation on German spot-market prices between 2001 and 2006, find a considerable reduction in wholesale electricity prices associated with higher levels of RES-E fed into the grid. Furthermore, they find that in 2006 cost savings due to RES-E feed-in actually outweighed the direct costs of the FIT. Similarly, De Miera, del Rio Gonzalez, and Vizcaino

(2008), using hourly historical data, find that the reduction of the wholesale price of electricity as a result of more RES-E generation being fed into the grid is greater than the increase in consumer prices for electricity that arise from the FIT scheme.

In contrast, OPTRES (2007) projects that the direct effect of the FIT will outweigh the indirect reduction of wholesale prices. Although it estimates that the total amount of avoided fossil fuels will reduce costs for the EU27 by €23 billion from the year 2020 onward, the study expects the costs of RES-E generation to be higher.

The merit-order effect has further implications for the long-term effects of supporting RES-E. FIT schemes do not run infinitely; the contract duration typically lies somewhere between 12 and 20 years. The scheme is supposed to enable developers to recover their capital investment over the contract period. Once the contract expires, the price of electricity from those generators is driven by companies' (actual) variable costs. Therefore, the long-term price effects of FITs may be much more beneficial to consumers if they result in a lot of renewable generation capacity installed that will still be around after the FIT contract period has ended.

This long-term effect also applies to hydro and nuclear facilities. Similar to RES-E technologies such as wind and solar PV, both technologies have extremely low operating costs compared to the high up-front capital investments required to build the plants. Once these plants have recovered the initial capital investments, they are able to generate electricity relatively cheaply. Moreover, a larger share of electricity generation from hydro and nuclear sources decreases wholesale prices as it replaces generation from fossil fuels with higher marginal costs.

Table 1. EU Member States Included in the Dataset

Austria	Finland
Belgium	France
Czech Republic	Germany
Denmark	Greece
Estonia	Hungary
Ireland	Slovak Republic
Italy	Slovenia
Netherlands	Spain
Poland	Sweden
Portugal	United Kingdom

III. DATA

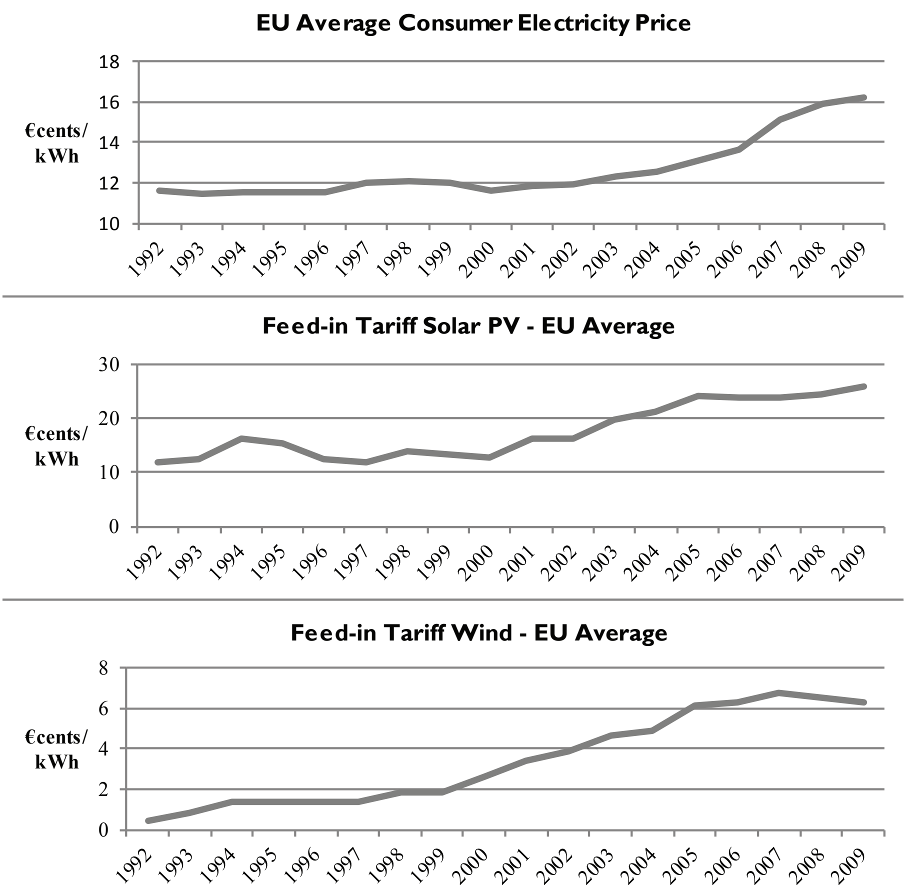
The dataset used in this study covers 20 European Union member states for the period of 1992 to 2009, and consists of data related to and representing the factors influencing electricity prices identified above.

Dependent Variable Selection

This study focuses on the effects of FIT policies on household consumer electricity prices for conceptual and technical reasons. Firstly, adverse price impacts at the household level are likely to have the most direct socioeconomic consequences as they directly affect welfare and inflation levels. As all EU member states are democracies, adverse price effects may potentially trigger backlash against FIT policies. Retail rates thus constitute the most politically relevant pricing unit for study. Although utility companies could mitigate the effect of FIT policies on retail consumers by cross-subsidizing rates for small-scale consumers at the expense of commercial consumers, such cross-subsidization is difficult to capture and typically works the opposite way in developed countries, with residential customers subsidizing businesses (IEA 2005).

Secondly, as electricity covered by the FIT is, in most cases, not actually traded on the spot market (because the payment amount is fixed), analyzing effects on spot prices would stop short of the full impact of FIT policies on electricity prices. Therefore, the dependent variable for this analysis is average yearly retail prices for electricity in €cents/kWh, for household consumers with an annual consumption of 3,500 kWh, including all taxes and levies. The data are obtained from EUROSTAT (2012) for EU member states for the time period 1992 to 2009.

Figure 2. Average Electricity Prices and FITs



Main Policy Variables

The main independent variables of interest are the tariff amounts for FIT in €cents/kWh in EU member states. I obtain data on tariff amounts for on-shore wind and solar PV from Groba, Indvik, and Jenner (2011). The dataset reports the mean value of the PV tariff across all size, location, and ownership categories, but fails to capture the complete extent of heterogeneity in FIT policies across countries. This shortcoming may result in a bias of the error term; however, as the unit of analysis is the country level, limitations of this kind have to be accepted to allow for feasibility.

Figure 2 shows that the average FIT has increased throughout the sample period in parallel with the average retail electricity price.

In an attempt to cover more of the existing policy heterogeneity, I also multiplied the tariff amount by the number of years generators receive the FIT under a country's respective policy regime and used the result as an alternative specification of the main policy variable. Data on contract duration was again obtained from Groba, Indvik, and Jenner (2011). As the exact amount of electricity generation under the FIT is not constant over the years due to changes in weather conditions and future payments were not discounted to the present, this measure fails to capture the exact annual payments made under the program, but nonetheless adds a critical dimension to the measure of overall payments provided by the respective policies.

Fossil Fuel Input Costs

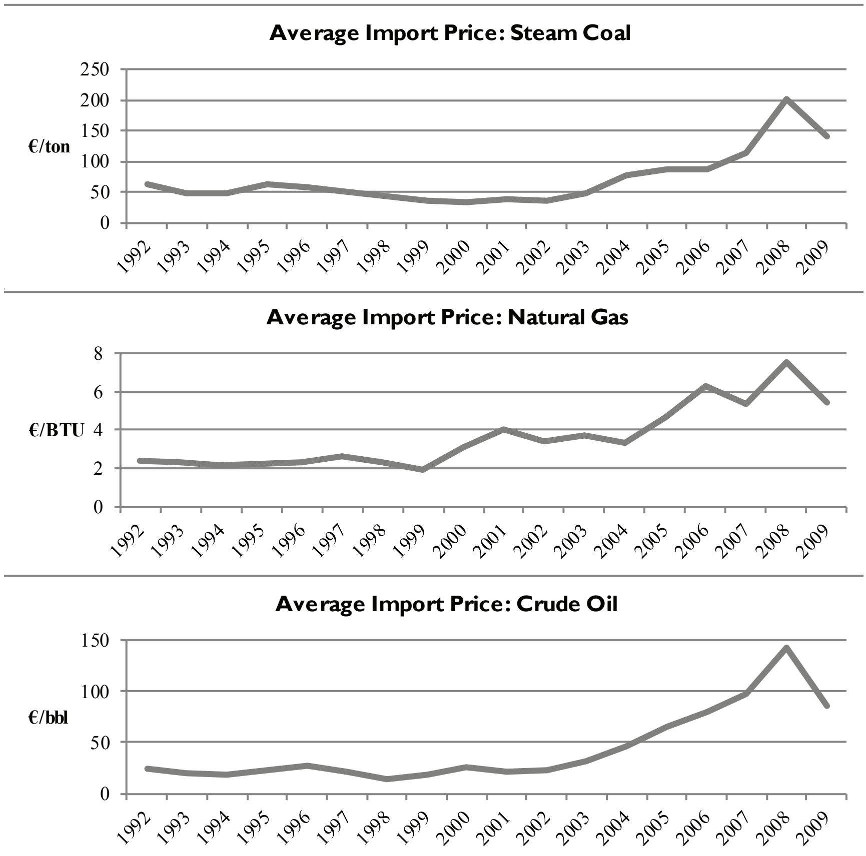
In order to approximate the variable costs for generating electricity in the economy, I include the costs of the fossil fuel inputs as the most important (variable) cost factor of generating electricity. To approximate these costs, I multiply the shares of electricity generation from coal, natural gas, and petroleum with those inputs' respective import prices, as most European countries import those fossil fuels. Both import prices and shares were obtained from the IEA's Electricity Information Statistics database. As IEA did not report import prices for each of the 20 member states and/or all 16 years in the panel, some values had to be imputed based on strong assumptions about member states' characteristics in terms of import costs.[2] The upward trend in household

[2] As coal and petroleum are traded on a global market, differences in import prices are largely driven by transportation costs, such as availability of shipping and rail capacity. Countries with similar geographic characteristics can therefore be assumed to have relatively similar import costs. Missing values for import prices were thus imputed based on the values reported for neighboring states with similar geographic characteristics. Analogous reasoning was applied for imputing natural gas import prices, which critically depend on availability of pipeline capacity.

For countries that did not report import prices for all panel years, I calculate an average "spread" multiplier between the existing observations and the European average as reported by IEA. Missing variables were imputed by multiplying this "transportation" multiplier with the European average.

These methods assume that the transport costs are driven by geographic factors and do not change over time. In cases where they do, this

Figure 3. Fossil Fuel Import Costs over Time



electricity prices is also matched by a continuous increase in fossil fuel import prices until 2008 (Figure 3), when the latter dropped sharply due to reduced demand in the wake of the global financial crisis.

Additional Explanatory Variables

In order to complete the energy mix, the shares of electricity generated from hydro and nuclear power are also included. The variable costs of these technologies are very low, so shares were not multiplied by any input prices although nuclear power relies on materials such as uranium or plutonium for nuclear fission. The shares of electricity generated from hydro and nuclear power are also obtained from the EIA electricity information statistics database.

I also include a capacity factor variable to capture the relative capital-intensity of electricity generation, using data from the IEA electricity information database. The capacity factor is the only variable in the model that attempts to capture the cross-country differences

may bias the error term. This is particularly problematic for countries that only reported values toward the end of the panel, as the methodology may not account for structural changes (such as rail or pipeline capacity) between the beginning and the end of the panel.

in the structural costs of generation, although it cannot capture the full capital costs embedded in retail rates.

I further include binary indicator variables for whether end-user prices were regulated by a government agency and for whether countries have enacted a requirement on utilities to generate a minimum share of the electricity from renewable sources. The information to construct these indicators is obtained from the country profiles of the IEA 2011 Electricity Information and ECME 2010 as well as from RES-Financing (2011) and OPTRES (2007), respectively.

IV. EMPIRICAL METHODS & MODEL SPECIFICATION

In order to capture some of the underlying capital costs of generating and supplying electricity in structurally different markets, controlling for country-level fixed effects (FE) is necessary. For example, a small country like the Netherlands, with relatively moderate weather and high population density, will have relatively low structural costs compared to large countries with extremely cold winters or extremely hot summers and low population density. These factors are unlikely to show much year-to-year variation.

While there is a relatively clear indication of a country fixed effect, accounting for time-specific changes in capital costs is more challenging. Although certain yearly effects related to input fuel prices are felt the same way across all countries, notably the costs for fossil-fuel inputs such as coal and petroleum, most time effects related to electricity prices occur within individual countries. These effects are largely related to a country's market structure and political system, for example in the competitiveness of wholesale or retail markets. Changes in these factors do not occur at the same time across countries. Whereas some countries like the UK saw multiple restructurings during the panel period, other markets have seen few changes in competitiveness (Cooke 2011). Omitting within-country structural changes in the costs of generating and supplying electricity (or any variable that can capture it) is likely to introduce autocorrelation into the error term of the model, as countries that have low embedded capital costs in year t are very likely to also have low embedded capital costs in year t–1. The Wooldridge test for first-order autocorrelation thus strongly rejected the null hypothesis of no first-order autocorrelation when omitting the embedded capital costs.

In order to capture the country- and year-specific underlying capital costs, eliminate the serial correlation, and address the clear upward trend in the dependent variable, I estimate a linear dynamic panel-data model that includes a lag of the dependent variable as a covariate as well as unobserved panel-level fixed effects. Bond (2002) emphasizes that even when coefficients on lagged dependent variables are not of direct interest, allowing for dynamics in the underlying process may be "crucial for recovering consistent estimates of other parameters." In my model, the previous year's retail price is understood to capture the unobserved variation in the embedded capital costs due to the strong relationship between retail prices and capital costs. As underlying capital costs are likely to only change slowly over time, the prior year's retail price offers a good approximation of the changes in embedded capital costs. All models soundly reject the null hypothesis that all coefficients except the time trend (lagged DV) are zero, tested through the chi-squared test reported by Arellano and Bond, and show that the rest of the model has explanatory power that goes considerably beyond the time trend.

Guided by this framework, I estimate the following base model:

$$RETAIL\ PRICE_{it} = \beta_1 RETAIL\ PRICE_{i,t-1} + \beta_2 FIT(wind)_{it} + \beta_3 FIT(solar)_{it} + \beta_4 Coal\ Input\ Cost_{it} + \beta_5 Natural\ Gas\ Input\ Cost_{it} + \beta_6 Petroleum\ Input\ Cost_{it} + \beta_7 Hydro_{share} + \beta_8 Hydro_{share} + \beta_9 Capacity\ factor + \beta_{10} RPS + \beta_{11} Regulator + \nu_i + \varepsilon_{it} \quad (1)$$

As the panel level effect ν_i is the same across time periods, it is by construction correlated with the lagged dependent variable because the dependent variable in year $t-1$ is also affected by ν_i, making the standard estimators inconsistent. In order to address this problem, I use the Arellano and Bond (1991) generalized method-of-moments (GMM) estimator, which was first proposed by Holtz-Eakin, Newey, and Rosen (1988). This method uses estimators constructed by first differencing to remove the panel-level effects and further lags of the dependent variable to create instruments of the lagged dependent variables and remove the autocorrelation. When the idiosyncratic errors εit are independently and identically distributed, the first differenced errors are first-order serially correlated in the Arellano-Bond specification.[3] However, assuming that ε_{it} is serially uncorrelated, the predetermined initial conditions imply that the lagged level $y_{i,t-2}$ will be uncorrelated with $\Delta\varepsilon_{it}$ and thus available as an instrument for the first differenced equation (Bond 2008). Serial correlation at order 1 thus does not invalidate the moment conditions used by the Arellano-Bond estimator, because only lags of two time periods and further are used as instruments. Apart from the lagged dependent variable, the first difference of all exogenous variables is used as standard instruments.

All models were tested for second-order autocorrelation with the Arellano-Bond post-estimation test for zero autocorrelation. The test is applied to the differenced residuals, and the null hypothesis is that there is no autocorrelation. As expected, the test for autocorrelation at order 1 in the first differences rejects the null hypothesis, but the test fails to reject

[3] As $\Delta\varepsilon_{it} = \varepsilon_{it} - \varepsilon_{i,t-1}$ and $\Delta\varepsilon_{i,t-1} = \varepsilon_{i,t-1} - \varepsilon_{i,t-2}$ both terms contain $\varepsilon_{i,t-1}$. Therefore the test for AR(1) is expected to be failed, and the test for AR(2) is decisive.

it at the second order, presenting no significant evidence of serial correlation in the first-differenced errors at order two or higher. The tests for autocorrelation thus present no evidence for model misspecification.

The Arellano-Bond method constructs the GMM estimator using as many lags of ε_{it} as are available in the panel. For long panels (panels with a large amount of time periods *t*) this potentially leads to over-identification. Over-identification in itself is generally desirable; however, there is potential danger of correlation between the over-identifying instruments and the residuals, which would invalidate the central assumption of the Arellano-Bond estimation that the instruments as a group need to be exogenous. In order to maintain instrument exogeneity, the number of lags used as instruments is restricted to 10. This represents the maximum number of lags that still allow the basic model specification to pass the Sargan test, which tests whether the residuals are uncorrelated with the set of constructed instruments. Subsequently, lags from two time periods back to 11 time periods are used to create the GMM type instruments described by Arellano and Bond (1991), in order to ensure instrument exogeneity.

The results are displayed in Table 2 next to the results of standard FE regression. The results for the Arellano-Bond and OLS estimations are relatively closely matched, though the coefficient estimate for the feed-in tariff for solar PV is significant under OLS but not

"The empirical analysis indicates there are clear adverse price effects associated with supporting electricity generation from wind through feed-in tariff schemes, although they are relatively small in magnitude."

under the Arellano-Bond estimate. This similarity can be explained by the fact that the Arellano-Bond estimator was designed for panels with many observations and only few time periods. However, the panel used in this analysis, although consisting of more *n* than *t*, was relatively evenly matched, covering 20 countries over 16 years. According to Rodman (2006), the correlation of the time trend with the error term will be less significant in panels with many time periods, as a shock to the country fixed effect that could affect the error term will decline over time.

V. EMPIRICAL RESULTS

Table 2 displays the empirical results of several alternative specifications of the main regression outlined in Equation 1.

Main Policy Variables

The empirical analysis indicates that there are clear adverse price effects associated with supporting electricity generation from wind through feed-in tariff schemes, although they are relatively small in magnitude. The results from model 1 show that an extra cent FIT for wind raises the retail electricity prices for residential consumers by approximately 0.06 cents. This corresponds to roughly 0.5

Table 2. Empirical Results

	Arellano-Bond Estimator				OLS with country FE			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
VARIABLES	I	I	I	I	I	I	I	I
lag-price	0.754***	0.742***	0.741***	0.729***	0.877***	0.874***	0.880***	0.874***
	(0.060)	(0.061)	(0.059)	(0.060)	(0.034)	(0.034)	(0.034)	(0.034)
tariff_wind	0.063**		-0.023		0.059**		-0.033	
	(0.031)		(0.053)		(0.024)		(0.039)	
tariff_pv	-0.006		0.002		-0.008**		-0.003	
	(0.006)		(0.010)		(0.004)		(0.008)	
totalpay_wind		0.005**		-0.003		0.004***		-0.003
		(0.002)		(0.005)		(0.002)		(0.004)
totalpay_pv		-0.000		0.000		-0.000**		-0.000
		(0.000)		(0.001)		(0.000)		(0.000)
tariff*regulator_wind			0.101**				0.112***	
			(0.051)				(0.037)	
tariff*regulator_pv			-0.010				-0.006	
			(0.008)				(0.006)	
totalpay*regulator_wind				0.008* (0.005)				0.009** (0.003)
totalpay*regulator_pv				-0.001				-0.000
				(0.000)				(0.000)
coal_input_cost	0.011	0.011	0.014*	0.014**	0.004	0.004	0.005	0.005
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
petro_input_cost	0.002	0.008	-0.019	-0.019	0.097*	0.102**	0.098**	0.105**
	(0.073)	(0.073)	(0.073)	(0.072)	(0.050)	(0.050)	(0.049)	(0.049)
gas_input_cost	0.349**	0.342**	0.417***	0.403***	0.162	0.150	0.197*	0.185*
	(0.148)	(0.147)	(0.145)	(0.145)	(0.106)	(0.106)	(0.106)	(0.107)
share_hydro	-2.623*	-2.604*	-2.709*	-2.780*	-1.875*	-1.811*	-1.627	-1.424
	(1.549)	(1.540)	(1.544)	(1.536)	(1.090)	(1.083)	(1.082)	(1.091)
share_nuclear	-7.313**	-7.160**	-6.641*	-6.617*	-3.975	-3.861	-3.031	-2.901
	(3.658)	(3.636)	(3.682)	(3.654)	(2.459)	(2.444)	(2.506)	(2.491)
capacityfactor	-0.400	-0.423*	-0.392	-0.386	-0.316	-0.282	-0.279	-0.244
	(0.246)	(0.246)	(0.247)	(0.248)	(0.195)	(0.194)	(0.193)	(0.193)
rps	-0.063	-0.022	-0.198	-0.133	0.024	0.044	-0.173	-0.168
	(0.309)	(0.308)	(0.314)	(0.312)	(0.230)	(0.231)	(0.247)	(0.248)
regulator	-0.697**	-0.694**	-0.974**	-1.016**	-0.368*	-0.379*	-0.841***	-0.893***
	(0.284)	(0.283)	(0.428)	(0.417)	(0.199)	(0.199)	(0.321)	(0.317)
Observations	287	287	287	287	320	320	320	320
Number of code	20	20	20	20	20	20	20	20
R-squared					0.880	0.881	0.884	0.883

Standard errors in parentheses.

