FERTILITY AND WOMEN'S ECONOMIC OUTCOMES IN THE UNITED STATES, PERU AND SOUTH AFRICA

by

Tanya S. Byker

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2014

Doctoral Committee:

Associate Professor Martha J. Bailey, Co-Chair
Professor David A. Lam, Co-Chair
Professor Murray V. Leibbrandt, University of Cape Town
Professor Jeffrey Andrew Smith
Professor Pamela J. Smock

DEDICATION

This dissertation is dedicated to Eva, who started kindergarten the same day I started graduate school. To Bas, who set early versions of this research to verse. To Ian, who has been infinitely supportive. To all three of them, who joined me on an adventure to South Africa. To Susan, who helped in far too many ways to list. To Elsie, who nourished us along the way. To Bern, Gaylen, Gayle, Joel, Rebecca, Kathy, John, and Henrietta who inspired and supported me. It truly took a family to make this dissertation possible.

ACKNOWLEDGEMENTS

This research was generously supported by an NICHD training grant to the Population Studies Center at the University of Michigan (T32 HD007339) and by a William and Flora Hewlett Foundation/Institute of International Education Dissertation Fellowship in Population, Reproductive Health and Economic Development.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

The Role of Paid Parental Leave in Reducing Women's Career Interruptions: Evidence from Paid Leave Laws in California and New Jersey

Abstract: I analyze the effects of paid parental leave on maternal labor supply. Using monthly longitudinal data from the Survey of Income and Program Participation, my event-study research design characterizes the evolution of labor-force participation around childbirth for women affected by paid leave laws in California and New Jersey. I find that paid leave laws are associated with a substantial increase in labor-force participation in the months directly around birth but have little impact beyond six months after birth. While US-style short-duration leave is unlikely to change prolonged exits from the labor force, my findings imply that paid leave laws induce some women to stay more attached, particularly low-skill women.

Acknowledgements:
I am grateful to Martha Bailey, Charlie Brown, Jamein Cunningham, Nic Duquette, Anne Fitzpatrick, Robert Garlick, Andrew Goodman-Bacon, David Lam, Sayeh Nikpay, Matt Rutledge, Pamela Smock and Jeff Smith for helpful comments and guidance.

1.1 Introduction

The influx of mothers with young children into the labor market has been a major driver of the steep rise in women's labor-force participation since the 1960s. Participation among married women with children under six rose from just 20 percent in 1960, to over 60 percent by 1990 (U.S. Census Bureau, 1995, 2012). The resulting increase in women's years of work experience contributed substantially to the narrowing of the gender gap in earnings (Blau and Kahn, 1997, O'Neill and Polachek, 1993). But as mothers' attachment to the labor force has grown, the tension between home production and market work has increased. Women remain far more likely to experience career interruptions during childbearing years than men, and women's actual labor-force experience remains 20 percent below men's (Bertrand et al., 2010, Blau and Kahn, 2004, Byker, 2012).

Parental leave laws are a common policy approach to mitigating the negative effects of childbearing on women's careers. Parental leave allows women to spend time away from work with a child while maintaining attachment to their employer through mandated job protection. In countries where paid leave is lengthy and generous—the OECD average is 80 weeks of leave, with 33 weeks of full-time equivalent paid leave (OECD, 2012)—these laws have been shown to increase women's labor-force participation. However, they also increase time away from work and may raise the cost of hiring women due to lost productivity during absences or costs of finding and training temporary replacements. There is mixed evidence on whether these factors suppress wage growth and hinder women's career advancement (Blau and Kahn, 2013, Lalive and Zweimüller, 2009, Ruhm, 1998).

In contrast to other industrialized economies, leave policies in the US are short and typically unpaid. The 1993 Family and Medical Leave Act (FMLA) guarantees 12 weeks of unpaid leave for qualified employees. Studies of FMLA find little evidence of an impact of the law on women's work outcomes (Baum, 2003a, Han et al., 2009, Waldfogel, 1999). Recently, several US states have expanded leave benefits beyond the federal mandate by offering short-duration paid leave to parents of newborn children—California's 2004 law mandates 55 percent wage replacement for leave of at most 6 weeks. Evaluations of the California law using the Current Population Survey (CPS) find evidence of increased use of parental leave, but are not able to provide conclusive evidence on how leave-taking impacts women's attachment to the labor force (Espinola-Arrendondo and Mondal, 2010, Rossin-Slater et al., 2013).

The challenge of estimating the impact of US parental leave policies stems from the laws' narrow window of eligibility and the likelihood that effects may be very different in the months immediately around birth than in the longer term. Section 1.3 of this paper presents a simple model of labor supply which shows how paid leave laws may affect women who take brief labor-force exits after giving birth but should have less impact on women who take prolonged exits. Because data limitations of previous studies make it difficult to disentangle these time-varying effects, this paper uses longitudinal data and a flexible estimation strategy to examine both the short and longer-term impacts of paid leave laws in California (2004) and New Jersey (2009).

The Survey of Income and Program Participation (SIPP) allows a fuller characterization of women's labor-force participation from 24 months before to 24 months after they give birth. Using an event-study methodology, I measure women's career interruptions on a month-to-month basis relative to their own labor-force participation prior to giving birth. Triple-difference event-study estimates show the impact of paid leave on career interruptions by comparing (1) birth-related changes in labor-force participation in California and New Jersey (2) before and after enactment of paid leave laws with (3) birth-related changes for women in states without paid leave policies.

The results show that California and New Jersey paid leave laws substantially reduce the incidence of short-term exits from the labor force lasting less than six months—women's labor-force participation increases by five to 10 percentage points in the three months before to three months after a birth. However, paid leave laws have little impact on sustained exits lasting six or more months after birth. Testing for heterogeneous effects, I find that the reduction in short-term interruptions occur almost exclusively among women with less than a college education. This is consistent with the fact that voluntarily-provided paid leave disproportionately benefited college-educated women before paid leave mandates were implemented. Paid leave laws extend these benefits to women lower in the skill distribution, allowing them to choose career interruption patterns that more closely resemble those of more-educated women.

Opponents claim that paid leave laws, which are funded through employee payroll taxes, are costly mandates that burden firms with inconvenience costs. In 2012, parents of nearly 40 percent of all children born in California filed paid leave claims, at a cost of $470 million. New Jersey paid $63 million for claims associated with nearly a quarter of all births. The potential benefits to households in terms of child health and wellbeing are difficult to quantify and the evidence is mixed (Baum, 2003b, Berger et al., 2005, Carneiro et al., 2011, Dunifon et al., 2013, Rossin, 2011, Ruhm, 2004). Advocates claim that paid leave also increases mothers' attachment to the labor force, but evidence has been lacking. While I find that short-duration leave of the type mandated in California and New Jersey is unlikely to change the behavior of women who take prolonged exits from the labor force, my results demonstrate that paid leave laws may induce some women to stay more attached to their jobs, in particular low-skill women. Increasing these women's attachment to the labor force has the potential to increase tenure, accumulated experience and long-term earnings growth, which suggests that paid leave policies have the potential to reduce the gender gap in earnings and earnings inequality across the education spectrum.