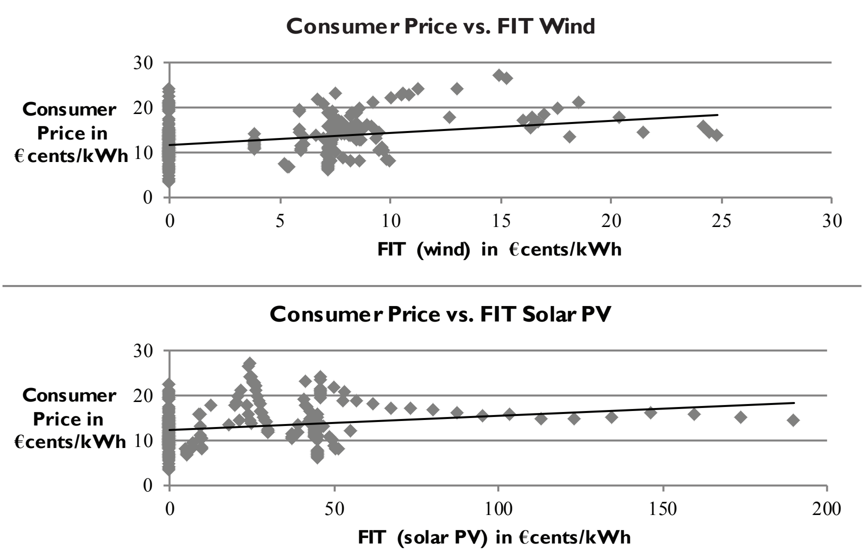
*** p<0.01, ** p<0.05, * p<0.10

percent of the average retail price for electricity throughout the panel. Given that the mean FIT tariff for wind is 3.6 cents, this implies that the presence of an FIT that pays exactly the mean tariff amount results in an electricity price that is 0.22 cents per kWh higher than in the absence of the FIT, approximately 2 percent of the average retail rate. For countries with successful FIT programs, such as Germany, that paid an average tariff of approximately 8 cents over the period of the panel, this corresponds to an increase in electricity prices of 0.48 cents per kWh, approximately 3 percent of the average retail price in Germany.

The coefficient of FIT for wind is highly statistically significant throughout all models and sensitivity tests, and the magnitude of the coefficient remains relatively unchanged. The coefficient decreases considerably when using the total payment amount over the contract duration, to approximately 0.04 percent of average retail prices in the panel. Increasing this value by 1 unit (either by increasing the tariff or by extending the contract period) increases retail electricity prices by 0.004 cents per kWh.

Interestingly, the coefficient for the feed-in tariff on solar PV is negative throughout all model specifications, although the magnitude of the coefficient is so small that it does not appear economically relevant. Statistical significance decreases when estimating the Arellano-Bond estimator, whereas the sign and magnitude of the effect remain relatively unchanged. However, the comparatively small magnitude of the effect may result from the fact that, currently, only a relatively small share of electricity is generated from solar PV, and many European countries have only put FIT policies in place relatively

Figure 4. Scatter Consumer Price vs. Tariff Amounts



recently. The mean of electricity generated from solar PV is 0.03 percent of total generation, although in market leaders Germany and Spain it is more than 0.17 percent. In contrast, the mean for wind generation is considerably higher at approximately 1.4 percent, with leaders Denmark and Spain at 10.3 and 3.9 percent, respectively. This result may change once countries develop greater solar generation capacity.

The negative coefficient may also be driven by the fact that all observations with tariff amounts greater than 60 cents are all from one country (Germany), which decreased its support for solar PV generation throughout the panel. The scatter graph provided in Figure 4 further illuminates this relationship. Nonetheless, my estimations indicate that, to date, supporting electricity generation from solar PV has had no effect on electricity retail prices.

Other Policy Variables of Interest

Higher shares of hydro and nuclear power in the electricity generation mix are associated with lower retail electricity prices; the coefficients for shares of electricity generated from both hydro and nuclear power were extremely large and negative across all models, with high to moderate levels of statistical significance.

> "The empirical results ... stand in stark contrast to the economic concerns regularly voiced by the opponents of such policies."

Retail price regulation is also associated with lower retail electricity prices. The coefficient on end-user price regulation is statistically significant and negative throughout all models. If end user prices are regulated by a government agency, they are approximately 0.7 cents per kWh lower than if they are not regulated.

The regulator variable is also of particular importance to the FIT debate. Models 3 and 4 show that in countries with retail price regulation, the effect of the FIT is highly statistically significant. In contrast, the effect of the FIT becomes insignificant for countries that do not regulate retail rates.

VI. DISCUSSION

The empirical results suggest that European feed-in tariff programs for electricity generated from wind and solar PV have had relatively little effect on retail electricity prices. This finding stands in stark contrast to the economic concerns regularly voiced by the opponents of such policies. In combination with the well-established success of FIT programs in spurring installation of RES-E generation capacity in countries such as Denmark, Germany, and Spain, the results presented in this paper support the view that well-designed FIT programs are not only the most successful, but also an economically viable policy option for supporting RES-E.

The results further suggest that there is a distinct difference in the price-effect of FIT legislations depending on the technology supported. The diverging findings for wind and solar PV are of great interest, as they point to the existence of the so-called “merit-order effect” as described by Ragawitz, Sensfuß, and Barbose (2008). While wind generates electricity mostly during off-peak periods at night, solar PV generates electricity during times when electricity demand is actually high, such as during clear, cold winter days and hot summer days. Therefore, solar PV can replace costly natural gas and petroleum plants, whereas wind electricity only replaces electricity generated from base-load coal, hydro, and nuclear plants with comparably low marginal costs, if these plants can even be shut down. Positive price-effects from replacing more costly technologies are therefore comparatively smaller for electricity generated from wind. However, the results for solar PV need to be treated with particular care given the small share of electricity produced by this technology.

The fact that the price increases associated with FITs for wind is only statistically significant when retail prices are still regulated suggests that regulators accommodate for increased costs incurred by utility companies by allowing them to charge higher retail rates. This result is particularly interesting in the light of recent developments in Spain, where the energy regulator CNE has failed to raise consumer prices appropriately for utilities to recover their costs. *The Economist* (2011) reports the resulting annual “electricity-tariff deficit” (the differential between utility companies’ costs and revenue) has risen dramatically to €5.6 billion ($8.3 billion). As a result, FIT payments have been cut retrospectively to alleviate utilities’ burdens, although the main cause for the “deficit” is likely rising raw-material prices. However, given that there is no significant effect of the FIT for wind if retail markets are liberalized (in fact, the coefficients throughout all models are negative), the empirical results indicate that efficiency gains from competition prevent retail rates to rise in the presence of FITs. This finding points toward a positive interaction between market liberalization at the retail level and RES-E support through FIT policies, which warrants some more focused exploration in the future.

“... well-designed FIT programs are not only the most successful, but also an economically viable policy option for supporting RES-E.”

The finding that the coefficient on retail regulation is negative stands in contrast to the literature on retail price deregulation, which suggests that market liberalization should lower prices rather than increase them. However, in the absence of functioning retail markets, utilities may be able to charge higher prices and thus extract rents from consumers. This is particularly true if retail electricity prices have previously been subsidized. Therefore, retail price regulation might shield consumers from higher prices.

“Even though rising retail rates are a serious concern for social welfare, this study suggests that in the last two decades such effects have largely been driven by factors other than FITs.”

The result that both hydro and nuclear power are associated with lower retail electricity prices across all model estimations also wields considerable explanatory power in explaining the price effects of FIT policies. With respect to its cost profile, electricity generation from wind and solar is very similar to hydro and nuclear. Most European nuclear and hydro plants are relatively old and thus have already paid off their enormous capital costs. Seeing that larger shares of hydro and nuclear generation are associated with a large decrease in retail rates in the long term, the same may hold true for wind and solar, once the initial investments necessary to expand capacity to significant levels are paid off. Given that the FIT programs were not established until the early 1990s and most countries have not even completed the first cycle of FIT contracts, it is unlikely that renewable capacity installed under any FIT regime has already recovered the initial capital investment. In this context, feed-in tariff schemes could play an instrumental part in getting this capacity installed and allowing for such low-cost generation in the future.

VII. CONCLUSION

Overall, the relatively modest price increases associated with FIT policies (for wind) found by this analysis should not overly concern European policy makers. Even though rising retail rates are a serious concern for social welfare, this study suggests that in the last two decades such effects have largely been driven by factors other than FITs. In light of the price-decreasing effects of larger shares of hydro and nuclear power in a country’s energy mix—and both wind’s and solar PV’s extremely similar cost profile—substantial investment into these technologies could actually result in lower electricity prices in the long run, as wind and solar generators replace more expensive natural gas and/or petroleum generation plants.

Considering the successful track record of FITs in increasing RES-E capacity in combination with these relatively low costs, feed-in tariff policies continue to be an extremely attractive policy tool for supporting RES-E. Nonetheless, policy is not an absolute and even successful policies need to be monitored and evaluated constantly. In particular, tariff rates need to be adjusted as technologies mature to accurately reflect development costs and ensure that the security guaranteed to investors under FIT schemes does not turn into excessive profits paid for by the consumer.

VIII. REFERENCES

Alagappan, L., R. Orans and C. Woo. 2011. What drives renewable energy development. *Energy Policy* 39: 5099-5104.

Arellano, M. and S. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies* 58, no. 2: 277-297.

BMU. 2011. Environment Minister Röttgen: Solar power support must be adjusted to market developments. Federal Ministry for the Environment, Nature Conservation, and Nuclear Safety, January 20. http://www.bmu.de/english/current_press_releases/pm/46964.php.

Bode, S. 2006. On the impact of renewable energy support schemes on power prices. HWWI Research Paper. Hamburg, Germany: Hamburgisches Welt-Wirtschafts Institut.

Bond, S. 2002. Dynamic panel data models: A guide to micro data methods and practice. *CEMMAP* Working Paper. London, UK: Institute for Fiscal Studies.

Butler, L. and K. Neuhoff. 2008. Comparison of feed-in tariff, quota and auction mechanism to support wind power development. *Renewable Energy* 33: 1856-1867.

Cooke, D. 2011. Empowering customer choice in electricity markets. IEA Information Paper. Paris, France: International Energy Association.

De Miera, G., P. del Rio Gonzalez, and I. Vizcaino. 2008. Analyzing the impact of renewable electricity support schemes on power prices: The case of wind electricity in Spain. *Energy Policy* 36: 3345-3359.

Del Rio, P. and M. Gual. 2007. An integrated assessment of the feed-in tariff system in Spain. *Energy Policy* 35: 994–1012.

Drukker, D. 2008. Econometric analysis of dynamic panel-data models using Stata. Stata Corp Presentation at the North American Stata Users Group Meeting, July 24 – 25, in College Station, TX.

EC. 2008. The support of electricity from renewable energy sources. Accompanying document to the Proposal for a Directive of the European Parliament and of the Council on the promotion of the use of energy from renewable sources. Brussels, Belgium: Commission of the European Communities.

EC Roadmap. 2008. Renewable Energy Road Map Renewable energies in the 21st Century: Building a more Sustainable Future. Communication from the Commission to the Council and the European Parliament. Brussels, Belgium: Commission of the European Communities.

ECME Consortium. 2010. The Functioning of Retail Consumer Markets for Consumers in the European Union. Brussels, Belgium: Report for the Directorate-General for Health and Consumers.

The Economist. 2011. Renewable energy: Sun-burned, July 26. http://www.economist.com/node/21524449.

EEG. 2009. Green-X. Deriving optimal promotion strategies for increasing the share of RES-E in a dynamic European electricity market. Vienna: Energy Economics Group of the Vienna University of Technology.

Eurostat. 2011. *Statistics Database*. Luxembourg: European Commission.

Frondel, M. et al. 2008. Economic impacts from the promotion of renewable energy technologies: The German Experience. *Energy Policy* 38: 4048-4056.

Fox-Penner, P. 1997. Electric utility restructuring—a guide to the competitive era. Vienna, VA: Public Utilities Reports Inc.

Gelabert, L., X. Labandeira and P. Lineares. 2011. Renewable energy and electricity prices in Spain. *Economics for Energy Working Paper*, Vigo, Spain: Energy for Economics.

Groba, F., J. Indivik, and S. Jenner. 2011. Assessing the strength and effectiveness of renewable electricity feed-in tariffs in European Union countries. DIW Discussion Papers. Berlin, Germany: German Institute for Economic Research.

Helm, D. and A. Powell. 1992. Pool prices, contracts and regulation in the British electricity supply industry. *Fiscal Studies* 13, no. 1: 89-105.

Holtz-Eakin, D., W. Newey, and H. Rosen. 1988. Estimating vector autoregressions with panel data. *Econometrica* 56, no. 6: 1371-1395.

International Energy Agency. 2011. *Electricity Information Database*. Paris, France: Organization for Economic Cooperation and Development (OECD).

International Energy Agency. 2005. Russian electricity reform—emerging challenges and opportunities. Paris, France: OECD.

Jamasb, T. and M. Pollitt. 2005. Electricity market reform in the European Union: Review of progress toward liberalization & integration. *The Energy Journal* 26, no. I: 11-42.

Jonsson, T., P. Pinson and H. Madson. 2010. On the market impact of wind energy forecasts. *Energy Economics* 32: 313-320.

Lesser, J. A. and X. Su. 2008. Design of an economically efficient feed-in tariff structure for renewable energy development. *Energy Policy* 36, no. 3: 981-990.

Lipp, J. 2007. Lessons for effective renewable electricity policy from Denmark, Germany and the United Kingdom. *Energy Policy* 35, no. 11: 5481-5495.

Mallet, Victor. 2010. Spain's solar power subsidy cuts attacked. *Financial Times*, December 22, http://www.ft.com/intl/cms/s/0/fdedc4d2-0dfb-11e0-86e9-00144feabdc0.html#axzz1qEmCx5Jm.

Newbery, D. 2002. Regulating unbundled network utilities. *The Economic and Social Review* 33, no. 1: 23-41.

Newberry, D. 1999. *Privatization, restructuring and regulation of network utilities–The Walras-Pareto lectures*. Cambridge, MA: MIT Press.

OPT RES. 2007. Assessment and optimization of renewable energy support schemes in the European electricity market. Report by order of the European Commission, DG Energy and Transport. Brussels, Belgium: European Commission.

Rathmann, M. 2007. Do support systems for RES-E reduce EU-ETS-driven electricity prices? *Energy Policy* 35: 342–349.

RES-Financing. 2011. Financing renewable energy in the European market. Report by order of the European Commission, DG Energy. Brussels, Belgium: European Commission.

REN21. 2010. Renewables global status report. 2010 update, technical report, Renewable Energy Policy Network for the 21st Century. Paris, France.

RES-Legal. 2011. Legal sources on renewable energy. Berlin: German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety.

Rodman, D. 2006. How to do xtabond2: An introduction to "difference" and "system" GMM in Stata. DGDev Working Paper. Washington, DC: Center for Global Development.

Sensfuß, F., M. Ragawitz, and M. Genoese. 2008. The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany." Energy Policy 36: 3086– 3094.

Stata Corp. 2007. Longditudinal/panel data, *Stata statistical Software: Release 10*. College Station, TX: Stata Corp LP.

Stoft, S. 2002. *Power systems economics—designing markets for electricity*. New York, NY: The Institute of Electrical and Electronics Engineers.

Traber, T. and C. Kemfort. 2009. Impacts of the German support for renewable energy on electricity prices, emissions, and firms. *The Energy Journal* 30, no. 3: 155-177.

Traber, T. and C. Kemfort. 2011. Gone with the wind?—Electricity market prices and incentives to invest in thermal power plants under increasing wind energy supply. *Energy Economics* 33: 249-256.

Weigt, H. 2009. Germany's wind energy: The potential for fossil capacity replacement and cost saving. *Applied Energy* 86: 1857–1863.

Wooldridge, J. 2009. *Introductory econometrics: A modern approach.* Mason, OH: Cengage Learning.

Can Incentives Increase Preference for Daughters?

Evidence from a Cash Transfer Scheme in Haryana, India

By Chand Tulal Mazumdar

Chand Tulal Mazumdar completed her Master of Public Policy from Georgetown University in 2012 and finished her thesis under the guidance of Thomas Wei, PhD. She is currently a Research Associate at the Abdul Latif Jameel Poverty Action Lab (J-PAL), where she works on a project that uses randomized evaluation to assess the impact of a performance-based incentive scheme on tuberculosis health workers in Northern India.

ABSTRACT

The gender imbalance caused by a skewed female-to-male sex ratio remains a persistent problem in India despite rapid economic development in recent times. The low child sex ratio results from both excess female infant mortality—due to malnutrition and neglect—and from sex-selective abortions, the roots of which lie in a strong cultural preference for sons over daughters. Although the government banned prenatal sex determination techniques in 1994, many feel that the policy has been ineffective due to inadequate enforcement. Apart from the ban, the government introduced other schemes and campaigns at both national and state levels that focused on changing parental perception and behavior toward daughters. Using two rounds of District Level Household Survey (DLHS) data, this study assesses the impact of Haryana's *Ladli* scheme—a conditional cash transfer scheme that provides incentive to parents for having a second daughter—on the likelihood of having daughters using a difference-in-differences approach with Punjab as a comparison state. The findings suggest that while the likelihood of having a daughter increased in Haryana compared with Punjab in the post-policy period, the effect is not statistically significant. However, restricting the sample to border districts in Haryana and Punjab shows some significant results.

I. INTRODUCTION

The human sex ratio is defined as the number of females to every 1000 males in a population. Low human sex ratios in the Indian population have been documented since the first Indian census in 1871(Kanitkar 1991). While earlier low ratios were attributed to pre-industrial factors such as low levels of education and healthcare provision that reinforced cultural and behavioral bias, persistently low ratios suggest that the problem may not disappear with economic development. Even though the 2011 Indian census showed an encouraging trend in the sex ratio, which increased seven points from the 2001 level to 940—the highest since 1971—it remains low compared to the estimated global average of 984. In addition, the bias remains stark in the child sex ratio—the number of females to every 1000 males in the 0-6 age group—which has declined to an all-time low of 914 (Government of India 2011).

According to demographers, both the natural sex ratio at birth and the population sex ratio are "remarkably consistent" across human populations in the absence of manipulation. The natural sex ratio at birth is 934–952 female births for every 1,000 male births. The slight excess of male births is balanced out in the population sex ratio as males have higher mortality rates than females. Thus, the population sex ratio is estimated to be 979–1,003 females for every 1,000 males (Hesketh & Xing 2006). In practice, both types of sex ratios vary widely across regions and countries (Table 1).

While revisiting the concept of "missing women" he introduced in 1990, Sen (2003) states that the total number of missing women has grown globally during the past decade, primarily due to an absolute growth in population. However, he adds that "another more important and radical change" has occurred during this period: while female disadvantage in mortality has been reduced drastically, this has been counterbalanced by natal disadvantage through prenatal sex detection and selective abortion (Sen 2003). In India, the low child sex ratio has resulted from both excess female infant mortality due to malnutrition and neglect, and sex-selective abortions, the roots of which lie in strong cultural preferences for

Table 1. Population Sex Ratio in 10 Most Populous Countries

Country	2001	2011
World	**986**	**984**
China	944	926
India	933	940
US	1,029	1,025
Indonesia	1,004	988
Brazil	1,025	1,042
Pakistan	938	943
Russia	1,140	1,167
Bangladesh	958	978
Japan	1,041	1,055
Nigeria	1,016	987

Source: Office of the Registrar General & Census Commissioner, Government of India (2011).

Table 2. Population and Child Sex Ratio in India, and Haryana and Punjab States				
	Total Population		Child Population (0-6)	
	2001	2011	2001	2011
India	933	940	927	914
Haryana	861	877	819	830
Punjab	876	893	798	846

Source: Office of the Registrar General & Census Commissioner, Government of India (2011).

sons (Jha et al. 2006; Arnold, Kishor, and Roy 2002; Jha et al. 2011). Similar observations have been documented in East Asian countries (Ebenstein 2007; Lin and Luoh 2008; Chunn and Das Gupta 2009), while male-biased sex ratios have been found among children of Asian immigrants in the US, Canada and, Norway (Almond and Edlund 2008; Almond, Edlund, and Milligan 2009; Singh et al. 2010).

Assumptions that discrimination against girls would diminish with economic development and female education have proven simplistic (Löfstedt, Shusheng, and Johansson 2004). For example, in South Korea sex ratios kept declining until a few years ago despite rapid development in industrialization, education, and urbanization, including women's participation in the formal labor force. Even though South Korea was included in the OECD countries by the mid-1990s, gender imbalance rose sharply during this period (Chung and Das Gupta 2007). This pattern is also evident in the Indian context, where the decline has continued despite rising living standards and higher levels of human development. In fact, the gender imbalance is more pronounced in wealthier states like Punjab and Haryana, and in urban areas where people have better access to prenatal tests to determine fetus sex (Table 2) (Haub and Sharma 2006; Subramanian and Selvaraj 2009).

Apart from widely documented non-economic factors, there are several hypotheses that sex selection occurs for economic reasons and is based on parents' intertemporal allocation decisions to optimize the family utility function. According to economic models of choice, parents tend to invest in the child with greatest potential returns, and this rationale can be extended to the unequal sex ratio in India. In developing countries, the gender gap in returns is due to both labor market forces and cultural practices where parents have to pay a dowry for their daughter's marriage. These daughters then move out of the family, while sons stay within the household with their wives. Thus,

"... while female disadvantage in mortality has been reduced drastically, this has been counterbalanced by natal disadvantage through prenatal sex detection and selective abortion."

“The impact of the falling ratio is important as it not only contributes to the deteriorating status of women in society, but also adds to increasing crime and violence.”

parents are more likely to receive the full return of investing in sons than in daughters under resource constraints (Rosenzweig and Schultz 1982; Strauss and Thomas 1995).

Studies have found that the sex outcome of the first pregnancy in India is close to the natural rate as parents do not selectively abort their first pregnancy as they feel they can try again to have a son. However, as family size in India has fallen substantially, it seems that selective abortion of girls is increasingly being used for second- or higher-order births to ensure at least one boy in the household, given a firstborn girl. On the other hand, it has been found that there is no significant decline in second-order birth sex ratio if the first-born was a boy (Jha et al. 2011; Ebenstein 2007).

The impact of the falling ratio is important as it not only contributes to the deteriorating status of women in society, but also adds to increasing crime and violence (Edlund et al. 2007; Hudson and Boer 2002), affects psychological wellbeing (Zhou et al. 2011), and creates long-run socio-demographic imbalances. The reduced number of women may push women into traditional family roles at the expense of education, training, and employment (Guilmato 2007). This deficit is already being felt in Punjab and Haryana, where young men have difficulties in finding brides and are increasingly resorting to unusual solutions, such as marrying across other caste groups, importing brides from other regions, or trafficking (Jagran Post Bureau 2011).

Though India legalized abortion in 1971, ultrasound technology did not become widely available until the mid-1980s, after which sex ratios at birth began to fall significantly below expected norms. Recognizing this trend, the Indian government passed the Pre-Conception and Pre-Natal Diagnostic Techniques (PNDT) (Prohibition of Sex Selection) Act in 1994, outlawing prenatal sex determination on January 1, 1996. However, the sex ratio has continued to decline, leading many to believe that the ban has been practically ineffective due to inadequate enforcement and insufficient punitive measures (Guilmato 2007). Even though stricter measures are being taken to enforce the ban nationwide, the Indian Planning Commission recently acknowledged that the government has failed to implement the ban. The Commission is now looking at alternate policy options, including giving incentives to families and health workers for the safe delivery of babies and adoption of female fetuses (Dhar 2011).

In addition to the national legislation, several schemes and campaigns exist at the state level to try to change parental perception and behavior toward daughters. The first such scheme was launched by Tamil Nadu in 1992, and similar schemes were implemented

Table 3. Salient Features of Haryana's *Ladli* Scheme

Objectives	To combat female feticide, increase number of girls in families, improve sex ratio, and raise status of the girl child in society.
Eligibility	All state residents are eligible if they have a second girl child born on or after August 20, 2005. Parents should ensure proper immunization and enroll both sisters in school. Parents receiving benefits from any other schemes are also eligible.
Incentives	Government will invest a cash incentive of Indian Rupees (INR) 5,000 (~US$100) per year for a period of 5 years or until the scheme is extended, in designated investment bonds (*Kisan Vikas Patra*), with an interest rate of approximately 8.29 percent, in the name of the second girl child and the mother. The accumulated amount will be given when the girl child turns 18. The incentive will expire if either girl gets married before she reaches the age of 18.
Beneficiaries	86,820 beneficiaries as of December 2009 since implementation in August 2005 (The Hindu 2010).

by Haryana in 2005, Madhya Pradesh in 2006, and Delhi in 2008. These schemes have tailored eligibility criteria and incentives to achieve a variety of development objectives, including improvement in educational attainment and health indicators, so that parents perceive greater benefits from having daughters, thereby reducing female feticide and improving the child sex ratio.

While there have been several empirical studies on the child sex ratio—its impacts, causes, and consequences—there are relatively few studies to assess the impact of legislation or policy interventions on improving the ratio. A recent study assesses the impact of the legal ban (PNDT Act) on sex-selective abortions (Nandi and Deolalikar 2011), and finds that contrary to general perception of the law being virtually ineffective, the act had a significantly positive impact on the child sex ratio. The study's authors estimate that in the absence of the Act, the gender imbalance would have further increased by 13-20 percentage points, or an additional 51,000 female fetuses would have been aborted. Another study measures the impact of a financial incentive program—*Apni Beti Apna Dhan* cash transfer scheme implemented in Haryana in 1994—on the child sex ratio and finds that it had a positive impact on both the number of daughters born and parental investment in daughters' health and education (Sinha and Yoong 2009).

This study contributes to this limited literature by estimating the impact of Haryana's *Ladli* scheme (Table 3), a conditional cash transfer (CCT) program, on the likelihood of mothers having daughters by using data from two rounds of the India District Level Household Survey (DLHS). This study also adds to the broader literature of evaluating CCT schemes in India, which are being advocated within multiple sectors and at multiple levels for alleviating poverty and achieving the United Nations Millennium Development Goals.[1]

II. DATA

This study uses figures from two rounds of the DLHS data, which are primarily designed to collect information on reproductive and child health at the district level in India. These surveys are conducted by the International Institute for Population Sciences (IIPS) of Mumbai with funding from India's Ministry of Health and Family Welfare[2] and consist of health interviews covering family planning, maternal and child health, reproductive health of ever-married women, and use of maternal and child healthcare services.

The third round of DLHS (DLHS-3) was conducted in 611 districts from late 2007 to late 2008 and sampled 720,320 households (1,000, 1,200, or 1,500 from each district) using multistage stratified sampling with probability proportional to size using the 2001 census data. From these households, 643,944 currently married and ever-married women, aged 15-49 years, were interviewed. DLHS-3 covered the period two to three years after Haryana's *Ladli* scheme was implemented.

The second round of DLHS (DLHS-2) was conducted between 2002 and 2004 in 593 districts and sampled 620,107 households (about 1,000 in each district) using the random sampling method outlined for DLHS-3. From these households, 507,622 currently married women, aged 15-44 years, were interviewed. DLHS-2 covered the period before the Haryana *Ladli* scheme was implemented in August 2005. While all ever-married women were interviewed for DLHS-3, only currently married women were interviewed for DLHS-2. Thus, for the purpose of this study, I use data for only currently married women.

Household interviews were conducted to gather information about the women's age, educational attainment, birth history, birth order, fertility preference, and child sex preference. The outcome of the most recent pregnancy (live birth, stillbirth, or spontaneous or induced abortion) and the survival of the child in the case of a live birth were also recorded. All rounds of DLHS data include a separate household interview that gathered information about demographic composition of households and socioeconomic characteristics, including asset ownership.

III. EMPIRICAL STRATEGY

To estimate the effects of the *Ladli* scheme on the likelihood of having a second daughter in Haryana, I pool the second and third rounds of DLHS

[1] Based on recommendations by the subgroup for girl child development created during the 11th five-year plan, the Government of India has implemented a pilot CCT scheme for families with girl children, Dhanalakshmi, on a pilot basis in 11 blocks in 7 states since 2008-09.
[2] DLHS-3 was also funded by United Nations Population Fund (UNFPA) and United Nations Children's Fund (UNICEF).

data with Haryana as the treatment group. I use Punjab as the control group because it is economically, demographically, and geographically similar to Haryana. Thus, the strategy for estimating the impact of the program on the likelihood of having a second daughter is difference-in-differences (DID):

$$DID = (\overline{Haryana}^{post} - \overline{Punjab}^{post}) - (\overline{Haryana}^{pre} - \overline{Punjab}^{pre})$$

where *post* and *pre* index individuals in the post-policy and pre-policy periods, and *Haryana* and *Punjab* indicate the average likelihood of having a second daughter for Haryana and Punjab, respectively. Thus, my identifying assumption is that the change in the likelihood of having a second daughter in Haryana between the pre- and post-policy periods would have been the same as the change in Punjab during the same period, if the policy had not been implemented.

The DID approach can be implemented with linear regressions to allow for controlling for other variables:

$$girl_i = \beta_0 + \beta_1 * Haryana_i + \beta_2 * Post2005_i + \beta_3 * (Haryana * Post2005)_i + X_i`\gamma + \varepsilon_i$$

where *i* indexes the individual eligible woman, and ε is the error term. *Girl* is a dummy variable indicating if the eligible woman had a daughter during a certain time period (described below). *Post2005* indicates post-policy period, and *X* represents individual-specific covariates such as age, education, husband's education, residence type, and standard of living. *Haryana* is a dummy indicating if the eligible woman lives in the treatment state (Haryana) or the control state (Punjab). β_1 and β_2 estimate the differences in the likelihood of having a daughter between the treatment-control groups and pre- and post-policy period groups, respectively. The coefficient of the interaction term, *Haryana***Post*2005, is the DID estimator (β_3), which indicates whether the change in likelihood of having a daughter between the pre- and post-policy periods is higher in the treatment group as compared to the control group. Thus, a positive estimate for β_3 indicates a relative increase in the likelihood of having a second daughter—suggesting that the *Ladli* program was effective—while a negative estimate for β_3 indicates otherwise. Since the dependent variable is a dummy variable, I run the linear probability model (LPM) to estimate the coefficients.

The sample of eligible women for the study consists of currently married women aged 15-44 years. I drop women older than 44 years in DLHS-3 (2007-08) for consistency, as DLHS-2 (2002-04) only interviews women up to 44 years. Then, I divide the data into two sub-samples (Full and Restricted). The Full sample includes all eligible women who had either one daughter or no daughter before 2002 for the pre-policy dataset (cut-off period: 2001) and before 2006 for the post-policy dataset (cut-off period: 2005). The Restricted sample contains only those eligible women who had exactly one daughter before the cut-off periods mentioned above. Thus, the outcome of interest (*girl)* is a dummy variable indicating whether a woman in the sample had at least one daughter after the cut-off period. The reason I include first daughters as well is because even though there is no incentive for having the first daughter, some women with no daughters may aim to have two daughters eventually to get the incentive. But with only three years of data since 2005 (which is the year when the program was enacted), they may just have had one of the two daughters by 2008, so having the first daughter still might indicate that the incentive is working.

I also run the regression specifications separately for the whole of Haryana and Punjab states and for the border districts alone, as the latter may be more comparable. Based on the latest state borders, I consider seven districts for Haryana: Sirsa, Fatehbad, Jind, Kaithal, Kurukshetra, Ambala, and Panchkula, and six districts for Punjab: Muktsar, Bhatinda, Mansa, Sangrur, Patiala, and Mohali.

IV. DESCRIPTIVE FINDINGS

Full States

Table 4 provides the summary statistics: there are 13,736 observations for Haryana and 11,784 observations for Punjab for the Full sample in the pre-policy period (2004), while there are 14,173 observations for Haryana and 14,067 observations for Punjab in the post-policy period (2008). The Restricted sample includes 6,230 women for Haryana and 5,746 women for Punjab in the pre-policy period, and 6,079 women for Haryana and 6,338 for Punjab in the post-policy period. While the mean of most of the variables of interest seem to have relatively similar magnitudes,

Figure 1. Summary Statistics of Dependent Variable (Full States)

Figure 1A

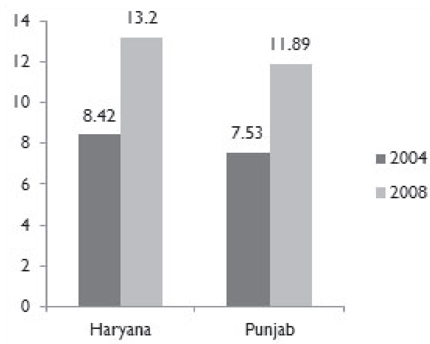
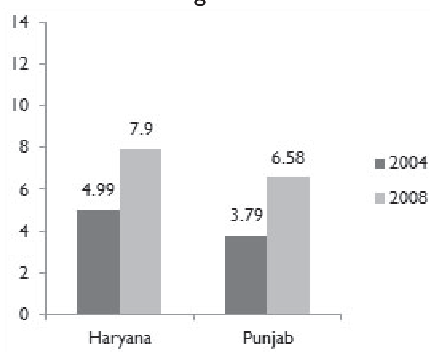


Figure 1B



Note: Figure 1A represents the Full sample that includes all women who had zero or one daughter before 2002 and 2006. Figure 1B represents the Restricted sample that includes all women who had exactly one daughter before 2002 and 2006. Each bar gives the percentage of eligible women in that particular year that had at least one daughter after the cut off periods in the pre- and post-policy periods (2001 for pre-policy and 2005 for post-policy period).

Table 4. Sample Summary Statistics (Full States)

	2004				2008			
	Full Sample		Restricted Sample		Full Sample		Restricted Sample	
	Haryana	Punjab	Haryana	Punjab	Haryana	Punjab	Haryana	Punjab
Observations	13,736	11,784	6,230	5,746	14,173	14,067	6,079	6,338
Age	28.085	30.132	30.823	32.465	28.717	30.31	31.944	33.065
Education	5.66	6.385	5.058	6.022	6.037	7.054	5.289	6.439
Husband's Education	8.333	7.607	7.963	7.456	8.491	8.11	8.046	7.737
Urban	0.302	0.316*	0.321	0.326^	0.253	0.29	0.265	0.294
Standard of living index								
- Middle	0.433	0.368	0.419	0.363	0.489	0.342	0.495	0.356
- Richest	0.432	0.556	0.437	0.562	0.451	0.644	0.442	0.631
Total sons	1.248	1.233^	1.383	1.319	1.204	1.141	1.378	1.247
Total children	1.789	1.799^	2.435	2.358	1.771	1.716	2.461	2.315

Notes: Full sample includes all women who had zero or one daughter before 2002 and 2006. The Restricted sample includes all women who had exactly one daughter before 2002 and 2006. Age is measured in number of years; education and husband's education are measured in number of years of schooling; region of residence is urban or rural; standard of living index is richest, middle, or poorest; total number of children, sons and daughters include total number of surviving children, sons, and daughters. T-tests of differences in mean between Haryana and Punjab were calculated for each variable in pre- (2004) and post- (2008) policy periods for both Full and Restricted samples. * = significant at 10% level (p<0.05). ** = significant at 5% level (p<0.01). ^ = not significant. All t-test scores are significant at the 1% level unless indicated otherwise.

formal testing of means shows that the differences in means between the states were statistically significant for all variables, except for *urban* and *total children* for the Full sample in 2004. I control for these variables in the regressions because of these differences.

Figures 1A and 1B show the percentage of women who had at least one daughter after the cut-off in pre- and post-policy periods, for the Full and Restricted samples of the whole states. For both Punjab and Haryana, this percentage of women was higher in the post-policy period relative to the pre-policy period. Additionally, the unconditional DID estimates, at 0.42 percentage points for the Full sample and 0.12 percentage points for the Restricted sample, show that the increase in Haryana (Full sample: from 8.42 to 13.2 percent; Restricted sample: from 4.99 to 7.9 percent) was more than that in Punjab (Full sample: from 7.53 to 11.89 percent; Restricted sample: from 3.79 to 6.58). However, none of the estimates are statistically significant as will be seen from the regression results in the next section.

Border Districts

Table 5 provides the summary statistics for the border districts: for the Full sample in pre-policy period (2004), there are 5,039 observations for Haryana and 4,574 observations for Punjab, while there are 4,838 observations for Haryana and 4,511 observations for Punjab in the post-policy period (2008). The Restricted sample includes 2,321 women for Haryana and 2,137 women for Punjab

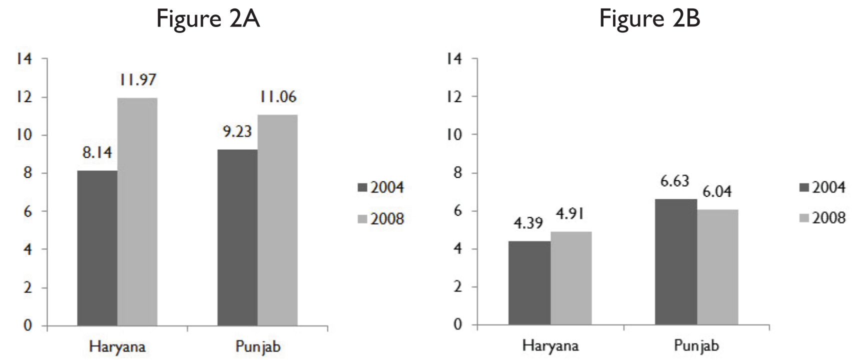
Table 5. Sample Summary Statistics (Border Districts)								
	2004				2008			
	Full Sample		Restricted Sample		Full Sample		Restricted Sample	
	Haryana	Punjab	Haryana	Punjab	Haryana	Punjab	Haryana	Punjab
Observations	5,039	4,574	2,321	2,137	4,838	4,511	2,083	1,969
Age	28.706	29.608	31.192	32.084	29.435	29.900**	32.62	32.756^
Education	5.137	5.632	4.598	5.302	5.869	6.461	5.176	5.837
Husband's Education	7.54	7.084	7.081	7.001^	7.98	7.794^	7.554	7.365^
Urban	0.275	0.303**	0.296	0.313^	0.24	0.291	0.254	0.293**
Standard of living index								
- Middle	0.421	0.379	0.42	0.371	0.482	0.372	0.479	0.394
- Richest	0.45	0.534	0.444	0.549	0.473	0.611	0.472	0.59
Total sons	1.246	1.211^	1.358	1.297*	1.175	1.113	1.31	1.224
Total children	1.791	1.774^	2.404	2.347*	1.73	1.665**	2.379	2.287

Notes: Full sample includes all women who had zero or one daughter before 2002 and 2006. The Restricted sample includes all women who had exactly one daughter before 2002 and 2006. Age is measured in number of years; education and husband's education is measured in number of years of schooling; region of residence is urban or rural; standard of living index is richest, middle, or poorest; total number of children, sons, and daughters include total number of surviving children, sons and daughters. T-tests of differences in mean between Haryana and Punjab were calculated for each variable in pre- (2004) and post- (2008) policy periods for both the Full and Restricted sample. * = significant at 10% level ($p<0.05$). ** = significant at 5% level ($p<0.01$). ^ = not significant. All t-test scores are significant at the 1% level unless indicated otherwise.

in the pre-policy period, and 2,083 women for Haryana and 1,969 for Punjab in the post-policy period. Similar to the full states, formal testing of means shows that the differences in means between the border districts are statistically significant for several variables. However, as expected, the districts appear to be more similar between the two states in terms of women's *age* for the Restricted sample in 2008, *husband's education* in almost all years and samples, *urban* areas in the Restricted samples of the pre-policy period, and *total boys* and *total children* in Full sample of the pre-policy period. Again, I control for these variables in the regressions because of these differences.