1.2 Parental Leave in the United States

The history of parental leave in the United States is brief as the first federal leave legislation was not passed until 1993. This section describes the policy landscape prior to the introduction of the nation's first paid leave laws in California and New Jersey in the mid-2000s, describes trends

1.2.A Unpaid Parental Leave

The Family and Medical Leave Act of 1993 (FMLA) was the first bill signed into law by President Bill Clinton. The bill was a major part of Clinton's first term agenda to enact policies that "both increased employment and strengthened families" (Clinton, 2013).

FMLA mandates that eligible employees, working for covered employers, receive up to 12 weeks of unpaid, job-protected leave within a 12-month period for specified family and medical reasons including the birth of a child. Upon return from leave, workers are guaranteed their original or an equivalent job with equivalent pay and benefits.[1] Coverage applies to public and private employers with at least 50 employees within 75 miles of the worksite. Workers are eligible if they have worked for a covered employer for at least 12 months and at least 1250 hours over the last 12 months (at least 25 hours on average for 50 weeks). The Department of Labor estimates that in 2012, 59 percent of US workers were both covered and eligible, and that 16 percent of those workers had taken an FMLA leave in that year (Klerman et al., 2012). Prior to FMLA, 13 states required some form of unpaid parental leave, and currently 18 states offer unpaid leaves with less restrictive coverage and/or eligibility restrictions, and in a few cases

[1] Health benefits are maintained under the same terms as if the employee had continued work.

slightly longer durations (Baum, 2003a, National Partnership for Women and Families, 2012).

In 2002 California passed the first law mandating paid parental leave in the United States. California is among five states with long-standing temporary disability insurance (TDI) programs that include pregnancy as an eligible "disability" for leave with partial wage replacement.[2] These TDI states have so far been the most likely to pass paid family leave legislation, which is facilitated by established political support, and administrative mechanisms for collecting taxes and making payments for this type of program. Two other TDI states have also passed paid parental leave laws -- New Jersey in 2008 and Rhode Island in 2013.[3]

1.2.B Trends in Voluntary Provision of Paid Leave Benefits

While there is no federal mandate for paid leave, the number of firms that voluntarily offer paid leave has increased substantially since the 1980s especially for more-educated workers. Figure 1.1.1 shows trends in usage of paid parental leave in the 12 weeks after a first birth among women who worked during pregnancy as reported in SIPP retrospective fertility modules.[4] In the early 1980s, less than 20% of all women reported using paid parental leave, with similar rates for women with and without a bachelor's degree. By the late 2000s, 44 percent of women with at least a bachelor's degree report using paid parental leave compared to only 26 percent among women with less than a college degree. Figure 1.1.2 includes other types of paid leave (including paid sick leave, vacation leave and other leave), and indicates that while use of any paid leave increased for both groups, more-educated women are far more likely to use paid leave

[2] The TDI states are California, Hawaii, New Jersey, New York, and Rhode Island.
[3] Washington State passed a paid leave law in 2007 providing a flat $250 benefit for up to five weeks, but implementation has been delayed while lawmakers seek funding for the program. Several other states have proposals for paid leave laws and Senators Dodd and Stevens proposed the federal Family Leave Insurance Act in 2007: http://www.apwu.org/issues-fmla/stmnt-dodd.pdf
[4] The data come from topical modules in the 1996, 2001, 2004 and 2008 SIPP panels. The survey question refers to "maternity" leave.

after a birth – 60 compared to 30 percent of less-educated women.

These findings of greater usage of paid leave among higher-skilled women from the SIPP concur with evidence on access to paid leave from the Bureau of Labor Statistics National Compensation Surveys. In 1993, only two percent of private employees had access to leave that was specifically designated as family leave, though most had access to paid vacation days and half had some paid sick days (Van Giezen, 2013). By 2012 the overall rate of access to paid family leave increased to 11 percent, and the rate in managerial and professional occupations was around 20 percent (around 30 percent in the financial and business sectors) while for service occupations the rate of access was only six percent (Bureau of Labor Statistics, 2012). Unequal access to paid leave was one of the drivers of new paid-leave legislation in California and New Jersey. These differences in access prior to implementation of the laws also suggest investigating whether there are heterogeneous impacts by education.

1.2.C Details of Paid Leave laws in California and New Jersey

California's Paid Family Leave (C-PFL) law was passed in September 2002 and went into effect in July 2004. The New Jersey Family Leave Insurance (NJ-FLI) law was passed in April 2008 and took effect July 2009. The California and New Jersey policies are similar in most respects. Both laws provide partial wage replacement for up to six weeks for time spent caring for sick family members or to "bond" with a newborn or an adopted child. C-PFL provides 55 percent of wages up to \$1067 per week, and NJ-FLI provides 66 percent of wages up to \$584 per week.[5] In both states, the six weeks of paid family leave extend existing temporary disability leave of ten weeks for a normal pregnancy – four before birth and six after -- under the same

[5] These are maximum weekly payments for 2013. The California cap is indexed to the state's average weekly wage. Unlike SDI benefits, paid family leave benefits are not taxed at the state level, but are still subject to federal taxes.

replacement rates.[6,7] In both California and New Jersey paid parental leave benefits are financed entirely by mandatory payroll taxes levied on all private employees, which were projected to amount to an additional .08% tax in California and .09% in New Jersey over existing deductions for TDI.[8,9]

While both laws have minimal eligibility and coverage restrictions, neither grants new rights to take a leave. Importantly, neither law provides job protection or continuation of fringe benefits. The wage replacement provided under the laws is to be used concurrently with leave granted either under FMLA or an employer's voluntary program. Therefore, these laws do not directly expand parental-leave access to previously un-served populations.

Both laws were enacted at least a year before they were implemented, and there was media coverage of the laws at the time of passage and when they went into effect. Employers are required to provide information about the laws to all new employees and anyone requesting leave. Surveys have shown that not all eligible worker are aware of C-PFL (Appelbaum and Milkman, 2011), but administrative records show a substantial and immediate increase in claims filed for bonding with newborns in both states. I estimate that in 2004 the number of claims was roughly equal to one third of all births occurring in California among women who were employed a year before birth; I estimate a similar rate in New Jersey in 2010, the first full year the law was in effect. By 2010, six years after implementation, the number of claims in California was roughly equal to half of the number of births among employed women in the

[6] There are possible extensions of disability insurance for cesarean sections or birth complications.
[7] New Jersey Law: http://lwd.dol.state.nj.us/labor/fli/fliindex.html California Law: http://www.edd.ca.gov/disability/paid_family_leave.htm
[8] The self-employed and state and local employees can opt in to the programs. Employers may substitute with their own program if it is at least as generous as the state mandate. In New Jersey non-family-leave disability insurance is also partially funded by employer contributions.
[9] Tax incidence projections: California: Rodriguez (2004), New Jersey: Deak (2008).

state.[10] In California in 2012 the average claim duration was 5.35 weeks with an average weekly benefit amount of $497; in New Jersey in 2011 the average weekly benefit was $489 and duration 5.3 weeks. Summary information on claims in California and New Jersey is show in Table 1.1.

Unlike TDI benefits for pregnancy, the new family leave laws apply to both mothers and fathers. However, the vast majority of bonding claims are made by women – 70 percent of claims in California, and 89 percent in New Jersey in 2012.[11]

1.3 Expected Effects of Paid Leave Laws

The stated goal of parental leave laws is to help women balance commitments to both job and family out of concern that in the absence of a mandate "too many" women are "forced" to choose one or the other (Senator Buono, 2008).[12] In the language of economics, paid leave laws expand the budget set to make new choices available. Opponents claim that paid leave is an entitlement transfer that pays women to work less.[13] But modeling the incentives created by the policies leads to an unexpected prediction – paid leave may lead some women to work *more* by pulling them into the labor force when they would otherwise have exited after a birth. Arriving at this prediction requires making distinctions in labor force status that are uniquely important for

[10] These estimates are based on National Center for Health Statistics state-level natality records by state, claims as reported by the California Employment Development Department and the New Jersey Department of Labor and Workforce Development, and the author's estimates in the SIPP of labor-force participation rates among women who give birth in each state.

[11] The number of fathers making claims in California has increased from 17 percent of claims in 2004 to 30 percent of claims in 2012. However, statistics are not available on the duration of leaves by gender, so that men may be taking a lower percentage of total leave weeks than this trend implies.

[12] On March 3, 2008, the day the New Jersey bill was approved by the state senate, Senator Barbara Buono, co-sponsor of the bill stated "Far too many individuals have had to make the tough decision between caring for a loved one and being able to maintain their income…These are choices that should never have to be made in the United States, and soon they will no longer be choices forced upon New Jersey families."

[13] This claim is bolstered by a recent working paper by Dahl et al. (2013) who find that paid maternity leave reforms in Denmark had little effect on a wide variety of outcomes including parental labor-force participation. However the Danish policy context is very different from US, and the authors point out "it is important to distinguish between the introduction of paid leave (and job protection) versus the continual expansions to a program."

parents of young children.

Working is ordinarily synonymous with being in the labor force, but for mothers of newborns the distinction is important. Clearly a woman who separates from her employer to spend time at home with her child is both *out* of the labor force and *not* working; but a woman who maintains her attachment to her employer and takes job-protected parental leave is *in* the labor force even though she is *not* working. This is more than a semantic distinction.[14] Parental leave is a legal arrangement whereby a woman maintains benefits and is guaranteed the same wages when she returns to work. She also maintains her firm-specific human capital/tenure and avoids the search cost of looking for a new job. The woman who is not working because she exited the labor force forgoes this value by severing attachment to her employer.

I present a simple illustrative model that generates predictions for the impact of paid leave on three work-related outcomes: 1) weeks worked, 2) weeks of paid leave, and 3) weeks in the labor force. The model uses the structure of US policies and evidence on how labor-force exits affect wages to identify predictions for all three outcomes. The goal of my empirical strategy is to test the model's predictions for how paid leave effects labor-force participation which compliments and extends the existing literature that has studied impacts on leave-taking but has been unable to provide conclusive evidence for other work outcomes.

1.3.A Labor Supply Model and Predictions

Consider a woman choosing how much labor to supply in the longer-term (T) after the birth of a child. Utility is derived from consumption and time spent at home.[15] Individual preferences

[14] Klerman and Leibowitz (1994) have an excellent discussion of the work-employment distinction among mothers of young children.

[15] U = u(C,H), where the utility function is strictly convex and monotonic.

are based on a woman's relative taste for time spent with children compared to time spent at work. These tastes are driven by factors such as cultural norms, professional identity, and satisfaction from time spent with children. The budget constraint is determined by a woman's potential income from working (Y*), unearned income (μ) such as spousal income, and the parental-leave policy environment.

Figure 1.2.1 illustrates a woman's budget constraints under different parental leave policies. I assume every woman must take a minimum, medically-necessary amount of time to give birth and recover from delivery. Point A represents the amount she can consume if she takes no other time away from work after giving birth, $A = Y^* = wT + \mu$, where w is her wage prior to giving birth. I first consider a case in which she can spend as much unpaid time away from work as she desires and can return to the same job at the same wage, resulting in a budget constraint,

$$C = w(T - h) + \mu$$

where h is number of weeks spent at home. In Figure 1.2.1 this constraint is represented by ABT.

In fact, employers are unlikely to hold a woman's job, and FMLA mandates job protection for only 12 weeks. So it is more realistic to assume that if a woman chooses to be at home for longer than 12 weeks she must separate from her employer and exit the labor-force. I assume that exits are costly which causes a kink point in the budget constraint at 12 weeks.[16] Exit costs can arise due to depreciation of skills while away from work, wage penalties due to a perceived lack of commitment to work "signaled" by labor-force exit, or search costs of looking for a new job (Bertrand, Goldin and Katz, 2010, Hotchkiss and Pitts, 2007, Mincer and Polachek, 1974). I characterize the exit cost as a wage penalty, resulting in a shallower slope in the budget

[16] The model has elements similar to a "conceptual framework" described in Han, Ruhm and Waldfogel (2009).

constraint after 12 weeks and therefore a discontinuous drop in consumption

$$C = (w - 1\{h > 12\} * penalty)(T - h) + \mu$$

this new budget constraint under FMLA conditions is represented in Figure 1.2.1 as ACDBT. [17]

With the introduction of paid leave, I assume for simplicity that the mother is guaranteed full wage replacement for 12 weeks – the same duration as unpaid leave under FMLA.[18] A woman can now spend up to 12 weeks at home and maintain the same level of consumption as if she took no time off:

$$C = 1\{h \leq 12\} * (wT + \mu) + \ 1\{h > 12\} * (w - penalty) * (T - h).$$

This new budget constraint under paid leave is represented in Figure 1.2.1 as AEDBT.

To generate predictions for how the introduction of paid leave laws will affect 1) weeks at home (or conversely time at work $= T - h$), 2) weeks on leave and 3) weeks in the labor force, assume the budget constraint under FMLA prior to the introduction of paid leave, and consider three different women with preferences as described in Figure 1.2.2. Figures 1.3.1 –1. 3.4 illustrate the behavioral response of each type of woman to the introduction of paid leave benefits.

Case 1 (Figure 1.3.1): Prior to the introduction of paid leave, a woman with preferences described by U_1 takes no time off beyond the medically necessary minimum, optimally choosing the bundle of consumption and no time at home represented by point A. The introduction of paid

[17] A fixed cost will give the same predictions.
[18] Partial wage replacement and different durations yield similar predictions. This duration is similar to the California and New Jersey laws.

leave expands the budget set, and her utility is increased by choosing the previously unattainable point E. She increases time spent at home without separating from her employer—leave-taking increases, time spent with children increases, but labor-force participation is unchanged. A similar case arises for any woman who optimally chooses a point between A and C on the budget constraint in the absence of paid leave. If a woman optimally locates at point C prior to the law, she will relocate to E after the law, with no change in leave-taking, time spent with children or labor-force participation, but with higher income and utility compared to point C.