Figures 2A and 2B show the percentage of women who had at least one daughter after the cut-off in pre- and post-policy periods for the Full and Restricted samples of the border districts, respectively. In both Punjab and Haryana, this percentage of women went up in the post-policy period relative to the pre-policy period. Additionally, the unconditional DID estimates, at 2 percentage points for the Full sample and 1.11 percentage points for the Restricted sample, show that the increase in Haryana (Full sample: from 8.14 to 11.97 percent; Restricted sample: from 4.39 to 4.91 percent) was greater when compared to Punjab (Full sample: from 9.23 to 11.06 percent; Restricted sample: from 6.63 to 6.04 percent). The estimate for the Full sample is statistically significant in this case, as we will see from the regression results in the next section.

Figure 2. Summary Statistics of Dependent Variable (Border Districts)



Notes: Figure 2A represents the Full sample that includes all women who had zero or one daughter before 2002 and 2006. Figure 2B represents the Restricted sample that includes all women who had exactly one daughter before 2002 and 2006. Each bar gives the proportion of eligible women in that particular year that had at least one daughter after the cut-off periods in the pre- and post-policy periods (2001 for pre-policy and 2005 for post-policy period).

V. REGRESSION RESULTS

Table 6 presents the regression results of the key outcome of interest (*girl*) if women from the Full and Restricted samples had at least one daughter after the cut-off period for the full states and the border districts alone. While all the coefficients have been detailed in the tables, the parameter of interest in each equation is the interaction term, *Haryana*Post*2005. This parameter measures the increase in likelihood of having at least one daughter or a second daughter for Haryana residents after the policy was implemented compared to Punjab. The LPM estimates are presented with the full set of controls including age, education, husband's education, standard of living index (richest, middle, and poorest), and region of residence (urban and rural).

Full States

For the full states, the LPM DID estimate of the Full sample shows that after the implementation of the policy, the likelihood of having at least a daughter in Haryana increased by 0.8 percentage points more between the pre- and post-policy periods compared with the change during the same period in Punjab; however, the estimate is not statistically significant. For the Restricted sample, the estimate shows that the likelihood of having a second daughter increased by 0.5 percentage points more between the pre- and post-policy periods compared with the change in Punjab during the same period, but the effect is not statistically significant.

Border Districts

In comparison, the LPM DID estimate for the border districts shows that for the Full sample, the likelihood of having at least one daughter increased by almost 2.3 percentage points more between the pre- and post-policy periods in Haryana compared with the change during the same period in Punjab, and the estimate is statistically

Table 6. LPM DID Estimates (Full States & Border Districts)				
Dependent Variable: Women had at least one daughter after the cut-off (2001 and 2005)				
	Full States		Border Districts	
	Full Sample	Restricted Sample	Full Sample	Restricted Sample
Haryana	-0.013***	-0.007	-0.021***	-0.015*
	(0.003)	(0.004)	(0.006)	(0.006)
Post2005	0.047***	0.037***	0.022***	0.020**
	(0.004)	(0.004)	(0.006)	(0.007)
Haryana Post2005* (DID coefficient)	0.008	0.005	0.023**	0.018
	(0.005)	(0.006)	(0.008)	(0.010)
Type of residence:				
- *Urban*	0.009**	0.015***	0.016**	0.014*
	(0.003)	(0.004)	(0.005)	(0.006)
Standard of living index:				
- *Middle*	0.002	-0.017*	0.014	-0.005
	(0.006)	(0.007)	(0.009)	(0.012)
- *Richest*	-0.012*	-0.035***	-0.008	-0.022
	(0.006)	(0.008)	(0.010)	(0.012)
R-squared	0.069	0.068	0.07	0.064

Notes: Full sample includes all women who had zero or one daughter before 2002 and 2006. The Restricted sample includes all women who had exactly one daughter before 2002 and 2006. For Full States: n = 53,760 for Full sample and n = 24,393 for Restricted sample. For Border Districts: n = 18,962 for Full sample and n = 8,510 for Restricted sample. For all specifications, the dependent variable is a dummy variable, girl, which takes the value of 1 if the woman in the sample had at least one daughter after the cut-off (2001 for pre-policy and 2005 for post-policy period) and 0 if not. The right-hand-side variables are a dummy for Haryana, post-policy period (2008), a Haryana post-interaction term, and covariates (age, education, husband's education, standard of living index, and region of residence). Age is measured in number of years; education and husband's education are measured in number of years of schooling; region of residence is urban or rural; standard of living index is richest, middle, or poorest. * = significant at 10% level ($p<0.05$). ** = significant at 5% level ($p<0.01$). *** = significant at 1% level ($p<0.001$).

significant at the 5 percent level. For the Restricted sample, the estimate shows that the likelihood of having a second daughter increased by 1.8 percent more between the pre- and post-policy periods in Haryana compared with the change in Punjab during the same period, but is not statistically significant.

Heterogeneity

Table 7 shows the DID estimates for various subgroups—urban, rural, richest, middle and poorest, both for the full states and border districts, and also for the Full and Restricted samples. For the full states, the results did not produce any statistically significant insights for most categories from the Full sample. However, the border districts show a significant increase in the likelihood of having a daughter in both the rural category and the richest category by more than two percentage points each.

VI. ROBUSTNESS CHECKS

Since the estimates for the full states do not produce any significant results,

Table 7. LPM DID Estimates, by category (Full States & Border Districts)

Dependent Variable: Women who had at least one daughter after the cut-off (2001 and 2005)

	Full States		Border Districts	
	Full Sample	Restricted Sample	Full Sample	Restricted Sample
Urban	0.012	-0.005	0.029	0.02
	(0.009)	(0.010)	(0.016)	(0.017)
Rural	0.006	0.010	0.021*	0.016
	(0.006)	(0.007)	(0.010)	(0.012)
Richest	0.005	-0.003	0.024*	0.016
	(0.007)	(0.007)	(0.011)	(0.011)
Middle	-0.005	-0.011	0.002	-0.009
	(0.008)	(0.010)	(0.014)	(0.017)
Poorest	0.028	0.078	0.047	0.114
	(0.032)	(0.043)	(0.051)	(0.066)

Notes: The table shows the DID estimates, or the interaction terms, from each of the categories–urban, rural, richest, middle, and poorest. The Full sample includes all women who had zero or one daughter before 2002 and 2006. For all specifications, the dependent variable is a dummy variable, girl, which takes the value of 1 if the woman in the sample had at least one daughter after the cut-off (2001 for pre-policy and 2005 for post-policy period) and 0 if not. The right-hand-side variables are a dummy for Haryana, post-policy period (2008), a Haryana post-interaction term, and covariates (age, education, husband's education, standard of living index, and region of residence). Age is measured in number of years; education and husband's education are measured in number of years of schooling; region of residence is urban or rural; standard of living index is richest, middle, or poorest. * = significant at 10% level (p<0.05). ** = significant at 5% level (p<0.01). *** = significant at 1% level (p<0.001).

I run robustness checks only for the border districts. To show that the parallel assumption of the DID method holds, I use the first round of the DLHS survey conducted in 1998-99 to ascertain that the trends in the dependent variables for the treatment and control states were not divergent prior to the baseline. For that purpose, I simulate the same empirical strategy for DLHS-1 and DLHS-2, and divide the data into two subsamples. The Full sample was divided into a pre-policy dataset that includes all eligible women who had either one daughter or no daughter before 1997 (cut-off period: 1996) and a post-policy dataset that includes all eligible women who had either one or no daughter before 2002 (cut-off period: 2001). The Restricted sample contains only those eligible women who had exactly one daughter before the cut-off periods as mentioned. The measures of the dependent variable, *girl,* remain the same. The DLHS-1 data is comparable to the DLHS-2 and DLHS-3 waves, with the only difference being that the standard of living index was not available for DLHS-1. As such, the type of house—*kachha* (mud house), *semipucca* (mix of mud, brick and cement), and *pucca* (brick and cement)—was used as the proxy variable.

The regression results in Table 8 show there is no significant effect in the likelihood of having a daughter in Haryana compared with the change in Punjab before the baseline period.

Apart from validating the parallel assumption trend to establish the robustness of my identification strategy,

"... the improved sex ratio may not be attributable to the *Ladli* scheme. However, there are some positive results for the border parts of the state."

one of the major factors to account for is that Punjab also implemented a program called *Balri Rakhshak Yojana* in 2005-06 to improve the sex ratio in the state. However, its eligibility criteria were much more stringent than Haryana's program, and as of 2008-09 it had only 212 beneficiaries since its launch (Government of Punjab 2009). In comparison, the Haryana *Ladli* scheme had nearly 50,000 beneficiaries in 2007-08 alone, during which it spent 119 percent of the amount budgeted for the fiscal year (Government of Haryana 2007-08), and had a total of 86,820 beneficiaries as of December 2009 (The Hindu 2010).

VII. DISCUSSION

Results from the 2011 census brought both good and bad news for Haryana—the state reported the best sex ratio figures in the last 110 years, yet it continues to rank the lowest among all 28 states (IANS 2011). According to the Haryana government, the *Ladli* scheme has turned around the state's sex ratio (Mahajan 2011). In 2010, the state government also decided to continue the available benefits of the scheme for another five years (The Hindu 2010). The results from this study suggest that the improved sex ratio may not be attributable to the *Ladli* scheme. However, there are some positive results for the border parts of the state. This is encouraging, as the identification assumption used for the study may be more likely to hold true for the border districts. These findings also raise the possibility that the *Ladli* scheme may not have affected all parts of Haryana uniformly.

It would be more plausible to attribute a positive impact to the policy if the Restricted sample—which constitutes the primary target group of the policy—had yielded significant results. Significant results for the Full sample that included women with both zero and one daughter at baseline, but not for the Restricted sample that included women with exactly one daughter at baseline, may imply that the number of first-born daughters (for women with zero daughters) went up significantly in the border districts of Haryana compared to Punjab. While this is an encouraging outcome because it indicates an increase in the overall likelihood of having a daughter, the data does not clearly reflect the impact of the policy as it is currently designed. However since the DLHS-3 data only covers a very brief period after the policy was implemented, it is possible that although there is no direct incentive to have the first daughter, some women with no daughters may aim to eventually have two daughters to get the incentive. Thus, data for women who had only one of the two daughters by 2008 might still indicate that the incentive is working. Nonetheless, empirical evidence suggests that parents do not usually abort their first daughter, but the incidence of such sex-selective abortion increases with two or more pregnancies, as noted in the

Table 8. LPM DID Estimates (Border Districts)

Dependent Variable: Women who had at least one daughter after the cut-off (1996 and 2001)		
	Full Sample	Restricted Sample
Haryana	-0.027***	-0.016
	(0.008)	(0.009)
Post2000	-0.078***	-0.049***
	(0.009)	(0.011)
*Haryana * Post2000* (DID coefficient)	0.008	0.003
	(0.010)	(0.011)
Type of residence:		
- *Urban*	0.004	0.024**
	(0.006)	(0.007)
House type:		
- *Semipucca*	0.013	0.010
	(0.007)	(0.009)
- *Pucca*	0.004	-0.004
	(0.008)	(0.010)
R-squared	0.061	0.065

Notes: Full sample includes all women who had zero or one daughter before 1997 and 2002. The Restricted sample includes all women who had exactly one daughter before 1997 and 2002. n = 18,063 for both regressions with Full sample and n = 8,358 for both regressions with Restricted sample. For all specifications, the dependent variable is a dummy variable, girl, which takes the value of 1 if the woman in the sample had at least one daughter after the cut-off (1996 to simulate pre policy and 2001 to simulate post policy period) and 0 if not. The right-hand-side variables are a dummy for Haryana, post-policy period (2000), a Haryana post-interaction term, and covariates (age, education, husband's education, type of house as a proxy for standard of living index, and region of residence). Age is measured in number of years; education and husband's education are measured in number of years of schooling; region of residence is urban or rural; house type is kachha (mud house), semipucca (mix of mud, brick, and cement), and pucca (brick and cement). * = significant at 10% level ($p<0.05$). ** = significant at 5% level ($p<0.01$). *** = significant at 1% level ($p<0.001$).

Introduction (Jha et al. 2011; Ebenstein 2007).

Thus, it is evident that, while there appears to be some improvement in the relative likelihood of having daughters in the post-policy period in the border districts, it is difficult to determine whether they can be directly attributed to the policy. Additionally, as mentioned earlier, the incentive amount at the current rate is only a fraction of the average total cost of raising a child, which may be one of the primary reasons the policy did not have the desired impact. Also, the insignificant results across the full state of Haryana raise some questions about policy implementation and overall effectiveness.

VIII. POLICY IMPLICATIONS

Based on these results, it can be argued that incentive schemes may have some positive effects on parents' views about having daughters and improve the skewed sex ratio, albeit at a gradual pace. Such direct transfers can be particularly beneficial for impoverished families as it may help them invest in the education of their daughters, wherein the tussle is always about prioritization of limited resources for the son. However, the most important policy implication here is the substantial empirical evidence that shows that even though son preference may be prevalent across the society, it is mostly the educated and well-off who have access to sex-selective techniques. Thus, incentives targeting only the poorest families may overlook the need to reinforce positive perceptions about daughters. That being said, it is true that such incentives, along with targeted awareness campaigns, can help the most impoverished section of society by equalizing the benefits of having sons versus daughters.

Secondly, with the declining fertility rate and the promotion of a two-child policy across the country, the criteria of tying the incentive to the second daughter, irrespective of the number of sons, may not be the most desirable solution. For example, unless some families have a strong preference for daughters, they are unlikely to have daughters if they already have two or more sons. In this case, the scheme may merely attract people who already have a preference for daughters.

"... it can be argued that incentive schemes may have some positive effects on parents' views about having daughters and improve the skewed sex ratio, albeit at a gradual pace."

Thirdly, since one of the objectives of the policy is to "raise the status of the girl child in society," it does not necessarily make sense for the incentive to be restricted to only the second daughter; ideally, incentives should be extended to all daughters. But as this option requires a higher budget and outlay, it could be a solution implemented by the federal government, either alone or in combination with the state governments. Also, considering the rapid increase in per capita income of the state[3] and the nation, and the fact that the incentive amount at its current level constitutes only 3.57 percent of the total cost of raising a child, the Haryana state government could consider doubling the incentive from INR 5,000 to INR 10,000 in line with Delhi's successful *Ladli* Scheme (Sekher 2010).

Finally, because the third round of DLHS data covered the period soon after the *Ladli* scheme was introduced, the effects observed in this study might be different from the current sex-ratio levels. To evaluate the longer-term effects of the program, further assessments need to be done using data from more periods of the DLHS and, ideally, also using other available datasets, including administrative data (SRS) and other health surveys, such as

[3] Since 2005-06, the per capita income in Haryana rose to Rs. 56,922 (~US$1,122) in 2007-08, was Rs. 94,680 (~US$1,868) in 2010-11, and was estimated to reach Rs. 109,227 (~US$ 2,155) in 2011-12 (Economic Survey of Haryana 2010-2011).

the Demographic and Health Surveys (DHS). It would also be beneficial to have survey data that asked respondents whether the incentive had any effect on their decision to have or not have a second daughter. Most importantly, when assessing longer-term impacts, it is not sufficient to judge the success or failure of the program based upon its impact on the number of female births. Instead, one must look at the overall development and social status of female children in society. If there is reason to believe that more girls were born but suffered lifelong neglect and discrimination, the final outcome of the policy might be not be considered successful, whereas an increase in actual fondness for daughters, even if generated among fewer parents, might be a more successful outcome for society.

In conclusion, while the skewed sex ratio tarnishes the economic growth story of a modern India, there are reasons to believe that the right policymaking process and targeted interventions can aim to reverse such bias. In this process, the important thing to remember is that while short-term policy measures can bring about a change in the immediate outcomes, policymakers should focus on the long-term outcomes—the true development of the girl child and her later life outcomes—as that will be the primary contributing force to hasten the normative transformation in society.

IX. REFERENCES

Almond, Douglas, and Lena Edlund. 2008. Son-biased sex ratios in the 2000 United States Census. *Proceedings of the National Academy of Sciences* 105, no. 15: 5681-5682.

Almond, Douglas, Lena Edlund, and Kevin Milligan. 2009. O sister, where art thou? The role of son preference and sex choice: evidence from immigrants to Canada. No. w15391. National Bureau of Economic Research.

Arnold, Fred, Sunita Kishor, and T. K. Roy. 2002. Sex-selective abortions in India. *Population and Development Review* 28, no. 4: 759-785.

Chung, Heeran, and Monica Das Gupta. 2009. Gender discrimination in sex selective abortions and its transition in South Korea. *Women's Studies International Forum* 32, no. 2: 89-97. Pergamon.

Department of Planning, Government of Punjab. 2009. New schemes for girl child. February. http://pbplanning.gov.in/pdf/New%20Schemes%20for%20Girl%20Child.pdf.

Dhar, Aarti. 2002. Should ban on sex determination tests be relaxed? *The Hindu*, October 8, New Delhi, India.

Chung, Woojin, and Monica Das Gupta. 2007. The decline of son preference in South Korea: The roles of development and public policy. *Population and Development Review* 33, no. 4: 757-783.

Ebenstein, Avraham. 2007. Fertility choices and sex selection in Asia: Analysis and policy. SSRN 965551.

Edlund, Lena, Hongbin Li, Junjian Yi, and Junsen Zhang. 2007. Sex ratios and crime: Evidence from China's one-child policy. No. 3214. IZA Discussion Papers.

ET Wealth. 2011. How much it costs to raise a child. *The Economic Times Wealth*, April 17, India, http://economictimes.indiatimes.com/photo.cms?msid=7998976.

Department of Women and Child Development, Government of Haryana. 2007-08. Ladli: A new scheme for the girl child, http://wcdhry.gov.in/new_schemes_F.htm

Government of India. 2011. Gender composition of the population: Census 2011. New Delhi: Registrar General of India.

Guilmato, Christophe Z. 2007. Characteristics of sex-ratio imbalance in India, and future scenarios. *4th Asia Pacific Conference on Reproductive and Sexual Health and Rights*: 1-28. Hyderabad: United Nations Population Fund.

Haub, C., and O.P. Sharma. 2006. India's population reality: Reconciling change and tradition. *Population Bulletin, Population Reference Bureau* 61, no. 3: 1-24.

Hesketh, Therese, and Zhu Wei Xing. 2006. Abnormal sex ratios in human populations: Causes and consequences. *Proceedings of the National Academy of Sciences* 103, no. 36: 13271-13275.

Hudson, Valerie M., and Andrea Den Boer. 2002. A surplus of men, a deficit of peace: Security and sex ratios in Asia's largest states. *International Security* 26, no. 4: 5-38.

IANS. 2011. Though India's worst, Haryana's sex ratio is best in 110 years. *The Economic Times.*

Jagran Post Bureau. Bride-starved Haryana imports girls from other states. *Jagran Post*, May 25. Bhiwani, Haryana, India.

Jha, Prabhat, Maya A. Kesler, Rajesh Kumar, Faujdar Ram, Usha Ram, Lukasz Aleksandrowicz, Diego G. Bassani, Shailaja Chandra, and Jayant K. Banthia. 2011. Trends in selective abortions of girls in India: Analysis of nationally representative birth histories from 1990 to 2005 and census data from 1991 to 2011. *The Lancet* 377, no. 9781: 1921-1928.

Jha, Prabhat, Rajesh Kumar, Priya Vasa, Neeraj Dhingra, Deva Thiruchelvam, and Rahim Moineddin. 2006. Low male-to-female sex ratio of children born in India: National survey of 1.1 million households. *The Lancet* 367, no. 9506: 211-218.

Kanitkar, T. 1991. The sex ratio in India: a topic of speculation and research. *Journal of Family Welfare* 37, no. 3: 18-22.

Löfstedt, Petra, Luo Shusheng, and Annika Johansson. 2004. Abortion patterns and reported sex ratios at birth in rural Yunnan, China. *Reproductive Health Matters* 12, no. 24: 86-95.

Lin, Ming-Jen, and Ming-Ching Luoh. 2008. Can Hepatitis B mothers account for the number of missing women? Evidence from three million newborns in Taiwan. *The American Economic Review* 83, no. 5: 2259-2273.

Mahajan, N. 2011. Ladli scheme achieves a turn-around in Haryana. *The Daily Pioneer,* June 22.

Nandi, Arindam, and Anil B. Deolalikar. 2011. Does a legal ban on sex-selective abortions improve child sex ratios? Evidence from a policy change in India. *Social Science Research Network*.

Rosenzweig, Mark R., and T. Paul Schultz. 1982. Market opportunities, genetic endowments, and intrafamily resource distribution: Child survival in rural India. *The American Economic Review* 72, no. 4: 803-815.

Sen, Amartya. 2003. Missing women—revisited: Reduction in female mortality has been counterbalanced by sex selective abortions. *BMJ: British Medical Journal* 327, no. 7427: 1297.