Case 2 (Figure 1.3.2): A woman with preferences described by U_2 has a higher relative taste for time spent with children than the previous case. Under FMLA she optimally chooses point B, staying at home for the full two years after a birth. The introduction of 12 weeks of paid leave does not alter her behavior. Figure 3.2 shows that it would require a much longer-duration paid-leave benefit to change the labor supply behavior of a woman who chooses a sustained time away from work after childbirth. Similar predictions arise for any woman who optimally locates far to the right of the kink. This behavior is also more likely to be optimal as spousal or other unearned income increases.

Case 3 (Figure 1.3.3): The third case is an example of a woman who, in the absence of paid leave, optimally locates to the right of the kink point, but to the left of the sustained interruption chosen in the previous example. Potentially due to tastes or lower spousal income, she chooses to return to work relatively quickly. She separates from her employer after a birth but returns to the labor force within four months, for example. With the introduction of a paid leave mandate, she attains higher utility at point E, choosing to maintain attachment to her employer and preventing an exit. The final result is that leave-taking increases and her labor-force participation increases. Meanwhile, time spent with children decreases since she returns to work more quickly under paid leave than she would if she exited and had to search for a new job.

The previous three cases described how different types of women behave in the same job situation; Figure 1.3.4 describes a different type of work setting. The previous cases assumed full knowledge of FMLA and that the mandate is perfectly enforced. If women are not aware of FMLA or employers refuse to honor the mandate (through implicit pressure or imposing penalties) the budget constraint prior to the CA and NJ laws will appear as in Figure 1.3.4 – a woman may take no time off and choose point A, but anytime at home requires separation from the employer (with the associated costs discussed above) so that the budget constraint jumps down to point H – AHBT. This final case is relevant given findings that some women are unaware of their rights under FMLA or have limited ability to bargain for those rights (Waldfogel 2001). This situation may be particularly true of low-skill workers. The introduction of a new law may raise awareness of parental leave and increase pressure on employers to grant leave, making the budget constraint AEDBT as under paid leave and shifting the optimal choice from point J to E. The result is increased leave-taking and increased labor-force attachment, and an ambiguous change in time spent with children (depending where she locates before the law).[19]

The predictions of the model guide an empirical test of my research question: does paid leave have an effect on interruptions in women's labor-force participation? In summary, it is possible

[19] This model does not account for general equilibrium effects such as shifts in labor supply due tax costs of the policy or shifts in labor demand due to increased "inconvenience costs" of hiring workers likely to use paid leave. Curtis, Hirsch and Schroeder (2013) study the impact of the California law on new hires in a general equilibrium framework of a workplace mandate that only benefits some workers, but whose cost is "nominally borne" by all employees in the form of a payroll tax. Using data from the Quarterly Workforce Indictors, they find that wages of new hires who are most at risk of using paid leave (young women) are 2 percent lower than new hires among groups unlikely to use paid leave (young men and older women) and that employment among new hires in the at-risk group are relatively higher by 1.5 percent. The results imply a small outward shift in the labor supply of women most likely to use paid leave in response to the law, and that their valuation of the benefit offsets the policy's tax cost. Das and Polachek (2014), on the other hand, conclude using the CPS that, while participation of young women increases due to the law, their rate and duration of unemployment increases. The paper hypothesizes this is due to employer aversion to hiring workers who may take paid leave.

for leave-taking to increase but have no impact on labor-force participation if, as in Case 1, all women who are induced to take leave would have stayed with their employer anyway. Among women who exit the labor force in the absence of paid leave, the model predicts that only women who take relatively brief exits will be induced to change their behavior as in Case 3. And because women who take extended exits are unlikely to be affected by the law, any impacts on labor-force participation should be concentrated in a brief period of months around births while I should not expect impacts on labor-force participation in the longer-term. An empirical test of the impact of short-duration paid leave laws on labor-force participation thus requires an accurate measure of birth timing and a strategy to disentangle these time-varying effects.

1.3.B Which Hypotheses Does Existing Literature on Paid Leave Test?

While there is an extensive literature on the labor market impacts of the ubiquitous and generous parental leave policies in Europe, paid leave mandates have only recently appeared in the US, and this paper contributes to the nascent literature analyzing their effects. As the previous subsection demonstrates, the introduction of short-duration paid leave will likely have different effects on labor-market outcomes than the expansions of already-lengthy leaves that are studied in the European literature (ex. Dahl et al., 2013, Lalive and Zweimüller, 2009).

Rossin-Slater, Ruhm and Waldfogel (2012) study the effects of California's paid leave law using cross-sectional data from the March CPS within a differences-in-differences specification.[20] The paper provides robust evidence that the law increased overall leave-taking by mothers with children under the age of one in California, with the effects concentrated among less-educated women, and suggestive evidence of large gains for non-whites and unmarried

[20] The treated group for short-term outcomes is women with children under one, and various control groups include women with older children, men, and women with newborns in other states. They test for longer-term outcomes among women with children aged one to three years.

women. This result confirms the model's prediction that the law should weakly increase time spent on leave. The CPS is well suited to study impacts on leave-taking due to detailed coding of labor-force status.[21] However, the lack of information on precise birth timing or labor-force attachment during pregnancy limit the paper's ability to studying time-varying impacts on labor-force participation. The study finds small insignificant effects on employment and participation in the first year and no significant impacts in the medium term.[22] But this could be due to either a zero effect on participation (whereby all leave-takers are as in Case 1), or imprecision due to inaccurate estimates of birth timing and averaging zero long-term impacts with positive short-term impacts since they estimate impacts for all mothers with children under one (or between one and three).

Baum and Ruhm (2013) analyze the California law using the longitudinal employment outcomes and accurate birth timing in the National Longitudinal Survey of Youth (NLSY-97). The findings about leave-taking, work, and suggestive evidence on job-continuity corroborate the predictions of the model. Because the NLSY-97 cohort was only 19 to 23 years old when the California law was implemented, the difference-in-differences results compare outcomes for women who gave birth before age 20 to those who gave birth at age 24 on average. The paper, thus, provides evidence for a relatively young sub-set of the women potentially affected by the law, and is not able to study heterogeneous impacts by education as most college-educated women give birth after age 24.[23]

[21] In the CPS, while all types of leave are coded separately from time at work, paid leave and unpaid leave cannot be distinguished (Rossin-Slater et al 2012, footnote 14).
[22] The paper speculates that findings of increased wages and hours (marginally significant) in the medium term may be due to increases in "job continuity" in the first year after birth, which is consistent with the predictions of my model if increases in short-term attachment lead to earnings growth.
[23] Baker and Milligan (2008) also use longitudinal data though they do not directly analyze *paid* parental leave. The paper analyzes a series of expansions in job-protected maternity leave in Canada from the 1960s to the 2000s where the increments go from "modest" leave durations similar to those in the US to much longer durations closer to

1.4 Data and Methodology to Flexibly Measure Interruptions in Labor-Force Participation

I measure month-to-month changes in women's labor-force participation from 24 months before until 24 months after a birth using an event study specification with the detailed longitudinal data of the Survey of Income and Program Participation (SIPP). This flexible measure of birth-related labor-force interruptions allows me to test the time-varying predictions of the model presented in the previous section.

1.4.A Survey of Income and Program Participation

The SIPP is a series of nationally representative household panel surveys each approximately 48 months long with sample sizes large enough to study state-level policies. Households are interviewed every four months and provide information for the current and each of the previous three months. To measure birth-related interruptions in labor-force participation both before and after the California and New Jersey laws I use the 1996, 2001, 2004 and 2008 panels.[24]

I construct a sample of all women age 18 to 45 who give birth during each of the panels. I use the month of birth of each household member and variables that indicate the relationship of mothers to children to determine the month each woman gives birth.[25] Some women give birth

European-style mandates. They use short 6-month panels and repeated cross-sections from the Canadian Labour Force Survey to estimate the impact of increasing durations of job-protected leave on leave-taking and on job continuity. Given a precise measure of birth timing, they look for effects in the six months centered around birth. They do not directly measure labor force participation, but their findings about job-continuity confirm the predictions of my model that the introduction of leave—even short-duration leave—causes some women to stay with their employer rather than exiting the labor force. While the authors find effects in the short term, they do not directly compare these to effects in the longer term which would be an interesting extension since the model predicts that as the duration of leave benefits rises, effects on women who would otherwise take long exits from the labor force after a birth should increase.

[24] The timing of each panel is as follows: 1996 Panel: April 1996-March 2000, 2001 Panel: February 2001-January 2004, 2004 Panel: February 2004-January 2008, 2008 Panel: September 2008-December 2012.

[25] If there are no other own children in the household when a woman gives birth, I code it as a first birth; otherwise I code it as a higher-order birth.

more than once during a SIPP panel, and I use the first recorded birth as the reference event for my analysis.[26] Differential access to paid leave prior to the California and New Jersey laws for women with different levels of education, suggest separating the sample by education. I define two categories -- less than a bachelor's degree and at least a bachelor's degree.[27]

I create a measure of labor-force participation for each month based on monthly labor-force status. A woman is only coded as *out* of the labor-force if she has "no job all month, no time on layoff, and no time looking for work." A woman is *in* the labor force if she is "with a job" at least one week of the month, including months when she is absent from work with or without pay (due to leave or layoff), or if she is looking for work.[28]

The SIPP labor-force status codes separately identify unpaid leave, but women on *paid* leave are simply coded as "with a job, worked for pay," making it impossible to measure paid leave use directly. Appendix Figure 1.2 shows results for the proportion of women coded as "with a job the entire month, worked all weeks," which includes *both* women who are actually working *and* those on paid leave and serves as a measure of attachment to a job.[29]

1.4.B Event Study Methodology

An event study specification (Jacobson et al., 1993) allows me to study the monthly pattern of labor force participation for women who give birth, and measure interruptions as

[26] The fact that a woman has another child may naturally affect her outcomes, but the choice to have another child may be jointly determined with other labor-force outcomes.
[27] When categorizing women by time-varying characteristics such as age or educational attainment, I use the mother's status in the month of birth as the reference level.
[28] I test the robustness of my findings to using employment—defined as being "with a job" at least one week of the month—as the outcome of interest rather than labor-force participation, which includes layoff and job search, and find that the overall trends are very similar. Results available upon request.
[29] From SIPP documentation: "a person worked each week in any month when they were (a) on the job the entire month, or (b) they received wages or salary for all weeks in the month, whether they were on the job or not." The SIPP asks questions about *usual* hours at work on a four-monthly basis, but because maternity leave is not "usual" the responses do not help to sort out time on paid leave versus working.

changes in participation relative to pre-birth attachment. I pool all observations for women who give birth during a SIPP panel and estimate the following regression model by least squares

$$(1.1) \qquad \boldsymbol{Y_{it} = \alpha_i + \sum_{j=-25}^{25} \delta_j B_{it}^j + \gamma_t + \epsilon_{it}}$$

where Y_{it} is labor-force participation for woman *i* in month *t*, α_i are individual fixed effects and γ_t are year fixed effects.[30] The B_{it}^j are a set of dummy variables indicating each observation's timing relative to a birth, where *j* ranges from 24 months *before* to 24 months *after* a woman gives birth. If b_i is the month a woman gives birth, then

$$B_{it}^j = \begin{cases} \mathbf{1}(\, t < b_i - 24) & \text{for } j \leq -25 \\ \mathbf{1}(t = b_i + j) & \text{for } -24 \leq j \leq -13 \text{ and} - 11 \leq j \leq 24 \\ \mathit{omitted} & \text{for } j = -12, \\ \mathbf{1}(t > b_i + 24) & \text{for } j \geq 25 \end{cases} ,$$

where I omit the dummy for 12 months prior to birth (or in some cases a series of pre-birth months).[31] All observations more than 24 months before a birth are captured by a single dummy and similarly for observation more than 24 months after birth (hence the sum from -25 to +25).

The event-study dummies, B_{it}^j jointly represent a timeline indexed to the date a woman gives birth and make it possible to estimate average outcomes for women who are *j* months before (or after) birth even if these women gave birth in different calendar months. For example, $B_{it}^j = 1$ if in period *t*, woman *i* gave birth *j* months earlier (or if *j* is negative, *j* months later.)

[30] For the binary labor-force participation outcome, I estimate a linear probability model. I calculate variance using a Huber/White heteroskedasticity-robust estimator clustered at the individual-mother level. This allows for arbitrary covariance over time within units, and allows for heteroskedasticity across units, which is inherent in the linear probability model.

[31] With smaller sample sizes in the state-level analysis I omit a series of months so that results are not overly influenced by the noisy single-month estimates in the early pre-birth period. Sample size gets smaller as I move away from the month of birth (as described in the data section). The maximum sample is in the months directly around birth. The results are robust to omitting other groups of pre-birth moths.

Since I omit B_{it}^{-12}, the δ_j coefficients map out the time path of changes in participation relative to participation a year before the birth. Estimating equation (1) including months prior to birth makes it possible to observe if changes around births are disruptions related to births, or rather continuations of pre-trends for women who go on to have births.

SIPP panels are approximately four years in length. As a result, using all of the births that occur in each panel will mean that not all women in my sample have information for the full 24 lead and 24 lag months because women give birth at different points over the course of the panel. The individual fixed-effects specification in equation (1), however, gives consistent estimates of changes in participation for an unbalanced panel as long as the reason why a woman has missing information is uncorrelated with the $\epsilon_{it}'s$. Aside from attrition, whether I have data for a woman in any month *j* only depends on when during the panel she gives birth. In other words, all that is required for consistency is that, conditional on giving birth during the panel and any time invariant characteristics, *when* over the course of the panel that birth falls is not correlated with omitted variables affecting participation. It is unlikely that women would time their births relative to the census bureau's schedule for fielding SIPP panels.[32] Panel attrition remains a legitimate concern, and in a robustness check I find that the main results of the paper are essentially the same for a sample that excludes all women who left the panel or were absent from the panel for more than three straight months.[33]

[32] While we may be worried that over time, age at first birth for different cohorts has shifted and that a one or two year difference in time of birth is relevant, by using fixed effects, I control for mothers' birth cohort. Another concern is that women may time births relative to the business cycle, which is controlled for using year fixed effects.
[33] Note that in 2004, the Census Bureau randomly dropped half of the sample for budget reasons. I do not count these women as having attrited from the sample in my robustness check. Also, some women enter the panel after the first wave because they enter a household that is in the panel. These women are also not excluded in the robustness check.