Sen, Amartya. 1990. More than 100 million women are missing. *The New York Review of Books* 37.

Sekher, T.V. 2010. Special financial incentives for the girl child in India: A review of select schemes. International Institute of Population Sciences, Mumbai and United Nations Population Fund.

Singh, Narpinder, Pripp Are, Brekke Torkel, and Stray-Pedersen Babill. 2010. Different sex ratios of children born to Indian and Pakistani immigrants in Norway. *BMC Pregnancy and Childbirth* 10.

Sinha, Nistha and Joanne Yoong. 2009. Long-term financial incentives and investment in daughters: Evidence from conditional cash transfers in Northern India. Policy Research Working Paper WPS 4860, The World Bank.

Staff Correspondent. 2010. Hooda govt. extends "Ladli" scheme. *The Hindu*, January 23.

Strauss, John and Duncan Thomas. 1995. Human resources: Empirical modeling of household and family decisions. In *Handbook of Development Economics*, ed. Jere Behrman and T.N. Srinivasan, 3B: 1983-1988. Elsevier.

Subramanian, S. V., and Sakthivel Selvaraj. 2009. Social analysis of sex imbalance in India: Before and after the implementation of the Pre-Natal Diagnostic Techniques (PNDT) Act. *Journal of epidemiology and community health* 63, no. 3: 245-252.

Zhou, Xu Dong, Xiao Lei Wang, Lu Li, and Therese Hesketh. 2011. The very high sex ratio in rural China: Impact on the psychosocial wellbeing of unmarried men. *Social Science & Medicine* 73, no. 9: 1422-1427.

Effects of Converted Primary Enforcement Seat Belt Laws on Traffic Fatalities

By Christopher L. McCall

Christopher L. McCall completed his Master of Public Policy at the Georgetown Public Policy Institute in 2012. Omar Robles, PhD, served as his thesis advisor. McCall also holds a BA in Political Science from the University of Minnesota. He recently began work as a Legislative Assistant in the Minnesota State Senate, working for Sen. Ann Rest, the Chair of the Tax Reform Division.

ABSTRACT

As states continue to strive for safer roadways, seat belt laws remain a major policy issue in state legislatures across the country. The effectiveness of seat belt usage in saving lives during motor vehicle accidents is well documented, but enforcement methods vary among states. Some states and other actors continue to resist tougher seat belt laws on the grounds that they are either ineffective or violate personal freedoms. As of 2011, 24 states and the District of Colombia have upgraded from secondary to primary enforcement of existing seat belt laws. This paper analyzes the impact of tougher enforcement laws on reducing roadway fatality rates. This research builds upon previous works as it considers an additional 10 states that have converted to primary enforcement since the last known study. Using state-level panel data collected from the National Highway Traffic Safety Administration and the United States Census Bureau, this paper employs a series of fixed-effect regression models to determine if there are significant benefits observed in states that adopt stricter seat belt enforcement laws. The results indicate that, in terms of both lives and money saved, states experience significant benefits after upgrading existing secondary laws to primary enforcement. Closer examination reveals that this impact is not homogenous across states and that adopting primary laws may have outsized benefits for highly-populous states and specific geographic regions such as the Southeast and Pacific Coast.

I. INTRODUCTION

Sheer numbers demonstrate the true devastation of motor vehicle fatalities in the United States. From 1994 to 2009, the National Highway Traffic Safety Administration recorded 661,403 fatalities linked to motor vehicle accidents (NHTSA 2011). This signifies an average of 41,338 Americans killed each year by preventable accidents. In light of this continuing threat to public safety, policymakers must identify and replicate the laws most effective in reducing deaths caused by car crashes.

Unlike some policy arenas, a majority of traffic safety laws fall squarely in the jurisdiction of state-level statutes. The resulting diversity among states means that some successful policies may be underutilized. While there are certain factors (i.e., weather, congestion, and infrastructure) that cause states to implement different laws in pursuit of effective highway safety policy, many high-risk behavioral patterns are exhibited by vehicle occupants throughout the nation, and thus can be countered with the same legislation. Of these behaviors, failure to wear a seat belt is among the most dangerous.

The argument for stricter seat belt legislation relies on the consensus view that buckling up saves lives. NHTSA estimates that seat belts provide the most protection of any single safety apparatus, reducing potential deaths in crashes by 45-60 percent (2002). In light of this evidence, New York became the first state to pass a law mandating seat belt usage in 1984. Since then, every state except New Hampshire has passed some form of mandate for adult occupants, and all states require the use of seat belts by minors.

> "The argument for stricter seat belt legislation relies on the consensus view that buckling up saves lives."

While adoption of adult seat belt laws has been almost universal, state statutes still vary in their design. Among the most pronounced differences is the degree of power given to state and local officers in enforcing the law. In some states, seat belt mandates are subject to primary enforcement, meaning that vehicle operators who do not buckle up can be stopped and cited solely for that offense. However, other states only make these laws subject to secondary enforcement, prohibiting officers from pulling over or fining a driver for not wearing a seat belt unless another infraction has also occurred. Secondary enforcement is a relatively rare legal practice that relies largely on the theory that compliance with a statute will occur simply because of its existence.

Advocates of primary enforcement contend that the mandate gives vehicle occupants added incentive to buckle up due to the heightened risk of being pulled over. This in turn makes them safer in the event of an accident, reducing bodily harm and preventing deaths. In the past, state legislatures were also enticed to adopt primary enforcement because doing so would unlock conditional highway funds from the federal government. As this program expired in 2009, advocates now focus on publicizing the fiscal savings to state budgets generated by preventing highway deaths.

Advocates of secondary laws tend to utilize two lines of argument. The first is that primary enforcement, along with the mere idea of mandating seat belt usage at all, is an infringement on personal liberty. This contention is value-driven and loses much of its merit if the driver is not alone in his or her vehicle. Statistical evidence demonstrates that unrestrained vehicle occupants pose an immediate threat to other passengers since their bodies may become dangerous projectiles in the event of a high-speed accident (Redelmeier 2004).

The second argument against upgrading to primary enforcement provides the null hypothesis for this paper: that primary laws will not create a meaningful difference in traffic fatalities for states already using secondary enforcement. If this contention can be rejected with quantitative analysis, it would indicate that adopting primary enforcement actually enhances public safety. Specifically, a statistically significant difference between the impacts of secondary and converted primary laws would imply that the latter is a better policy mechanism for reducing roadway fatalities.

II. MOTIVATION

Past studies rely on panel data that is no longer up to date, thus creating a need for renewed research on seat belt legislation. My research

> "While adoption of adult seat belt laws has been almost universal, state statutes still vary in their design. Among the most pronounced differences is the degree of power given to state and local officers in enforcing the law."

utilizes data from 1994 through 2009 while other reports only contain observations through 2002. In that time period an additional 10 states have converted from secondary to primary enforcement, doubling the number of jurisdictions that have taken such action and providing a more robust sample for measuring the impact of upgraded enforcement mechanisms. Further, by using 1994 as the first year of data, I am able to use secondary enforcement as a nationwide baseline as opposed to a now defunct model where states have no seat belt laws.

The continuing political debate surrounding this issue suggests that this research may have real policy implications. As of my writing, 17 states have yet to adopt primary enforcement, representing roughly a quarter of the nation's population. Research accounting for the additional states that have adopted primary statutes in the last decade will enhance decision making for those legislatures deciding whether to upgrade existing laws. Furthermore, updated research can benefit states that are currently debating whether to repeal or downgrade existing seat belt laws. While my research does not directly test the consequences of this kind of policy reversal, my findings could help determine whether such actions would be detrimental to public safety.

III. LITERATURE REVIEW

Seat belt laws receive a large amount of attention in transportation policy. However, the focus of research has shifted over time as the national debate has moved from the actual utility of seat belts to the effectiveness of different policies encouraging their usage. Research specifically measuring the impact of converted primary enforcement laws has only emerged in the last decade.

Cohen and Einav (2003) run regression models on the effectiveness of varying seat belt laws using panel data from all 50 states for the years 1983-1997. Using these laws as instruments for actual usage rates, they find a 13.5 percent increase in seat belt usage for states that convert from secondary to primary enforcement, but their study lacks a direct measurement of the impact that tougher laws have on fatality rates. Research by Farmer and Williams (2004) specifically addresses this question by analyzing a sample of states from 1989 to 2003. Their results suggest there is approximately a 10 percent reduction in fatality rates among the treatment group compared to the control group.

Houston and Richardson (2006) also focus their research on the benefits of upgrading existing secondary laws, using panel data from 1990 to 2002. They control for demographic data, fixed effects, and policy controls such as graduated driver's license programs and blood-alcohol content limits for drunken driving arrests. They find that for the 10 states (along with the District of Columbia) that upgraded to primary laws between 1990 and 2002, annual fatality rates decreased by an average of 4.7 percent for all vehicle occupants and 5.1 percent for drivers.

IV. DATA

All data used in my models regarding state-level traffic deaths can be found in the Fatality Analysis Reporting System (FARS), a database maintained by NHTSA. Specifically, FARS provides state-by-state data for the annual rates of fatalities per 100 million vehicle miles travelled (VMT). I utilize public data available online for the years 1994 through 2009.

The Insurance Institute for Highway Safety provides a timeline for the adoption of seat belt laws in every state, allowing me to create indicator variables for the panel data. I ultimately choose to drop state-year observations for which there is no seat belt law in place, eliminating the state of New Hampshire from the panel data as well as another six observations from states that did not pass some form of mandate prior to 1994.[1] This is done to avoid having a control group that is too small to provide a reliable baseline. Instead, state-year observations with secondary enforcement laws become the baseline

[1] Initial seat belt laws fail to predate 1994 in Kentucky, Maine, Mississippi, North Dakota, and South Dakota

to which those with converted primary laws are compared.

While there may be concerns about bias or incorrect measurement in the NHTSA data, any problems would be systemic rather than varying by state due to a uniform reporting method at the federal level. Regardless, FARS provides the best available data due to federal reporting mandates, hence its usage here. Using the VMT rate rather than raw fatality totals provides a built-in control mechanism to account for any change in deaths that may be unrelated to occupant behavior, such as variance in population or changes in the level of roadway usage.

The state-level demographic data used in my models are publicly available online from the US Census Bureau. Populations are estimated for each year using projections calculated by the Bureau. Additional descriptive statistics taken from this source include median household income and population proportions for race and gender.

V. METHODOLOGY

I utilize my panel data to establish a difference-in-differences model, creating two control groups: 1) states that have only used secondary enforcement, and 2) pre-treatment observations from states that have converted from secondary to primary laws. My model also applies state and time fixed effects to control for national trends and unique state characteristics.[2]

[2] An OLS version of each model is also used to help understand the impact of omitting state fixed effects from the regressions.

Examples of year-based fixed effects include improvements in vehicle safety features and national economic shocks, which could trigger systemic increases or decreases in fatality rates. State fixed effects related to roadway fatalities may include climate, quality of infrastructure, and differences in "driving cultures" that cause drivers to be more or less responsible behind the wheel due to varying behavioral norms. The initial model used to measure the impact of converting to primary enforcement on fatality rates (*FatalitiesVMT)* is:

$$FatalitiesVMT_{st} = \beta_0 + \beta_1 SecToPrim_{st} + \Sigma\beta_s FE_s + \Sigma\beta_t FE_t + \varepsilon$$

The indicator variable *SecToPrim* has been coded to "1" only when observations have a primary enforcement law on the books that was updated from a previously standing secondary enforcement statute, as opposed to states where the original seat belt law utilized primary enforcement.[3] It should be noted that states that convert to primary laws will activate both the standalone primary variable and *SecToPrim*.[4] The latter

[3] For states with effective starting dates for their primary law that are later than April 30, the appropriate binary variable(s) for the new law will be activated beginning in the following calendar year to ensure that the law's immediate impact is not underestimated.

[4] While a "primary" variable is created, the way I construct my dataset means that all observations that activate this variable are already controlled for in fixed-effects regressions. This occurs through one of two ways. First, states whose original seat belt law used primary enforcement will have a value of "1" for all of their observations, making the primary law a fixed effect which is already controlled for in my models. Second, states for which the variable is

Table 1. Descriptive Statistics for State Time Series Data: 1994-2009 (n=794)

Variable	Mean	Std. Dev.	Minimum	Maximum
VMT Fatality Rate	1.57	0.43	0.61	2.94
Total Fatalities	827.57	814.13	29	4333
Seat Belt Law: Original Primary Enforcement	16.12%	36.80%	0	1
Seat Belt Law: Secondary Enforcement	65.62%	47.53%	0	1
Seat Belt Law: Converted Primary Enforcement	18.26%	38.66%	0	1
Median Household Income (2010 dollars)	$50,820.58	$7,677.52	$34,280.75	$73,598.42
State Population	5,683,292	6,240,953	474,982	36,961,229
Black (%)	11.79	11.69	0.34	64.9
Hispanic (%)	8.05	8.95	0.52	44.9
Male (%)	49.12	0.87	46.7	52.70

variable's coefficient (β_1) is of the most interest to my research because it examines whether states experience a statistically significant decline in the rate of fatalities per 100 million VMT after upgrading from secondary seat belt enforcement laws.

I then add demographic controls to this model to evaluate the robustness of my findings. In full, my final model will take on the following form:

$$FatalitiesVMT_{st} = \beta_0 + \beta_1 SecToPrim_{st} + \beta_2 Population_{st} + \beta_3 lnIncome_{st} + \beta_4 Black_{st} + \beta_5 Hispanic_{st} + \beta_6 Gender_{st} + \Sigma\beta_s FE_s + \Sigma\beta_t FE_t + \varepsilon$$

All of the variables added in this model are controls generated using demographic data from the US Census Bureau.[5] *Population* utilizes the Bureau's annual projections and accounts for differences in fatality rates that may be associated with having more or less total people on the roadways. I take the log of median household income to create *lnIncome,* which is included to account for the theory put forth by Christopher Ruhm that economic conditions in a state may influence the likelihood of motor vehicle accidents or crash-related fatalities (Cohen and Einav 2003).[6]

Values for *Black* and *Hispanic* are calculated by utilizing data compiled by the US Census Bureau to find the proportion of a state's population that identifies as each of those particular races in a given year. These variables are included due to observational data suggesting lower compliance rates for

activated upon upgrading a secondary law will have perfect correlation between that variable and *SecToPrim*.

[5] I initially attempted to use state median age as an additional control but ultimately omitted it. The measurement is not sophisticated enough to produce significant results because it does not tease out the possible impacts caused by high concentrations of particularly young or old drivers.

[6] Log is used to create more easily interpreted results, which display the impact of a percentage increase in income.

Table 2. Comparative Mean Statistics for State Time Series Data: 1994-2009				
Variable	Secondary (n=521)	Primary (n=128)	Converted Primary (n=145)	National Mean (n=794)
VMT Fatality Rate	1.64	1.43	1.45	1.57
Total Fatalities	701.33	1342.45	826.64	827.57
Black (%)	9.92	7.33	23.42	11.79
Hispanic (%)	5.98	18.83	5.99	8.05
Male (%)	49.17	49.31	48.75	49.12
Median Household Income (2010 dollars)	\$50,314.02	\$52,714.30	\$50,969.00	\$50,820.58
State Population	4,424,147	10,918,542	5,586,069	5,683,292

seat belt usage among both of these groups, particularly young passengers of both races (American Academy of Pediatrics 2011). Finally, *Gender* is calculated as the male percentage of a state's population and included due to lower seat belt usage rates among men and a demonstrated higher risk of them being killed in motor vehicle accidents (Borenstein 2007).

VI. HYPOTHESIS

In all models that control for state fixed effects, I hypothesize that the coefficient for *SecToPrim* (β_1) will be statistically significant with a negative relationship. This result would indicate that upgrading existing secondary enforcement laws to primary enforcement is associated with reductions in occupant fatality rates. It is difficult, however, to predict whether the magnitude of these correlations will be similar to past research.

VII. DESCRIPTIVE STATISTICS

Table 1 provides descriptive statistics for the data relevant to my research. For this dataset, I have dropped all observations for which there is no adult seat belt law, removing 22 observations from my model.

The staggeringly high standard deviations for fatalities and population highlight the importance of using the VMT fatality rate as my dependent variable rather than raw totals. There are also notably large standard deviations for the two race variables. The remaining demographic variables are far less volatile across state lines.

Table 2 displays the means for selected variables when the data is divided into subgroups based on enforcement law.[7]

This table demonstrates how relying upon total fatalities would give

[7] California was the first state to convert to a primary law and the only one to do so before 1994. I have coded the state's observations as if there was always a primary law because my model cannot capture a before-and-after effect.

misleading results. The nation's three largest states by population (California, Texas, and New York) all adopted primary legislation before the time frame of this study, meaning their larger fatality counts would be wrongfully linked to primary enforcement if the data were not standardized using VMT fatality rates. These three states also skew the average proportion of Hispanics living in jurisdictions with primary laws.

The stratified data for VMT fatality rates, found in Table 2, tell a story much closer to my hypothesis. Secondary states have the highest average rate of fatalities, while those with converted primary laws produce figures similar to states whose seat belt legislation originated with primary enforcement. However, regression analysis is still necessary to establish causation.

On average, African Americans represent a much higher proportion of residents in states with converted primary laws. This finding may not be coincidental. Since African Americans have lower observed compliance rates with seat belt laws than the national average, primary laws are often advertised to policymakers as a

Table 3. The Impact of Converted Seat Belt Enforcement Laws on Traffic Fatalities

Dependent Variable:	Fatalities per 100 Million VMT (National Average = 1.571)			
Independent Variable	Model 1 OLS	Model 2 State FE	Model 3 OLS	Model 4 State FE
Secondary to Primary Enforcement	-0.037	-0.049	-0.111***	-0.049*
	(0.039)	(0.030)	(0.027)	(0.025)
Log (population)	--	--	-0.028***	-0.744**
			(0.010)	(0.354)
Log (median income)	--	--	-2.000***	0.266*
			(0.064)	(0.152)
% Blacks	--	--	0.013***	0.034***
			(0.001)	(0.010)
% Hispanics	--	--	-0.001	-0.002
			(0.001)	(0.014)
% Male	--	--	0.250***	-0.107**
			(0.015)	(0.051)
Year FE	Yes	Yes	Yes	Yes
State FE	No	Yes	No	Yes
n	744	794	744	794
Adj. R-squared	0.117	0.596	0.673	0.628

Note: For exact definitions of the variables, refer to the methodology section.
Observations clustered by state.
Significance of the coefficient estimate at the 0.01 level ***, at the 0.05 level **, and at the 0.10 level *.

solution to improving usage among this demographic (Ellis et al. 2000). Despite concerns raised about the potential for increased racial profiling by law enforcement officers, research from states with upgraded laws suggests that primary enforcement does not lead to a disproportionate number of citations being issued to black or Hispanic drivers (NHTSA 2006).

VIII. REGRESSION RESULTS

Table 3 presents results for my proposed regressions of traffic fatality rates on enforcement laws and demographic controls. Models 1 and 3 use Ordinary Least Squares (OLS), which do not control for state fixed effects. Both models find negative relationships between fatality rates and upgrading to primary laws, but the coefficient for the *SecToPrim* indicator variable is only statistically significant in Model 3. The coefficient is also greater in magnitude in Model 3 than in Model 1. The only difference between these models is the inclusion of basic demographic controls in Model 3, suggesting that their omission makes it more difficult to identify the true impact of primary enforcement laws and creates an overall positive bias on the coefficient of interest when using OLS.

The OLS coefficient in Model 3 suggests that converting to primary enforcement of seat belt laws is correlated with a 0.111 decrease from the average fatality rate in states with secondary enforcement. The implied result is a 6.99 percent decrease in total motor vehicle deaths, which matches my expectations of a negative correlation. The five demographic controls also increase the overall explanatory power of the model (adjusted R-squared) by a large degree.

In Models 2 and 4, state fixed effects are taken into account such as climate and infrastructure. The *SecToPrim* coefficient (-0.049) does not change after adding an array of demographic controls in Model 4, suggesting that this result is robust. However, while the coefficient is not significant in Model 2, it is marginally significant at the 5.8 percent level after controlling for demographic variables. The different coefficient values for *SecToPrim* between Models 3 and 4 suggest that omitting state fixed effects creates a negative bias on the coefficient. If we choose to accept the results in Model 4, then the decision to control for state and time fixed effects allows us to conclude that converting to a primary enforcement law results in a 0.049 decrease from the average VMT fatality rate in secondary states, a 3.09 percent reduction.