1.4.C Patterns of Birth-Related Interruptions in Labor-Force Participation by Education

My model predicts that paid leave should affect the labor-force participation of women who take brief birth-related exits from the labor force. The event study methodology is well suited to studying time-varying patterns in outcomes. Figure 1.4 shows the results from estimating a basic version of equation (1)—including only individual fixed effects and event-study dummies—among all women who give birth in the 2004 and 2008 SIPP panels plotted separately by education category and by first and higher-order births.[34] By adding the δ_j coefficients back to the average level of participation in the left-out period, the plots show the pattern of women's average labor-force participation from 24 months before to 24 months after births. Figure 1.4.1 shows that birth-related interruptions are common among women in both education categories. For women with at least a bachelor's degree, labor-force participation falls 23 percentage points from 92 percent a year before first births to 70 percent two years after higher order births (24 percent drop). Less-educated women start with lower levels of participation prior to birth—79 percent a year before first births—but experience a similar drop of 24 percentage points (or 30 percent) by two years after higher order births.

Examining the shape of the plots in Figure 1.4.1 reveals that the behavior of more- and less-educated women differs not only in levels of participation, but also in terms of the *pattern* of birth-related interruptions. Women with at least a bachelor's degree exhibit relatively stable levels or participation up to three months before first births, when participation drops 15 percentage points from month -6 to month +3, and then remains relatively flat at this lower level for a full two years after first births. A similar, though muted, pattern starting from a lower level is seen for higher-order births. For women with less than a bachelor's degree, participation

[34] Appendix Table 1.1 gives summary statistics for the birth sample from the SIPP used in these figures.

starts to fall earlier, around a year before births. And in contrast to the smooth level-shift after birth seen for more-educated mothers, the profile for less-than-bachelor's women exhibits a dip-and-rebound pattern around both first- and higher-order births before leveling off around six months after births. For first births, between months -3 and +6, participation falls from 63 percent to 55 percent and then rebounds to 63 percent.

Figure 1.4.2 plots the results of estimating equation (1) pooling first and higher-order births, and including year fixed effects. By simply plotting the δ_j coefficients, Figure 1.4.2 abstracts away from level differences (normalizing them to zero at -12) between the education categories and focuses on the different patterns in birth-related labor-force interruptions. Recalling the behaviors described in the labor supply model above, the profile of participation of more-educated women appears to be an average of Cases 1 and 2 – women who do not exit the labor force at all, and women who exit for extended periods respectively. Meanwhile the pattern for less-educated women appears to include at least some women who exhibit the behavior described in Case 3 – women who take brief exits, returning to the labor force by six months after giving birth.[35,36] These women are the types the model predicts are most likely be impacted by the introduction of a paid leave law.

[35] I describe the relatively flat profile from +6 to +24 as evidence of "stable" behavior – sustained labor-force exits, with the remainder of the population continuously in the labor force. A stable average could also arises from equal numbers of women exiting and entering the labor force in each month. Since I have longitudinal data, I can test to ensure that this type of "churning" is not the underlying behavior driving the averages in the figures. I calculate the duration of months in or out of the labor force for women at various points after birth. The distribution of durations in a given state (in or out of labor force) at 12, 18 months, for example, is very consistent with the three cases discussed in Section III.A, rather than with offsetting churn behavior.

[36] Interestingly, this dip and return pattern from -3 to +6 was in evidence *across* the education spectrum in the 1980s prior to the rise in access to voluntarily-provided paid leave among college educated workers (Byker, 2012). A figure comparing 1980s to 2000s birth-related interruption patterns for women with at least a bachelor's degree is shown in Appendix Figure 1.1.

1.5 Empirical Strategy for Estimating the Impact of Paid Leave Laws in California and New Jersey

1.5.A Identification Strategy : Triple-Difference Event Study

The goal is to identify the impact of laws that are implemented (1) in certain states—California and New Jersey (2) in specific years—2004 and 2009 (3) which women are eligible for in some time periods—after birth—but not others—before birth. This scenario suggests a triple-difference estimator such as has used by Gruber (1994) to analyze employer mandates, Ruhm (1998), Waldfogel (1999) and Baum (2003a) to study geographic variation in parental leave policies, and most recently Curtis et al. (2013) to analyze the California paid leave policy. Introducing a third difference beyond the standard differences-in-differences strategy allows for relatively weak identification assumptions, but the results are often difficult to interpret because they are in terms of relative differences between a treatment and control group. Because these previous studies use repeated cross sections, they construct the third difference by comparing outcomes of women with very young children to women with older children (or even to men). I observe the same women over time, so that women's *own* pre-birth outcomes serve as the control for outcomes during the eligible period, leading to a more intuitive and easy to interpret comparison than is often used in a triple-differences strategy. By comparing pre-birth participation to post-birth participation, my third difference is a natural measure of labor-force interruptions. I compare women's birth-related interruptions in labor-force participation in California and New Jersey before and after the laws with birth-related changes for women in states without paid leave policies to produce differences-in-differences-in-differences (DDD) estimates of the impact of the laws.

By measuring birth-related changes in labor-force participation in an event-study framework, I reveal whether the laws have different impacts in the months immediately around birth as compared to one or two years after a birth. Given the predictions of my theoretical framework—impacts in the short-term, but not the longer-term—this strategy will be able to detect effects that other studies cannot due to cross-sectional data limitations that require averaging estimates among all women with children under one, or children aged one to three. Finally, this is the first study I am aware of to analyze the New Jersey law.

1.5.B Specification and Identification Assumptions

The DDD strategy controls for factors that may be correlated with outcomes through a series of fixed effects – year, state and state-by-year. With a set of month-relative-to-birth dummies B_{it}^{j} defined as in equation (1), I estimate the following equation

$$
\begin{aligned}
Y_{its} = \theta_i + \boldsymbol{\beta'}_{\mathbf{1}}\boldsymbol{Year_t} + \boldsymbol{\beta'}_{\mathbf{2}}\boldsymbol{State_s} + \sum_{j=-25}^{25} \delta_j B_{it}^{j} \\
+ \boldsymbol{\beta'}_{\mathbf{3}}\boldsymbol{Year_t} \times \boldsymbol{State_s} + \sum_{\boldsymbol{j=-25}}^{\mathbf{25}} \boldsymbol{\gamma_j B_{it}^{j}} \times \boldsymbol{Year_t} + \sum_{\boldsymbol{j=-25}}^{\mathbf{25}} \boldsymbol{\pi_j B_{it}^{j}} \times \boldsymbol{State_s} \\
+ \sum_{j=-25}^{25} \alpha_j B_{it}^{j} \times Policy_{ts} + \epsilon_{its}
\end{aligned}
\tag{1.2}
$$

where Y_{its} is labor-force participation for individual *i* in living in state *s* in period *t* and θ_i is set of individual fixed effects. $\boldsymbol{Year_t}$ is a vector of year-specific indicator variables and $\boldsymbol{State_s}$ is a vector of state indicator variables.[37] $Policy_{it}$ is an indicator equal to one if a paid parental leave law is in effect in period *t* in state *s*. The vector of coefficients $\boldsymbol{\alpha_j}$ provides monthly estimates of the treatment effect of the laws for each month before and after birth. I estimate equation (2)

[37] Note that technically state fixed effects are subsumed in the individual fixed effects as each woman is categorized by the state in which she gives birth.

using ordinary least squares clustering standard errors at the individual level.[38]

Because the policy treatment varies at the state-by-year-by-month-relative-to-birth, I am able to control for (observed and unobserved) state-specific shocks using state-by-year fixed effects. As a result, the identification assumption is that there are no unobserved contemporaneous shocks that only affect women around birth (but not before birth) in policy states in the same years that the laws go into effect. An example of policies that would impact women differentially after birth than before are TANF work requirement exemptions based on age of youngest child which vary over time and by states. I examine these exemptions by state and determine that rules did not change in the same period as the paid leave laws were enacted in California, New Jersey or any of the control states.

1.5.C *A*nalysis Sample

Table 1.2 provides numbers of observations by education for my analysis sample of women in the SIPP who give birth in California and New Jersey before and after policies are enacted in their respective states. I also include a set of control states as in Rossin-Slater, Ruhm and Waldfogel (2013) for the main specification – New York, Florida and Texas – which are the next three largest states after California. Using the largest states provides the most precision in the estimation given that the SIPP's sample sizes per state become small for many states. I also test the robustness of my results to two different methods for choosing control states. While the synthetic control method (Abadie et al., 2010) is not directly suited to testing my hypotheses (given that there are multiple treatment states), I use a similar strategy to select comparable states by matching on California and New Jersey's pre-policy characteristics including political

[38] I do not cluster at the state level because, as I describe in the next section my analysis involves two policy states and less than 10 control states while inference with cluster-robust standard errors is based on the assumption that the number of clusters goes to infinity (Cameron et al., 2008). I am able to address much of the concern posed by (Bertrand et al., 2004) relating to serial correlation in the state specific shocks in standard DD estimators by including state-by-year fixed effects.

leanings, demographic make-up, and labor-force participation by gender and among women who give birth. Secondly, I use other SDI states as control states (though other than New York, they are very small states). Note that Texas and New York which are included in my main specification by the size criteria are also suggested by either the matching or SDI criteria. The results are very similar across these different choices of control states.

1.6 Estimates of the Impact of Paid-Leave Laws on Birth-Related Interruptions in Labor-Force Participation

Figure 1.5 visually builds up the empirical strategy from simple differences to the DDD. Recall that there are two time dimensions in the analysis: 1) time relative to a birth event, and 2) before and after a paid leave law is enacted. In Panel A, I split the sample of all women who give birth in California or New Jersey into two groups, those who give birth before a paid leave law is enacted in their state, and those who give birth after the law is enacted.[39] Figure 1.5.A plots event studies of labor-force participation around birth separately for each group. The pre-law group exhibits a sharper dip in participation from around six months before birth to around 4 months after, while the post-law group exhibits a smoother interruption pattern. The shaded area between the two event study lines is a DD estimate of the effect of the laws using two simple differences: 1) women's change in labor-force participation between non-eligible and eligible periods and 2) the difference in those relative changes for groups who give birth before and after the law is enacted.

The dashed line at the bottom of Panel A plots the vertical difference between the two event studies: (post-law LFP) **–** (pre-law LFP) for each month relative to birth. These DD estimates show that the laws cause a "bump" of five to eight percentage points in participation in the months centered around birth compared to relatively flat pre-birth and post-birth impact

[39] Recall that there was at least a year between the passage and enactment of the laws in each state and evidence of publicity preceding the implementation.

estimates. A joint test of significance of the coefficients on months -3 to +3 has a p-value=0.05.[40]

Next I introduce the control states. If the DD results are due to national trends over time in birth-related interruptions, then the pre-post differences would falsely attribute the increase in labor-force participation to paid leave laws. Estimating equation (2) including women who give birth in other states allows me to control for trends in economic conditions that affected all states. In this specification I also include state-by-year fixed effects to control for any state-specific shocks that could be correlated with passage of the laws.

Panel B of Figure 1.5 repeats the dotted DD line from panel A and adds a plot of the coefficients on the policy interactions ($\boldsymbol{\alpha_j}$) from estimating equation (2). Pre-birth months -24 to -18 are omitted as the pre-birth base of comparison; the results are robust to omitting other combinations of months. These DDD estimates confirm the findings of the DD that the laws have a statistically significant impact on labor-force participation in the six months centered on birth—a joint test of the significance of months -3 to +3 has a p-value=0.06. There are small and insignificant impacts prior to month -6. If the law does not draw women into the labor force in the pre-birth period, then the identification strategy assumes that there should be no impact of the law when women are not eligible, and we should expect the impact of the law 12 to 18 months *before* birth to be zero. As discussed in Section 1.3.A, Curtis, Hirsch and Schroeder (2013) find a very small increase in hiring of women at future risk of leave-taking, so a small and insignificant estimated impact in the year before birth lends credibility to the mechanism of my identification strategy.[41]

---

[40] The test is based on a specification where months -24 to -18 are omitted.

[41] As seen in Section 1.4.C, birth-related exits start to occur in the few months before birth (how early depends on the education level and parity of birth) which is why it makes sense to see impacts of the law in the months directly before birth. I expect to see zero effects in the *earlier* pre-birth periods such as pre-pregnancy periods.

Meanwhile there are smaller and insignificant impacts of the laws after month +6. A joint test for months +12 to +18 has a p-value=0.24 with an average point estimate of five percentage points. Table 1.3 summarizes the results of Figure 1.5. California and New Jersey's paid leave laws substantially reduce short-term interruptions, but my results suggest they have less impact on sustained interruptions. This pattern of time-varying impacts is consistent with the predictions of my labor supply model.

Based on evidence of greater access to paid leave among more-educated women in the absence of the laws, I estimate equation (2) separately for women with at least a bachelor's degree and women with less than a bachelor's degree. Figure 1.6 shows the results of this analysis overlaying the DDD estimates for the full sample with the results for the two education groups. Figure 1.6 reveals that the impact of the laws is driven exclusively by the changes in labor-force participation among the less-educated women.[42] Joint tests of the significance of the DDD coefficients from months -3 to +3 for more-educated women have a p-value=0.84, while the joint test for less-educated women has p-value=0.05. Table 1.3 summarizes the results of Figure 1.6Figure 1.6, comparing estimated impacts and significance levels before and after birth by education.