Models 3 and 4 both control for the same demographic factors and produce similar findings with regard to statistical significance. Contrary to expectations from past research, neither regression shows a significant relationship between Hispanics and statewide traffic fatalities. Compared to Cohen and Einav's model, my results suggest a more robust relationship between the proportion of African Americans and traffic fatalities. One possible explanation for this finding comes from NHTSA observational data that demonstrates that adult black vehicle occupants have the lowest usage rate of seat belts. This problem persists even when children (age 12 and younger) are in the vehicle, meaning young passengers will likely emulate such risky behavior and face an increased probability of death in an accident (Glassbrenner 2008).

Table 4. Select Results of Using Dynamic Estimates

Independent Variable	Model 5 OLS	Model 6 State FE
-3 years from passage	-0.045	-0.026
	(0.061)	(0.038)
-2 years from passage	-0.105*	-0.069*
	(0.059)	(0.040)
-1 year from passage	-0.052	-0.013
	(0.058)	(0.038)
0 (year enacted)	-0.084	-0.046
	(0.058)	(0.038)
+1 year from passage	-0.105*	-0.081*
	(0.058)	(0.046)
+2 years from passage	-0.149**	-0.090*
	(0.058)	(0.045)
+3 years from passage	-0.157**	-0.076*
	(0.062)	(0.042)
+4 or more years	-0.157***	-0.054
	(0.033)	(0.045)
n	744	794
Adj. R-squared	0.677	0.630

Note: Control variables from Models 3 and 4 also included. Significance of the coefficient estimate at the 0.01 level ***, at the 0.05 level **, and at the 0.10 level *.

The male coefficient changes signs after applying fixed effects, though it is significant in both Models 3 and 4. Studies have indicated that men are more prone to severe crashes than women, but in crashes of equal severity women are more likely to be injured or killed (IIHS 2006). Thus, the true effect of this variable is hard to estimate and an omitted variable bias likely exists in the coefficient. Regardless, the impact of gender rates has been isolated from the true effect of seat belt laws in my model.

The variable for logged median household income is positive and statistically significant after accounting for state fixed effects. This also occurs in Cohen and Einav's regression analysis, and they partially attribute the finding to the theory that better economic conditions lead to increased car usage, making drivers more vulnerable to accidents and casualties (2003).

Finally, the coefficient for logged population indicates a negative relationship between population size and traffic fatalities, a finding that increases in significance after controlling for state fixed effects. This finding may defy expectations, as having more drivers on the road would increase traffic, creating more opportunities for motor vehicle accidents. However, the negative correlation could have several legitimate explanations. For instance,

Table 5. Reduced-Form Results for Era of Passage		
IV: Era Law was Upgraded	OLS Coefficient	State FE Coefficient
Early (Pre-2002)	-0.417*	-0.043
	(0.245)	(0.031)
Late (Post-2002)	-0.333	0.002
	(0.247)	(0.051)
T-test Early=Late	0.069	0.428
n	744	794
Adj. R-squared	0.675	0.628

Note: All variables from Models 3 and 4 also included.
Significance of the coefficient estimate at the 0.01 level ***, at the 0.05 level **, and at the 0.10 level *.

states with lower populations are likely more rural, requiring residents to spend more time driving on highways and operating their vehicles at faster speeds. This leads to a higher risk of being involved in fatal accidents as opposed to fender benders, which are more common on local roads with lower speed limits. Additionally, a higher percentage of residents in populous, urbanized states are likely using alternative modes of transportation such as public transit, bicycles, or walking.

IX. SUPPORTING EVIDENCE

The following tests intend to address potential challenges to my initial findings and search for differing trends in the impact of converted primary laws based on divergent state characteristics.

Dynamic Effects

To test the exogeneity of upgraded seat belt laws, I apply dynamic estimates to Models 3 and 4 to reveal if my main coefficient captures preexistent trends in fatality rates rather than acting as an independent effect. Unlike Models 1 through 4, the *SecToPrim* variable is replaced with indicator variables that denote the year that state observations occurred relative to when a converted primary law was passed.

Table 4 presents the results of the dynamic estimates test. When applying state fixed effects in Model 6, I find small, negative coefficients prior to states converting to primary enforcement. However, fatality rates fall at a larger and more consistent pace in the years immediately after those primary laws are enacted. Only the +4 year variable fails to show a significant reduction in fatality rates, indicating a possible diminishment in the effect of the law over time. This may be evidence that the ability of seat belt laws to change behavior is not consistent as the overall compliance rate rises. The riskiest drivers will be tougher to sway into abiding by the law than those who immediately start buckling up following the adoption of primary enforcement (Dee 1998).

Table 6. ***SecToPrim*** **Results Stratified by State Size (Population)**							
State Size in Population	Avg.VMT Fatal	Total States	Applicable States	N	Observations w/*SecToPrim*=1	*SecToPrim* OLS	*SecToPrim* State FE
Small	1.693	21	5	329	26	-0.051	-0.046
(under 3 million)						(0.076)	(0.078)
Medium	1.516	19	10	330	87	-0.130***	-0.052
(3-9 million)						(0.031)	(0.032)
Large	1.409	10	4	135	32	-0.131***	-0.079**
(over 9 million)						(0.035)	(0.030)

Note: Control variables from Models 3 and 4 also included.

Significance of the coefficient estimate at the 0.01 level ***, at the 0.05 level **, and at the 0.10 level *.

Overall, these findings suggest the possible existence of a small systemic decline in fatalities per VMT regardless of state legislation. However, following the adoption of converted primary laws, the rate of this decline noticeably increases in magnitude and significance, suggesting desirable behavioral changes in the wake of stricter regulations.

Adopting Statutes in Differing Eras

I also stratify states based on when laws were adopted to test for endogeneity between the impact of converted primary laws and the era in which those statutes were adopted. This test was conducted to detect any outsized effect for early adopting states, which may have passed stricter laws sooner because they had a larger problem with compliance or traffic fatalities. However, Table 5 shows that when controlling for fixed effects, no significant difference is detectable.

Differing Effects by Population Size

While my base model shows an inverse relationship between state population and fatality rates, it fails to capture any differing impact in upgrading seat belt laws based on state size. I

Table 7. ***SecToPrim*** **Results Stratified by Census Region**						
Region	Total States	Applicable States	N	Observations w/ *SecToPrim*=1	*SecToPrim* OLS	*SecToPrim* State FE
Northeast	8	2	126	11	0.105**	0.015
					(0.049)	(0.03)
Midwest	12	3	190	27	-0.054	-0.078
					(0.052)	(0.084)
South	17	12	270	97	-0.133***	-0.078**
					(0.035)	(0.028)
West	13	2	208	10	-0.087	-0.022
					(0.077)	(0.066)

Note: Control variables from Models 3 and 4 also included.

Significance of the coefficient estimate at the 0.01 level ***, at the 0.05 level **, and at the 0.10 level *.

stratify the states into three groups based on whether they make up less than 1 percent of the national population, more than 3 percent, or some proportion in between. The base regression models are then run separately for each group. Table 6 demonstrates the results.

After controlling for state fixed effects, small and medium states each produce coefficients similar to the aggregate figure of -0.049 from Table 3, though neither is statistically significant. However, the coefficient in large states is highly significant and exceeds the magnitude captured in Table 3, indicating a reduction of -0.079 from the average fatality rate in secondary enforcement states. This amounts to a 5.28 percent reduction in traffic deaths among the nation's most populous states that adopt primary enforcement. These results suggest that upgraded seat belt laws are more effective in highly populated states than they are in smaller ones, meaning larger states may have more incentive to change their policies than less populous ones.

Table 8. Projected Average Annual Impact of Upgrading Seat Belt Laws for States Still Using Secondary Enforcement

(Simulating a 3.09% reduction in fatalities for year 2009)

State	Projected Reduction in Annual Fatalities	Economic Savings*
Arizona	23.62	$22,446,287.11
Colorado	16.54	$14,177,103.14
Idaho	6.25	$6,216,649.28
Massachusetts	10.67	$9,047,763.40
Missouri	28.74	$22,229,846.19
Montana	6.52	$5,386,655.78
Nebraska	7.25	$5,889,169.13
Nevada	7.12	$6,573,922.31
North Dakota	4.08	$3,771,666.51
Ohio[a]	57.84	$47,957,954.98
Pennsylvania[a]	65.42	$59,823,257.78
South Dakota	4.38	$3,265,537.06
Utah	7.29	$7,252,993.59
Vermont	2.13	$2,210,152.76
Virginia	20.79	$20,658,480.55
West Virginia	11.2	$9,723,691.11
Wyoming	4.19	$3,216,896.68
Total:	**284.03**	**$261,701,915.05**

[a] Calculated at a reduction rate of 5.28% due to differing measured impact of laws in states with large populations.

*Economic savings calculated using Center for Disease Control state-level data from 2005 on total costs of traffic deaths.

Source: http://www.cdc.gov/motorvehiclesafety/statecosts/

Regional Differences

While the base models in Table 3 control for state fixed effects, they do not detect possible regional differences resulting from varying driving cultures or conditions. I thus group states into Northeast, Midwest, South, and Western regions as defined by the US Census Bureau, and replicate my models separately for each area.

Table 7 shows that roughly two-thirds of state-year observations with converted primary laws occur in the South. As a result, the sample size of applicable observations is rather small in other regions and it is difficult to discern a noticeable effect for upgraded enforcement statutes. However, in Southern states where there are a number of converted primary observations, there is a significantly larger decrease in fatality rates than the national average demonstrated in Table 3.[8] This suggests an outsized effect

[8] I also stratify states using the Census Bureau Divisions, which compartmentalize each region into more homogenous subgroups. Doing so reveals statistically significant drop-offs in fatality rates for states with converted primary laws in the Southeast and along the Pacific coast.

"These results suggest a sizeable potential impact for states that choose to upgrade their seat belt laws. In addition to saving lives, states and families also stand to experience a large reduction in the economic costs suffered from fatal roadway crashes."

for converted primary laws in this part of the country. It is hard to make sweeping conclusions about differing regional effects from this test. However, in light of the interesting results, I would encourage further research on this topic.

X. DISCUSSION

In Table 8, I indicate the potential implications of my results for states that, as of this writing, still utilize secondary enforcement seat belt laws.[9] Using the output from Model 4 in my primary results, as well as the findings from Table 6 on state size, I project the reduction in annual deaths for these states based on a fixed percentage reduction of their 2009 fatality counts. For most states I calculate reductions using an estimated decline of 3.09 percent in fatality rates. For the two large states of Ohio and Pennsylvania, I use my findings from Table 6 to apply an estimated 5.28 percent decrease in deaths per 100 million VMT.

Table 8 also displays my projected fiscal savings using CDC data from 2005 on the state-level financial burdens of highway deaths. The price faced by state governments when a traffic fatality occurs includes medical expenses, legal processes, and work-loss costs. Since these figures are slightly outdated, my estimated economic savings may actually be somewhat conservative due to inflation.

These results suggest a sizeable potential impact for states that choose to upgrade their seat belt laws. In addition to saving lives, states and families also stand to experience a large reduction in the economic costs suffered from fatal roadway crashes. The implications of primary enforcement are particularly powerful for populous states like Ohio and Pennsylvania. Even if the projections for these two states utilized the original coefficient from Model 4, they would still be projected to save a combined 74 lives and roughly $68.7 million annually.

XI. CONCLUSIONS

The results from this paper indicate that states that convert from secondary to primary enforcement of seat belt laws experience a significant decrease in traffic fatality rates and related economic costs. This effect is exogenous from subtle pre-existing trends. Furthermore, additional analysis shows that this drop in fatalities is more prominent in specific subgroups of states, including those with the largest populations and those in the South or the West Coast. States that adopt tougher seat belt

[9] States that have converted to primary enforcement between the end of my data period and the writing of this paper include Arkansas, Florida, Kansas, Minnesota, Rhode Island, and Wisconsin.

laws also stand to experience notable financial savings by foregoing the heavy economic costs associated with traffic deaths.

The implications of this research are important for state policymakers. While the sunset date has passed for receiving federal funds in exchange for adopting primary laws, my research provides evidence that such laws are still worthwhile because they significantly reduce traffic deaths and fiscal costs to state governments. If every state currently using secondary enforcement were to upgrade their statutes, my analysis suggests that it would save approximately 284 lives annually. These figures, even when given some room for error, are sizeable and worth consideration by all policymakers.

Further research on this topic is warranted. States that have only recently adopted stricter laws have had little time to measure the results of their new policies. By replicating my regression analysis in the future, more robust data can be used to verify my findings. I also recommend that future researchers with more time and resources attempt to add controls for several policies that work in conjunction with seat belt mandates to reduce fatalities. These laws include speed limits, drunk driving statutes, and distracted driving restrictions.

Even without controlling for these factors, it is clear from this research that converting seat belt laws from secondary to primary enforcement significantly reduces statewide traffic fatalities. Therefore, continued efforts to adopt stricter seat belt legislation in the remaining states with secondary enforcement are both admirable and worthwhile.

XII. REFERENCES

American Academy of Pediatrics. 2011. Minority children less likely to wear a car seatbelt, putting them at greater risk of severe injury. *AAP.org*, October 15, http://www.aap.org/pressroom/seatbelts.pdf.

Borenstein, Seth. 2007. Women drivers? They're safer than men. *Msnbc.com*, Jannuary 20, http://www.msnbc.msn.com/id/16698153/ns/technology_and _science-science/t/women-drivers-theyre-safer-men/.

Cohen, Alma and Liran Einav. 2003. The effects of mandatory seat belt laws on driving behavior and traffic fatalities. *The Review of Economics and Statistics* 85, no. 4: 828-843.

Ellis HM, B. Nelson, O. Cosby, L. Morgan, W. Haliburton, P. Dew. 2000. Achieving a credible health and safety approach to increasing seat belt use among African Americans. *Journal of Health Care for the Poor and Underserved* 11, no. 2: 144-50.

Farmer, Charles M. and Allan F. Williams. 2004. Effect on fatality risk of changing from secondary to primary seat belt enforcement. Insurance Institute for Highway Safety, http://www.gahighwaysafety.org/pdf/iihsseatbeltsof.pdf.

Glassbrenner, Donna and Tony Jianqiang Ye. 2008. Seat belt use in 2007–race and ethnicity results among occupants traveling with children. National Highway Traffic Safety Administration, http://www-nrd.nhtsa.dot.gov/Pubs/810896.pdf.

Houston, David J. and Lilliard E. Richardson Jr. 2006. Reducing traffic fatalities in the American states by upgrading seat belt use laws to primary enforcement. *Journal of Policy Analysis and Management* 25, no. 3: 645-659.

Insurance Institute for Highway Safety. 2011. Safety belt and child restraint laws. *IIHS.org*, http://www.iihs.org/laws/safetybeltuse.aspx.

Insurance Institute for Highway Safety. 2006.

Fatality facts 2006–gender. *IIHS.org,* http://www.iihs.org/research/fatality.aspx?topicName=gender&year=2006.

National Highway Traffic Safety Administration. 2006. Primary enforcement saves lives: The case for upgrading secondary safety belt laws. *NHTSA.gov,* http://www.nhtsa.gov/people/injury/enforce/PrimaryEnforcement/images/PrimaryEnforcement.pdf.

National Highway Traffic Safety Administration. 2008. States with primary enforcement laws have lower fatality rates (updated). NHTSA National Center for Statistics and Analysis, http://www-nrd.nhtsa.dot.gov/Pubs/810921.PDF.

National Highway Traffic Safety Administration. 2008. Seat belt use in 2008–demographic results. NHTSA National Center for Statistics and Analysis, http://www-nrd.nhtsa.dot.gov/Pubs/811183.PDF.

National Highway Traffic Safety Administration. 2011. Fatalities and fatality rates by STATE, 1994-2009. *NHTSA Fatality Analysis Reporting System Encyclopedia.*, http://www-fars.nhtsa.dot.gov/Main/index.aspx.

Redelmeier, Donald A. and Harriette G. C. Van Spall. 2004. Is there a greater risk of dying in a car crash when other passengers are unrestrained?" *Canadian Medical Association Journal* 170, no. 12: 1793.

United States Census Bureau. 2011. Population estimates data. *Census.gov* 1994-2009, http://www.census.gov/popest/data/intercensal/index.html.

United States Census Bureau. 2012. Census regions and divisions of the United States. *Census.gov,* http://www.census.gov/geo/www/us_regdiv.pdf.

The Relationship between Parental Receipt of the Earned Income Tax Credit and Children's High School Dropout Status

By Galen Savidge-Wilkins

ABSTRACT

Galen Savidge-Wilkins completed the Master of Public Policy at the Georgetown Public Policy Institute in 2012. Adam Thomas, PhD, served as his thesis advisor. Currently, Savidge-Wilkins is an Analyst with Abt Associates where he works on research and evaluation of federal homeless assistance, income security, and workforce development programs.

The Earned Income Tax Credit was established to provide low-income families with relief from the payroll tax, but it has grown over time to become the largest means-tested cash transfer program in the United States and one of the most substantial federal supports for the working poor. Although the EITC has been studied extensively, the literature has largely focused on its ability to encourage work, particularly among mothers (Eissa and Liebman 1996; Eissa and Hoynes 1998; Meyer and Rosenbaum 2001). There is a growing research literature on other effects of the EITC, including its impact on child cognitive ability and maternal health (Dahl and Lochner 2011; Evans and Garthwaite 2011). However, there is little research into how this component of the safety net impacts a key outcome for children: high school graduation. Using data from the National Longitudinal Survey of Youth on mothers and their children to construct a rich personal history for each child, this study examines the relationship between parental receipt of the Earned Income Tax Credit and children's likelihood of graduating from high school on time. Within this study's sample, EITC receipt is found to be most strongly associated with on-time graduation when these benefits are received during two life stages: before the children enter school and when children are in middle school. The results indicate that a $1,000 increase in average yearly real EITC receipt before the children enter school is associated with a 6.80

percentage point increase in the likelihood of finishing high school on time, and with a 1.56 percentage point improvement when that same increase in average real EITC receipt occurs during middle school. Analyses of specific disadvantaged subgroups yield statistically significant results during the same life stages and a stronger positive relationship between EITC receipt and high school graduation.

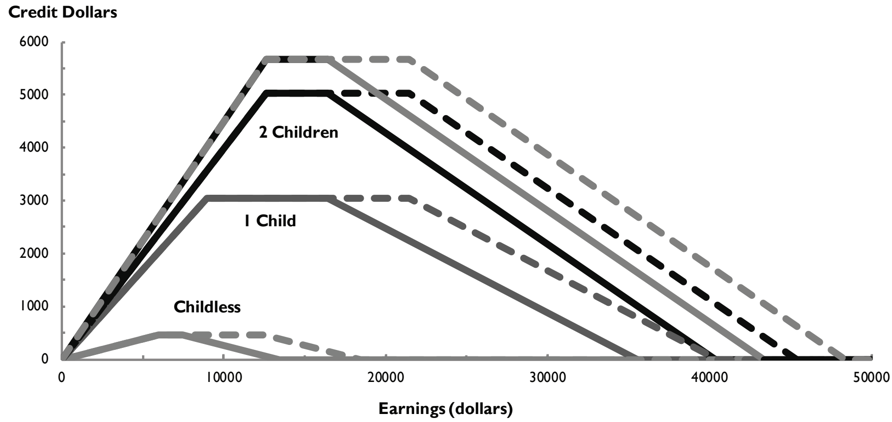
I. INTRODUCTION

As the Great Recession eases for some and drags on for many more, it is becoming clear that the impacts of the downturn and prospects for recovery vary enormously across individuals. One of the widest divides, even prior to the economic downturn that began in the winter of 2007, was between adults who completed high school and their peers who dropped out. Failing to complete high school is associated not only with a negative cultural perception but also with important adverse outcomes that include poor labor market performance and an array of negative social issues ranging from imprisonment to poor health later in life (Amos 2008). When high school dropouts enter the workforce, they do so with lower cognitive abilities, fewer demonstrated academic skills, and more pronounced behavior problems that may impact soft skills in the work world (Sellers 2011). Average earnings for high school dropouts in 2009 were $19,540, as compared to $27,380, $36,190, and $46,930 for individuals with high school diplomas, associate's degrees, and bachelor's degrees, respectively (Amos 2008). As of 2011, with employment growth relatively stagnant, high school dropouts face 14.9 percent unemployment compared to 10.3 percent for high school graduates and 5.4 percent for bachelor's degree holders (Bureau of Labor Statistics 2011).

"When high school dropouts enter the workforce, they do so with lower cognitive abilities, fewer demonstrated academic skills, and more pronounced behavior problems ..."

A wide variety of interventions are aimed at mitigating the effects of personal, academic, and family factors associated with dropping out of high school. This research focuses on identifying the relationship between high school graduation and the Earned Income Tax Credit (EITC), a significant family support policy that is not directly intended to enhance educational attainment or to prevent dropout among the children of the recipients, but can substantially boost the incomes of poor families. The Brookings Institution estimates that in 2010 $374 billion, or 11 percent of all federal expenditures, went to programs

Figure 1. 2010 Earned Income Tax Credit by Filing Status and Number of Children



Source: EITC parameters taken from http://www.taxpolicycenter.org/taxfacts/displayafact.cfm?Docid=36

or policies benefitting children (Isaacs et al. 2011). This includes $58 billion in expenditures on the EITC, the largest means-tested cash transfer program in the federal budget. Using data from the National Longitudinal Survey of Youth on mothers from the 1979 cohort and their children, I construct a rich personal history for each child and his or her family to examine the relationship between parental receipt of the Earned Income Tax Credit and their children's likelihood of graduating from high school on time.

II. BACKGROUND

The Earned Income Tax Credit

The Earned Income Tax Credit is a provision of the tax code that is designed to benefit low- and moderate-income workers. Established in 1975, the EITC was originally intended to be an offset to the payroll tax for low-income workers (Hotz and Scholz 2002). The program has since grown to be the single largest cash assistance program targeted at low- and moderate-income workers in the United States, disbursing $58 billion in refunds and reduced tax liability in 2009 (Internal Revenue Service 2011). The EITC's benefit schedule functions through the tax code and is designed to encourage work. Its refundable structure means poor filers with little tax liability receive the balance of the credit in the form of a check from the IRS. The value of the credit varies by income, marital status, and number of dependent children, but the average refund check received in 2009 was $2,770 for a filer with children and $259 for a childless filer (Center on Budget and Policy Priorities 2011).

The structure of the EITC for every category of filer can be broken down into three primary pieces: the phase-in stage, the plateau stage, and the phase-out stage. Figure 1 depicts these stages for both single and married filers. The phase-out stages are shown as the solid lines on the left for single filers and the dashed lines on the right for married

"Estimates indicate that cash delivered through the EITC lifts millions of children and families out of poverty every year; this effect has grown during economic downturns, possibly shielding a generation of children from the potentially damaging long-term effects of poverty."

filers. During the phase-in stage, the credit increases in value as a percentage of every dollar earned, effectively acting as a wage subsidy, up to the maximum dollar amount of the credit. In the plateau region the credit remains constant until earned income reaches the phase-out point. Beyond this point, each additional dollar of earned income reduces the value of the benefit until the amount of the credit received reaches zero. The three-stage design and the variations that exist across the spectrum of tax filers receiving the EITC create a variety of incentives for beneficiaries of this conditional cash assistance program.

High School Dropout Rates

The National Center for Education Statistics at the US Department of Education tracks two important measures that fit the general perception of a high school dropout rate. The *status completion rate* tracks 18-24 year olds who have earned a diploma, irrespective of when it was earned. The *average freshman graduation rate* estimates the proportion of high school students who graduate with a diploma four years after starting the ninth grade (Chapman 2010). Nationally in 2008, the status completion rate was 89.9 percent, and the average freshman graduation rate was 74.9 percent (Chapman 2010). According to the data, high school dropouts are more likely to be male than female, dropout rates are higher among Hispanics and African Americans than their white and Asian-American peers, low-income students are more likely to drop out than their more economically well-off peers, and students from the West and South drop out more frequently than students in the Midwest and Northeast (Rumberger 2004). The dropout rates by gender, race, socioeconomic status, and region all declined between the mid-1970s and 1990, rose slightly until 1995, and have declined steadily since then (Chapman 2010).

III. LITERATURE REVIEW

The focus on the EITC's potential to improve outcomes for the children of its recipients is motivated by a growing body of literature linking poverty during childhood to a range of negative outcomes later in life (Duncan 1998; Holzer et al. 2007; Moore 2009). Estimates indicate that cash delivered through the EITC lifts millions of children and families out of poverty every year; this effect has grown during economic downturns, possibly shielding a generation of children from the potentially damaging long-term effects of poverty (Eamon and Wu 2009). With the focus on the EITC, I hope to contribute to the understanding of how near-term increases in family income impact a particular long-term outcome—high school completion—for children in low- and moderate-income families.

The EITC is designed to encourage work and to ensure that individuals, particularly parents, do not fall below a certain base income if they participate in the labor force. Research on the impact of the EITC has focused primarily on mothers' decisions to enter the labor force (Eissa and Liebman 1996; Eissa and Hoynes 1998; Meyer and Rosenbaum 2001). Other studies focus on the EITC's impact on various parental and child outcomes, such as maternal health and children's cognitive abilities, and hypothesize that the EITC can be beneficial for children and their families (Dahl and Lochner 2011; Evans and Garthwaite 2011).

Research surrounding the theoretical assumptions that encourage work generally exploits statutory changes to the EITC during major expansions as an exogenous source of variation to isolate the impact of the program (Eissa and Liebman 1996; Eissa and Hoynes 1998; Hotz and Scholz 2005). Comparing single mothers to other single women before and after the 1986 expansion, Eissa and Liebman find an increase in labor force participation among single mothers receiving the EITC and no measurable reductions in hours worked among those already in the labor force (1996). Additionally, Eissa and Nichols examine wage data around major expansion periods and find no clear evidence that increases in EITC benefits are offset by a reduction in wages among the recipients (2005).

In an attempt to provide insight into how EITC receipt changes behavior and outcomes among recipients and their children, there is a growing body of research examining how refund checks are spent. Families seem to treat the money delivered through the EITC differently than other sources of income. The literature shows that the EITC is largely received and spent during February and March, indicating that many families anticipate a tax refund (Goodman-Bacon and McGranahan 2008). The families receiving the credit generally spend the large infusion of income on durable home goods, transportation needs like auto repair, and payment of bills, debt, and rent (Goodman-Bacon and McGranahan 2008; Smeeding et al. 1999). These expenditures tend to increase economic stability or provide support when a family faces hardships that could adversely impact their children's education (Smeeding et al. 1999; Mendenhall et al. 2012).

A meta-analysis of existing research concludes that the major expansions of the EITC in the early 1990s were more important contributors to the increased labor market participation of single women than welfare reform or a booming economy in that decade (Hoynes 2008). Because the EITC plays such a major role in the economic lives of low-income mothers, researchers have sought to identify other aspects of life that the credit impacts. Increases in the credit have been linked to improvements in the self-reported mental and physical health when comparing mothers with one and two children around the time of the 1993 expansion (Evans and Garthwaite 2011). Indirect relationships like this one are not limited to adults. Dahl and Lochner instrument for family income using exogenous changes in EITC benefit levels and find that additional income increases cognitive development as measured by reading and math scores among the children of recipients (2011).

The implications of child poverty are important to understanding the ways in which the EITC has the capacity to address educational outcomes as it boosts family incomes. Research indicates that poverty—measured through mediating factors such as physical and mental health, home environment, and educational resources—has a negative relationship with a range of child cognitive development measures (Yeung et al. 2002). The anti-poverty impact on children's outcomes could be substantial, given that the EITC reduced the child poverty rate by nearly 20 percent in 2005 (Eamon and Wu 2009). The hypothesis that the EITC will have a positive impact on the school achievement of recipients' children is based in the program's demonstrated ability to lift children out of poverty and increase family stability.

High School Dropout Rates

Examining the factors associated with high school completion provides more insight into my particular research question. Studies estimating the effects of income on children using sibling fixed-effects models find that increased income, particularly during early childhood, is associated with higher levels of school completion (Duncan et al. 199; Levy and Duncan 1999). Increased family income is also associated with improvements in individual and household characteristics that are associated with school success such as stress, supportive parenting, and acquisition of developmentally stimulating home resources (Yeung et al. 2002). Individual student predictors, such as earlier academic performance and behavioral characteristics, combine with demographics, family resources, and school characteristics to provide a broad picture of what determines whether students will complete high school (Rumberger and Lim 2008). The general consensus within the literature is that failing to complete school is the result of an additive process in which multiple factors from early childhood through the teenage years (e.g., educational performance, school quality, and family characteristics) can push a child to drop out of school (Rumberger 2004).

IV. DATA & METHODS

I utilize data from the National Longitudinal Survey of Youth 1979 (NLSY79) and NLSY Child and Young Adult Survey (C-NLSY) to match information on mothers with information on their children. This allows me to build individual observations containing detailed family information that can be used to measure the relationship between the EITC and dropout status. I am then able to construct a detailed cross section in which the unit of analysis is the child.

The NLSY79 is a nationally representative sample of non-institutionalized individuals who were between 14 and 22 years of age when the survey was first administered in 1979 (Bureau of Labor Statistics 2008). The survey collects detailed information about labor market activity, earnings, a range of family outcomes, and individual background data. The C-NLSY is composed of the children born to women in the original NLSY79 survey. The C-NLSY, which was first administered in 1986, asks questions similar to the questions asked in the NLSY79, but it includes more detail on parenting and individual developmental milestones for each child.

Matching all available data from NLSY79 and C-NLSY yields a full sample of 4,929 mothers and 11,495 children. The population of analysis is constructed from children who have valid information regarding high school completion and who have reached the age at which high school graduation is possible. On-time graduation in the model is defined as having completed 12th grade before the child's 20th birthday. This binary variable represents a close approximation of the average freshman graduation rate, the on-time graduation measure tracked by the National Center for Education Statistics (Chapman et al. 2010). Limiting the sample to those children and their parents, I retain 2,464 mothers with

"Studies estimating the effects of income on children using sibling fixed effects models find that increased income, particularly during early childhood, is associated with higher levels of school completion."

4,015 children. Because many mothers in this sample have more than one child tracked by the C-NLSY, parental and family characteristics are replicated for siblings in a given year.

The Earned Income Tax Credit is only tracked by the NLSY79 biennially starting in 1999. In order to create a consistent series of EITC benefits over the full range of time, data necessary to measure eligibility for the EITC—mother's income, spouse's income, qualifying child dependents, and marital status—are imputed for every year in the survey. Because the NLSY began interviewing biennially starting in 1994, personal income and spouse's income are both interpolated for years in which the survey was not conducted by using an average of income from the years surrounding that in which the value is missing.

In order to impute EITC information, I utilize the tax micro-simulation tool TAXSIM, administered by the National Bureau of Economic Research and maintained by Daniel Feenberg. TAXSIM utilizes data on income, marital status, and number of dependents to calculate each family's eligibility level for the EITC in a tax year (Feenberg and Coutts 1993). TAXSIM does not, however, determine the likelihood of filing a return or the specific take-up of the EITC. Therefore, every value obtained in this imputation process is the level of EITC eligibility for a family in a given year and not necessarily an exact benefit amount received. Estimates of the take-up rate range from 80 to 87 percent, and there is insufficient evidence within the NLSY79 to distinguish recipients of the credit from non-recipients (Hotz and Scholz 2002). I utilize the approach taken by Dahl and Lochner, which is to use TAXSIM-generated eligibility values for the analysis and implicitly assume full take-up of the EITC (2011).

Model

I construct a linear probability model depicted below in which the dependent variable of interest is the categorical variable completed 12 years of education before the age of 20.

$$Pr_{(1\text{-}0)}\ \textit{On-Time High School Graduate}_i = \beta_0 + \beta_1\ \textit{Parental EITC Receipt}_i + \beta_2 \textit{Family Economic Measures}_i + \beta_3 \textit{Maternal Background Characteristics}_i + \beta_4 \textit{Child Background Characteristics}_i + e_i$$

The key independent variable of interest is *Parental EITC Receipt*. In order to reduce omitted variable bias in the estimates of the relationship between EITC receipt and child educational attainment, I control for other variables that are plausibly correlated with receiving the credit and high school graduation. The personal history for each child is represented by three vectors of covariates: *Family Economic Measures*, *Maternal Background Characteristics*, and *Background Characteristics* of the children in the sample.

Table 1. Graduation Rates Among Children in Study Sample

Characteristic		Number of Cases	Percent Graduated
Full Sample:			
	Average Freshman Graduation Rate	4,015	47.8
	Status Completion Rate	4,015	70.1
By Subgroup:			
	Male	2,029	41.4
	Female	1,986	54.4
	White	1,751	51.3
	Black	1,353	44.1
	Hispanic	911	46.5
	Parents in Poverty Before Birth	1,683	39.5
	Father was not in Household at Birth	1,160	46.7
	Attended Head Start School	890	37.5

AFGR computed using number of children completing 12 years of education before age 20.

SCR computed using number of children completing 12 years of education before age 25.

Rates given in subgroup analysis computed in AFGR.

Table 2. Mean Real EITC Levels by Life Stage				
Life Stage	Full Sample		Ever Received EITC	
	Average given year of receipt	Average total receipt in stage	Average given year of receipt	Average total receipt in stage
Lifetime EITC	$542.50 $(613.60)	$9,140.80 $(10,542.60)	$687.90 $(614.30)	$11,592.00 $(10,608.40)
EITC in Early Childhood	$333.60 $(450.50)	$3,181.10 $(4,311.30)	$423.10 $(468.60)	$4,034.10 $(4,486.70)
EITC in Late Childhood	$814.80 $(1,035.00)	$5,959.70 $(7,822.30)	$1,033.40 $(1,064.20)	$7,557.90 $(8,094.40)
EITC Before School Years	$232.50 $(331.90)	$1,358.50 $(1,945.50)	$294.80 $(348.40)	$1,722.80 $(2,042.60)
EITC in Elementary School	$484.90 $(782.60)	$1,822.50 $(2,975.90)	$614.90 $(834.80)	$2,311.30 $(3,178.30)
EITC in Middle School	$747.10 $(1,103.90)	$2,765.10 $(4,189.10)	$947.40 $(1,164.40)	$3,506.60 $(4,433.40)
EITC in High School	$859.20 $(1,260.40)	$3,194.60 $(4,787.60)	$1,089.70 $(1,328.00)	$4,051.20 $(5,059.40)
	n=4015		n=3166	

Standard deviations in parenthesis.

Real values given in 2008 dollars, calculated from the CPI-U RS.

V. DESCRIPTIVE ANALYSIS

Table 1 contains a breakdown of the two distinct dropout rates: the average freshman completion rate and the status completion rate among the children in the sample. Specific rates are also listed among subpopulations of interest. Table 2 contains descriptive statistics on the EITC variables that are used as the independent variables of interest in later regression models. I present frequency distributions for a range of categorical background characteristics among the mothers in the sample in Table 3, and background characteristics specific to the children in the sample are illustrated in Table 4. Summary statistics for continuous variables used as background characteristics of both mothers and their children are presented in Table 5.

High school graduation data representing the dependent variable in later analyses are displayed in Table 1. The table shows that 47.8 percent of children in the sample reported that they had reached the equivalent of an on-time high school graduation, 12 years of education before age 20. This rate is well below the national Average Freshman Completion Rate, which was 74.9 percent in 2008 (Chapman et al. 2010). The substantially lower rate in the data is likely due to over sampling of minorities and more economically disadvantaged mothers in the original NLSY79 cohort. On-time graduation rates among white, African American, and Hispanic children in the sample are lower than those of their peers nationally (Chapman et al. 2010).

The EITC measures used as the independent variables of interest in the multivariate analyses—real average yearly EITC receipt during various childhood life stages—are displayed in Table 2. Average total receipt is also presented to provide a reference point for the amount of EITC the families are receiving, given that many receive the credit in more than one year over these same stages. The simplest measure, average EITC receipt throughout childhood, is presented first. Average EITC is then broken into two groups: receipt over early childhood (years from birth through age 9) and late childhood (ages 10 through 17). Next, average EITC receipt is shown over four distinct child life stages whose lengths roughly approximate different stages in students' school lives. These stages include high school (ages 14 through 17), middle school (ages 10 through 13), elementary school (ages 6 through 9), and before school (birth through age 5). Lower real values of average EITC receipt in earlier stages of the child's life are likely due to the fact that, on average, changes in policy that increased the real value of the EITC came into effect later on average in these children's lives.

Tables 3, 4, and 5, found in the Appendix, display the comprehensive list of parental and child characteristics that are used as covariates in the analyses because of the plausible correlation with both high school graduation and EITC receipt. These variables represent measures of personal disadvantage and health, indicators of background and economic activity, and measures of personal ability. In one clear cross-generational comparison in the data, I include mother's years of educational attainment to gauge success in school and find that 74.3 percent of the sample completed at least 12 years of school. The comparable 70.1 percent status completion rate among the children in the sample represents a small decline from their mothers. However, that decline is much more notable considering that graduation rates have steadily increased over the last three decades. This surprising result could indicate that the children in this sample face more persistent disadvantages than their mothers.

VI. RESULTS

The analysis of the full sample begins with the simplest form of the independent variable of interest, average real EITC receipt. I then allow average real EITC receipt to vary over two and then four stages of a child's life. In Table 6, I present the first specification of the average annual real EITC received over the course of a child's life. I change the specification to two life stages, early and late childhood, in Table 7, and I allow the effect of average EITC receipt to vary over four stages of a child's life in Table 8. In Table 9, I display

Table 6. Full Sample: Lifetime Specification				
Full Sample	(1)	(2)	(3)	(4)
VARIABLES	On-Time Graduate	On-Time Graduate	On-Time Graduate	On-Time Graduate
Average EITC Lifetime	-0.0559***	0.0339**	0.0202	0.0532***
	(0.0129)	(0.0137)	(0.0141)	(0.0201)
Family Income		X	X	X
Maternal Characteristics			X	X
Child Characteristics				X
Constant	0.509***	0.307***	0.571***	0.558***
	(0.0105)	(0.0160)	(0.0619)	(0.0669)
Observations	4,015	4,015	4,015	4,015
R-squared	0.005	0.069	0.142	0.321

Robust standard errors in parentheses.

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

results in which I restrict the sample to specific populations of interest in order to directly control for unobserved heterogeneity associated with different forms of disadvantages. In each of these subgroup analyses, I use the four life stages EITC specification. In Table 9, I use the sample of children whose families lived in poverty two years before they were born, whose fathers were not present in their household the year that they were born, and who attended Head Start Preschool. I then use samples of male, white, African American, and Hispanic children. In order to ensure accurate estimates within all models, I utilize a linear probability model, and then I estimate a corresponding logit model to verify that there were no substantial changes from the ordinary least squares estimates in sign or significance.

Intuitively, the most likely explanation for a negative relationship between EITC receipt and graduating from high school in a naïve model is omitted variable bias; a program designed to provide cash assistance potentially worth thousands of dollars to parents would likely not causally reduce their children's educational attainment. Therefore, when displaying estimates for every results table of the full sample and the subgroup analysis, I construct four models in which a progressively larger number of covariates are added. The first stage in each specification includes a basic regression of the independent variable of interest on the likelihood of on-time high school graduation. In the second model, basic income variables are added to control for family income during the specified period. The third model includes a vector of primarily maternal characteristics, including labor force participation, maternal background, family race, and experience during pregnancy. The most complex model adds child-specific measures such as family structure during the child's life, measures of cognitive ability, behavior

Table 7. Full Sample: Two Life Stage Specification

Full Sample	(1)	(2)	(3)	(4)
VARIABLES	On-Time Graduate	On-Time Graduate	On-Time Graduate	On-Time Graduate
Average EITC	-0.0326***	0.00561	0.00813	0.0170*
Late Childhood	(0.00862)	(0.00878)	(0.00885)	(0.00958)
Average EITC	-0.00562	0.0478**		0.0643***
Early Childhood	(0.0197)	(0.0198)	0.0157 (0.0204)	(0.0231)
Family Income		X	X	X
Maternal Characteristics			X	X
Child Characteristics				X
Constant	0.507***	0.305***	0.570***	0.559***
	(0.0106)	(0.0160)	(0.0619)	(0.0669)
Observations	4,015	4,015	4,015	4,015
R-squared	0.005	0.069	0.142	0.322

Robust standard errors in parentheses.

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

problems, region of birth, childbearing, and neighborhood characteristics.

The basic bivariate regression in the first model of Table 6 follows the intuition that receiving a larger average benefit over the course of a child's life is highly statistically significantly associated with a decreased likelihood of graduating from high school due to the fact that eligibility is based on disadvantage. In the full model containing all control variables shown in Table 6, the coefficient of 0.0532 indicates that a $1,000 increase in the average EITC benefit received over the course of a child's life is associated with a 5.32 percentage point increase in the likelihood of a child graduating from high school on time.[1] With the baseline graduation rate within the overall sample at 47.82 percent, a $1,000 increase in average EITC, holding all else constant, is associated with an 11.1 percent increase in graduation likelihood.

The full model in Table 7 shows a positive coefficient on early childhood EITC receipt of 0.0643 that is highly statistically significant at the p <0.01 level. This indicates that a $1,000 increase (in 2008 dollars) in the average EITC received during the period of birth through age 9 is associated with a 6.43 percentage point increase in on-time high school completion.[2] The coefficient on EITC receipt in late childhood is statistically

[1] A $1,000 increase in average real EITC over a child's life amounts to a 145.3 percent increase. This is calculated using average real EITC receipt conditional on ever receiving the credit, shown in Table 2.

[2] A $1,000 increase in average real EITC during early childhood amounts to a 236.4 percent increase. This is calculated using average real EITC receipt conditional on ever receiving the credit, shown in Table 2.

Table 8. Full Sample: Four Life Stage EITC Specification

VARIABLES	(1) On-Time Graduate	(2) On-Time Graduate	(3) On-Time Graduate	(4) On-Time Graduate
Average EITC in High School	-0.0315***	-0.00142	0.000201	0.00564
	(0.00715)	(0.00733)	(0.00722)	(0.00706)
Average EITC in Middle School	0.00206	0.00957	0.0105	0.0156*
	(0.00916)	(0.00914)	(0.00903)	(0.00863)
Average EITC in Elementary School	-0.0218*	-0.00659	-0.0156	0.0122
	(0.0131)	(0.0132)	(0.0130)	(0.0122)
Average EITC Before School Years	0.0262	0.0868***	0.0590**	0.0680**
	(0.0278)	(0.0281)	(0.0278)	(0.0288)
Family Income		X	X	X
Maternal Characteristics			X	X
Child Characteristics				X
Constant	0.508***	0.302***	0.568***	0.562***
	(0.0107)	(0.0164)	(0.0621)	(0.0669)
Observations	4,015	4,015	4,015	4,015
R-squared	0.008	0.070	0.143	0.322

Robust standard errors in parentheses.

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

significant at the p<0.1 level, but it is smaller in magnitude at 0.017.

Table 8 displays the results of allowing the relationship between average EITC receipt and high school graduation to vary over four life stages.[3] The full model in Table 8 indicates that, holding all else constant, a $1,000 increase in average real annual EITC in years prior to age 6 is associated with a 6.80 percentage point increase in the likelihood that a child will complete high school on time, and the same increase during middle school years is associated with a 1.56 percentage point increase.[4] The coefficient estimates on EITC receipt during high school years and elementary school years are not statistically significant in the full model.

[3] Additional regressions were estimated using different specifications for the years that made up each life stage, e.g. middle school representing ages 11-13 and elementary school covering ages 5-10. There are some marginal changes to the magnitude of the coefficients for each life stage, but there are no substantive changes in significance and no changes in the sign of any life stage coefficient.

[4] A $1,000 increase in average real EITC before a child enters school amounts to a 339.2 percent increase, and a $1,000 increase in average real EITC during middle school amounts to a 105.5 percent increase. This is calculated using average real EITC receipt conditional on ever receiving the credit, shown in Table 2.

Table 9. Sub-Sample Analyses: Four Life Stage EITC Specification							
	1	2	3	4	5	6	7
VARIABLES	Poverty Pre-Birth	Father Not Present	Head Start	Male	White	Black	Hispanic
Average EITC in High School	0.00290 (0.00947)	-0.00419 (0.0105)	-0.00929 (0.0136)	0.00801 (0.0101)	0.0133 (0.0126)	0.00169 (0.0111)	0.00645 (0.0146)
Average EITC in Middle School	0.0194* (0.0114)	0.0247* (0.0135)	0.00902 (0.0162)	-0.00373 (0.0118)	0.0152 (0.0156)	0.0207 (0.0139)	0.0114 (0.0166)
Average EITC in Elementary School	0.0176 (0.0169)	-0.0156 (0.0208)	0.000477 (0.0222)	0.0105 (0.0164)	0.0577** (0.0201)*	-0.0224 (0.0203)	0.0162 (0.0252)
Average EITC Before School Years	0.0750* -0.043	0.114** (0.0469)	0.0234 (0.0579)	0.112*** (0.0400)	0.0232 (0.0502)	0.0829* (0.0466)	0.119* (0.0629)
Family Income	X	X	X	X	X	X	X
Maternal Characteristics	X	X	X	X	X	X	X
Child Characteristics	X	X	X	X	X	X	X
Constant	0.569*** (0.103)	0.508*** (0.113)	0.540*** (0.155)	0.470*** (0.0952)	0.609*** (0.128)	0.712*** (0.144)	0.419*** (0.147)
Observations	1,683	1,532	890	2,029	1,751	1,353	911
R-squared	0.321	0.319	0.358	0.319	0.299	0.387	0.359

Robust standard errors in parentheses.

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

Coefficient estimates of the relationship between average EITC receipt and high school graduation for the 1,683 child poverty before birth subsample in Table 9 are slightly larger in magnitude than the results from the full sample. A $1,000 increase in average EITC during middle school is associated with a 1.94 percentage point increase in graduation likelihood, and the same increase in EITC prior to entering school is associated with a 7.50 percentage point increase. This is a 24.4 and 10.3 percent increase, respectively, over the baseline relationship established in the full sample. Limiting the sample to children whose fathers did not live in the child's household during the first year of life, I obtain coefficient estimates for average EITC received during middle school and before entering school that are larger in magnitude than those in the full sample but at roughly the same significance levels. A $1,000 increase in average EITC is associated with an 11.4 percentage point increase in graduation likelihood when it occurs before school years, and the increase in EITC is associated with a 2.47 percentage point increase in on-time graduation when it occurs during middle school. Among this more disadvantaged subgroup, the rate of on-time high school completion is 46.7 percent. This increase of 11.4 percentage points amounts to a 24.4 percent increase in the likelihood of graduation. Only the estimates with fewer covariates yield statistically significant results for the Head Start subsample. This may be due in part to the small sample size of this subgroup; the 890 Head Start attendees are just more than half the number of children whose families experienced poverty prior to their birth.

Focusing on the male subsample in Table 9, I estimate that a $1,000 increase in the average real EITC is associated with an 11.2 percentage point increase in the likelihood of graduating on time. Stratifying the sample by race and obtaining estimates from the full model yields different results for each subgroup. Among white children, a $1,000 increase in the EITC during elementary school is associated with a 5.8 percentage point increase in on-time graduation. This highly statistically significant ($p<0.01$) result is the first incidence in which EITC receipt during elementary school had a measurable relationship with graduation. Among African Americans the relationship more closely mirrors disadvantaged subgroups, as a $1,000 increase in EITC receipt prior to attending school is associated ($p<0.10$) with an 8.3 percentage point increase in on-time graduation. Results indicate that Hispanic children have a similarly statistically significant coefficient on EITC receipt before school years ($p<0.10$), but the magnitude is larger than any other significant estimate in the full sample or subgroup analysis. Among Hispanic children, a $1,000 increase in average EITC is associated with an 11.9 percentage point increase in the likelihood of graduation.

VII. DISCUSSION

The results of this study support the hypothesis that receipt of the Earned Income Tax Credit is associated with an increased likelihood of graduating from high school on time. I obtain robust results from a series of regressions in a linear probability model because the economic and demographic information available in the NLSY allowed me to include a rich set of covariates from the extensive personal history of each child and his or her family. The relationship between EITC receipt and on-time graduation is strongest when receipt occurs during early childhood, specifically before children enter school, and when the recipients are in more disadvantaged subgroups, including African Americans, Hispanics, children whose families lived in poverty before their birth, and children whose fathers were not living with them at birth. This analysis also indicates that increases in the real value of EITC receipt during middle school years are associated with an increased likelihood of graduation, but the magnitude of that relationship is smaller than during early childhood. The change in sign associated with the inclusion of a range of covariates supports the idea that the naïve bivariate correlation between EITC receipt and on-time graduation is influenced by significant omitted variable bias. Broadly speaking, these results contribute to the growing literature on the relationship between family economic interventions and school outcomes.

“… this analysis indicates that our safety net may have a greater ability than previously thought to address some of our nation’s most persistent drivers of poverty and inequality.”

Significance in the early childhood and middle school stages of children’s lives substantively supports the existing literature because of the rich set of covariates utilized in the models and based on the fact that results are robust when controlling for different forms of unmeasured heterogeneity in the subgroup analyses. The results of this study support the hypothesis that increased income is most beneficial to long-term educational outcomes during early childhood years (Haveman and Wolfe 1995; Duncan et al. 1998; Levy and Duncan 2000). The estimates indicate that a $1,000 increase in average EITC receipt in the years before a child enters school is associated with an increase in on-time graduation of 6.80 percentage points in the full sample, 7.50 percentage points among children whose families lived in poverty before their birth, and 11.4 percentage points in the group whose fathers were not present at birth.

The approach of using life stages when measuring EITC receipt also supports recommendations from practitioners in the dropout prevention field. The National High School Center and the Annie E. Casey Foundation both cite the eighth and ninth grade transition point into high school as a crucial intervention point in reducing the dropout rate (Kennelly and Monrad 2007; Shore and Shore 2009). The estimates indicate that a $1,000 EITC receipt during middle school is associated with an increase in on-time high school completion of 1.56 percentage points for the full sample, 1.94 percentage points in the family poverty sample, and 2.47 percentage points in the sample without fathers present.

Substantively, the analysis focuses on identifying the relationship between graduation and EITC receipt during children’s life stages. Table 2 indicates that a $1,000 increase in EITC benefits amounts to just below one standard deviation in average EITC receipt during middle school, but it represents a roughly three-standard-deviations increase in EITC receipt before the child entered school. However, the real value of the EITC has grown significantly with statutory changes. Although a $1,000 increase in average EITC receipt in the years before a child enters school would represent a more than three-fold increase, a family receiving the maximum credit possible saw that credit increase more than $3,000 in real terms from 1979 to 2008. Understanding the substantive implications of a $1,000 increase in average EITC is important within each life stage, but when comparing results from different life stages it may be more important to look at the significance of the coefficients than at the magnitude.

Careful cost-benefit analysis should be conducted to determine whether increases in real value of the EITC are economically beneficial. The long-term social costs associated with higher dropout rates suggest that increases in EITC spending could be offset by reductions in other public expenditures in areas such as criminal justice, public benefits, and health care (Alliance for Excellent Education 2011). Similarly, the adverse labor market consequences associated with failing to complete high school appear to be increasing, indicating that the costs of EITC expansion could be offset by increases in economic output and tax revenue (Holzer et al. 2007). However, if increasing the EITC is to be used as an anti-dropout policy tool, more careful analysis should be conducted to determine whether statutory changes should involve a change to the real value of the credit as took place in 1986, or a change to the real value for parents with a certain number of children as in 1993 and 2009. Non-statutory changes, such as outreach and education for low income families about the importance of filing tax returns or assistance in properly claiming the credit, could also be viable approaches for entities that seek to improve outcomes for families through interventions outside of the school. Beyond specific programmatic implications, this analysis indicates that our safety net may have a greater ability than previously thought to address some of our nation's most persistent drivers of poverty and inequality.

VIII. REFERENCES

Alliance for Excellent Education. 2011. The high cost of high school dropouts: What the nation pays for inadequate high schools. Issue Brief, October. www.all4ed.org/files/archive/publications/HighCost.pdf.

Amos, Jason. 2008. Dropouts, diplomas, and dollars: US high schools and the nation's econothe. Alliance for Excellent Education, August. http://www.all4ed.org/files/Econ2008.pdf.

Bureau of Labor Statistics. 2008. The NLSY79 sample: An introduction. http://www.nlsinfo.org/nlsy79/docs/79html/79text/79sample/introsample_1_a.html.

Bureau of Labor Statistics. 2011. Education pays... US Department of Labor. http://www.bls. gov/emp/ep_chart_001.htm.

Center on Budget and Policy Priorities. 2011. Policy basics: The Earned Income Tax Credit. http://www.cbpp.org/cms/index.cfm?fa=view&id=2505.

Chapman, Chris, Jennifer Laird, and Angelina KewalRamani. 2010. Trends in high school dropout and completion rates in the United States: 1972–2008. US Department of Education, National Center for Education Statistics, Compendium Report.

Duncan, Greg, W. Jean Yeung, Jeanne Brooks-Gunn and Judith Smith. 1998. How much does childhood poverty affect the life chances of children? *American Sociological Review* 63, no. 3.

Dahl, Gordon, and Lance Lochner. 2011. The impact of familyincome on child achievement: Evidence from the Earned Income Tax Credit. Working paper no. 14599, National Bureau of Economic Research.

Eamon, Mary Keegan, and Chi-Fang Wu. 2009. Effectiveness and limitations of the Earned Income Tax Credit for reducing child poverty in the United States. *Children and Youth Services Review* 31: 919-926.

Eissa, Nada, and Hilary Hoynes. 1998. The hours of work response of married couples: Taxes and the Earned Income Tax Credit. National Bureau of Economic Research.

Eissa, Nada, and Jefferey Liebman. 1996. Labor supply response to the Earned Income Tax Credit. *The Quarterly Journal of Economics* 111.

Eissa, Nada, and Austin Nichols. 2005. Tax-transfer policy and labor market outcomes. Urban Institute Publications.

Evans, William, and Craig Garthwaite. 2011. Giving mom a break: The impact of higher EITC payments on maternal health. Russell

Sage Foundation.

Feenberg, Daniel, and Elizabeth Coutts. 1993. An introduction to the TAXSIM Model. *Journal of Policy Analysis and Management* 12.

Goodman-Bacon, Andrew and Leslie McGranahan. 2008. How do EITC recipients spend their refunds. *Economic Perspectives* 2Q.

Haveman, Robert, and Barbara Wolfe. 1995. The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature* 33.

Holzer, Harry, Diane Whitmore Schanzenbach, Greg J. Duncan, and Jens Ludwig. 2007. The economic costs of poverty in the United States: Subsequent effects of growing up poor. Center for American Progress.

Hotz, Joseph, and John Karl Scholz. 2002. The Earned Income Tax Credit. In *Means tested transfer programs in the United States*, edited by Robert Moffitt, 141–198. Chicago: University of Chicago Press.

Hotz, Joseph and John Karl Scholz. 2005. Examining the effect of the Earned Income Tax Credit on the labor market participation of families on welfare. Working Paper #11968, National Bureau of Economic Research.

Hoynes, Hilary. 2008. The Earned Income Tax Credit, welfare reform, and the employment of low-skilled single mothers. Revised draft of paper prepared for Chicago Federal Reserve Bank Conference, November 2007.

Isaacs, Julia, Eugene Steuerle, Stephanie Rennane, and Jennifer Macomber. 2011. Kids share: Report on federal expenditures on children through 2010. The Brookings Institution, The Urban Institute, http://www.brookings.edu/research/reports/2010/07/14-kids-share-isaacs.

Internal Revenue Service. 2011. EITC statistics. US Department of the Treasury, http://www.eitc.irs.gov/central/eitcstats/.

Kennelly, Louise, and Maggie Monrad. 2007. Approaches to dropout prevention: Heeding early warning signs with appropriate interventions. National High School Center.

Levy, Dan, and Greg Duncan. 2000. Using sibling samples to assess the effect of childhood family income on completed school. Working paper, Institute for Poverty Research.

Mendenhall, Ruby, Kathryn Edin, Susan Crowley, Jennifer Sykes, Laura Tach, Katrin Kriz, and Jeffrey Kling. 2012. The role of the Earned Income Tax Credit in the budgets of low-income families. *Social Service Review*.

Meyer, Bruce, and Dan Rosenbaum. 2001. Welfare, the Earned Income Tax Credit and the labor supply of single mothers. *The Quarterly Journal of Economics* 116.

Moore, Kristen Anderson, Zakia Redd, Mary Burkhauser, Kassim Mbwana, and Ashleigh Collins. 2009. Children in poverty: Trends, consequences, and policy options. *Child Trends* 2009-11.

Morris, Pamela, Greg Duncan, and Christopher Rodrigues. 2004. Does money really matter? Estimating impacts of family income on children's achievement with data from random-assignment experiments. Next Generation Project, MDRC Evaluation.

Rumberger, Russel. 2004. Why students drop out of school. In *Dropouts in America: Confronting the graduation rate crisis*, edited by Gary Orfied, 131–155. Cambridge, MA: Harvard Education Press.

Rumberger, Russel. 2004. What can be done to reduce school dropouts? In *Dropouts in America: Confronting the graduation rate crisis*, edited by Gary Orfied, 243–254. Cambridge, MA: Harvard Education Press.

Rumberger, Russel, and Sun Ah Lim. 2008. Why students drop out of school: A review of 25 years of research. Policy Brief 15, California Dropout Research Project.

Sellers, Kacy. 2011. National graduation rate rebounds; 1.2 million sudents still fail to earn diplomas. Editorial Projects in Education Research Center.

Shore, Rima, and Barbara Shore. 2009. Kids count indicator brief: Reducing the high school dropout rate. Annie E. Casey Foundation.

Smeeding, Timothy, Katherine Ross Phillips, and Michael O'Connor. 1999. The EITC: Expectation, knowledge, use, and economic and social mobility. Paper 147, The Center for Policy Research.

Yeung, Jean, Miriam Linver, Jeanne Brooks-Gunn. 2002. How money matters for young children's development: Parental investment and family processes. Child Development 73, no. 6.

IX. APPENDIX

Table 3. Background of Mothers in Study Sample

Demographic Characteristic		Number of Cases	Percentage of sample
	Total Children:	4,015	100
	Total Mothers:	2,464	100
Race:	White	1,106	44.9
	Black	826	33.5
	Hispanic	532	21.6
Early Life Experiences:	Born in US	3,691	91.9
	Born Outside of US	324	8.1
	Spoke English at Home	2,972	74.1
	Spoke Another Language	1,043	24.9
	Non-South	2,480	61.7
	South	1,535	38.3
	Non-SMSA	1,157	28.8
	SMSA	2,858	71.2
	Central City (of SMSA)	868	21.6
Education:	Did Not Complete High School	1,035	25.7
	High School Grad or More	2,980	74.3
Pregnancy:	Had a Child as a Teen	354	8.8
	Had a Child >19 y/o	3,661	91.2
Marital Status:	Ever Married	3,438	85.6
	Never Married	577	14.4
	Ever Widowed or Divorced	1,861	46.4
	Never Widowed or Divorced	2,154	53.6

Race is measured as number of mothers in each group.

The unit of analysis in this and every other case is children whose mothers have these characteristics.

Table 4. Background of Children in Study Sample			
Demographic Characteristic		Number of Cases	Percentage of Sample
	Total:	4,015	100
Gender:	Male	2,029	50.5
	Female	1,986	49.5
Race:	White	1,751	53.1
	Black	1,353	27.7
	Hispanic	911	19.2
Region:	Born in Northeast	577	14.4
	Born in South	1,565	39
	Born in West	763	19
	Born in Midwest	984	24.6
Birth Outcomes	Born to Teen Mom	354	8.8
	Mom >19 y/o	3,661	91.2
	Received Prenatal Care	11,311	98.4
	Mom Drank During Pregnancy	4,897	42.6
	Mom Smoked During Pregnancy	3,759	32.7
	Born Underweight	1,058	9.2
	Born Healthy Weight	10,437	90.8
Parents 2 Years Before Birth:	In Poverty	1,683	41.9
	Receiving Public Benefits	709	17.6
Life Outcomes	Attended Head Start	890	22.2
	Lived in an Unsafe Neighborhood	986	24.6
	Had a Sibling who Dropped Out	1,653	41.2
	Had a Child as a Teen	663	16.5

Table 5. Summary Statistics of Continuous Variables for Mothers and Their Children

Demographic Characteristic		Mean	Median	Std. Dev.
Mothers	Number of Children	3.1	3	1.3
	Age at First Birth	20.8	20	3.2
	AFQT Percentile	32.9	27	25.7
	Highest Grade Completed	12.7	12	2.7
	Years Receiving Public Assistance	6.5	3	8.2
	Average Yearly AFDC/TANF	$441.80	$1,758.50	$2,634.80
	Average Yearly SNAP	$688.50	$1,371.80	$1,675.30
	Average Yearly SSI	$692.70	$ 0.00	$1,389.80
Children	PIAT Math Percentile	48.4	48.4	22.8
	PIAT Reading Percentile	54.1	54.1	26.1
	Behavioral Problems Index	61.8	63.8	22.3
	Highest Grade Completed	12.5	12	2.1

Average yearly public assistance is in 2008 real dollars and is calculated conditionally on a family ever receiving benefits.

Children's highest grade is taken from the highest reported level of educational attainment achieved irrespective of when it was achieved.

The Georgetown

PUBLIC POLICY REVIEW

Georgetown Public Policy Institute
37th and O Streets NW
100 Old North
Washington, DC 20057

E-mail: info@gppreview.com
Website: www.gppreview.com