The significant bump in labor-force participation in the months directly around birth is the inverse of the short-term exit-and-return pattern seen for less-educated women in Figure 1.4. My results imply that paid leave laws reduce short-term intermittency around birth for mothers with less than a bachelor's degree, making their labor-force participation profile around birth resemble the pattern for more-educated mothers.

Appendix Figure 1.2 shows DD and DDD estimates for the outcome "with a job entire

[42] In results, not shown, I find that while the impacts are not-statistically distinguishable, the effects seem to be stronger for women with some college (including associate's degrees) than among the least-educated women with at most a high school diploma.

month, worked all weeks." Recall that in the SIPP this category includes *both* women who are actually working *and* those on paid leave (see footnote 29). These results show that paid leave laws give rise to a similar bump around birth in the proportion of less-educated women who have a paying job, which indicates that the increase in labor-force participation shown in Figure 1.5 is largely driven by increased attachment to jobs.

1.7 Conclusion

This paper estimates the impact of recent state-level paid parental leave laws on women's birth-related career interruptions using the high-frequency longitudinal structure of the SIPP. Exploiting observations on a month-to-month basis, my results show that short-duration paid parental leave substantially reduces the incidence of short-term labor-force exits occurring in the seven months centered around birth but has less impact on longer-term interruptions. Results for the impact of the laws on the proportion of women with a paying job indicate that this represents an increased attachment to jobs. These impacts are only present for mothers with less than a college degree who are less likely to have access to private paid leave in the absence of a mandate.

If parental leave is a benefit valued by workers then, in the absence of market failures, we would expect leave to part be of an optimal compensation package voluntarily negotiated between workers and employers obviating the need for a mandate. There may be market failures including externalities in terms of child outcomes or a coordination problem among hiring firms that lead to suboptimal parental leave provision in a competitive market.[43]

If, as advocates claim, firms benefit from lower employee turnover costs when they offer leave, there may be countervailing forces to overcome these potential market failures. As shown in Section 1.2.B, access to privately-negotiated leave has increased substantially for some

[43] These potential market failures are discussed in detail in Ruhm (1998).

women since the 1980s implying that the joint value of leave is captured by some firms and workers without a mandate. The fact that voluntary paid leave is predominantly seen in high-education jobs suggests that market failures continue to prevail in low-education sectors, or low-skill workers do not value the benefits and/or employers of low-skill workers do not sufficiently value retaining them.

The substantial impact of paid leave I find among less-educated women indicates that they *do* value parental leave but may not be able to afford unpaid leave or are not able to bargain for paid leave in the absence of a mandate. This could have implications for the distribution of earnings across education if increased attachment to the labor-force leads to greater tenure with employers, increased accumulated experience and long-term earnings growth. If these follow-on gains to increased attachment accrue to the women impacted by the new paid leave laws in California and New Jersey, it could imply that the mandate corrects a market failure. If, on the other hand, there are no gains from increased attachment, the mandates could lead to distortions like statistical discrimination against women whom employers perceive as at risk for leave-taking. This could lead to less hiring of women, fewer promotions, and flatter wage profiles.

The findings of this paper about the short and medium-term labor supply impacts of short-duration paid leave, thus, take an important step in understanding the costs and benefits of these policies and suggest longer-term outcomes that will be useful to analyze as time since implementation increases.

Figure 1.1.1. Use of Paid Parental Leave after First Births

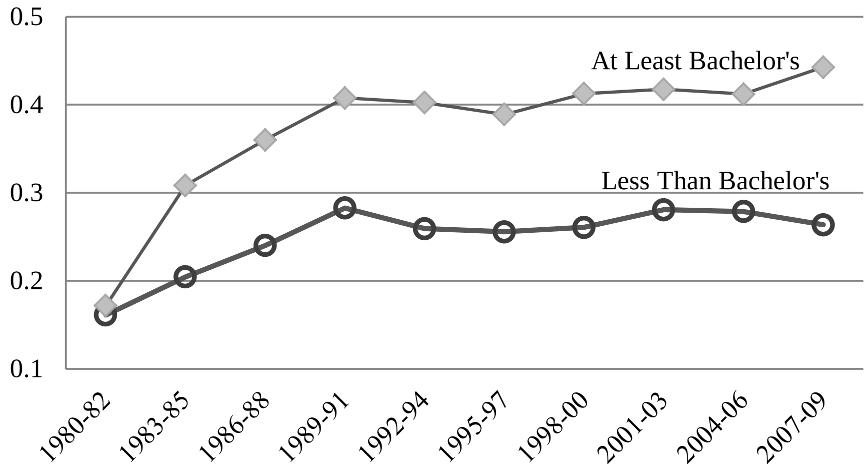
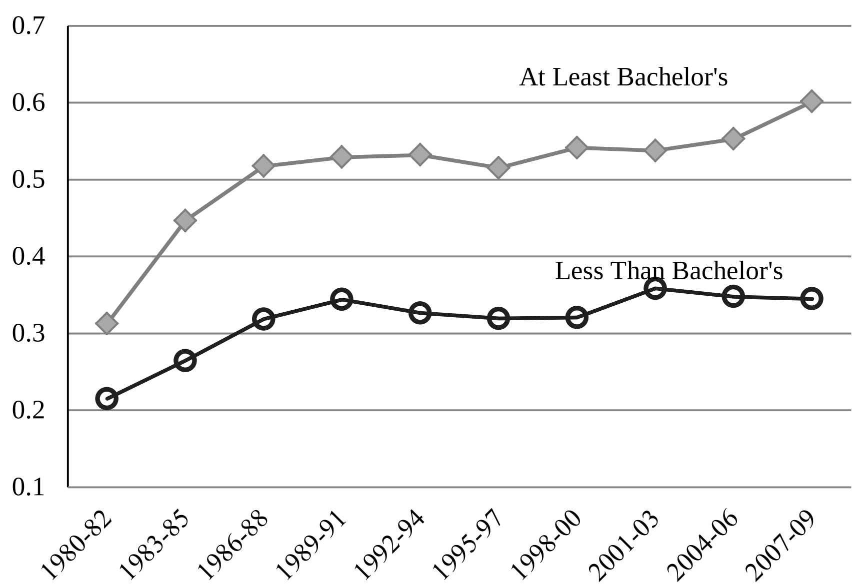


Figure 1.1.2. Use of Any Paid Leave after First Births



Notes: The figures show trends in of paid leave usage in the 12 weeks after a first birth among women who worked during pregnancy as reported in SIPP retrospective fertility modules in the 1996, 2001, 2004 and 2008 SIPP panels. The survey question refers to "maternity" rather than "parental" leave. "Any paid leave" includes the following categories: paid maternity leave, paid sick leave, paid vacation leave and other paid leave. SIPP sampling weights are used.

Figure 1.2.1. Budget Constraints Under Unpaid and Paid Parental Leave Laws

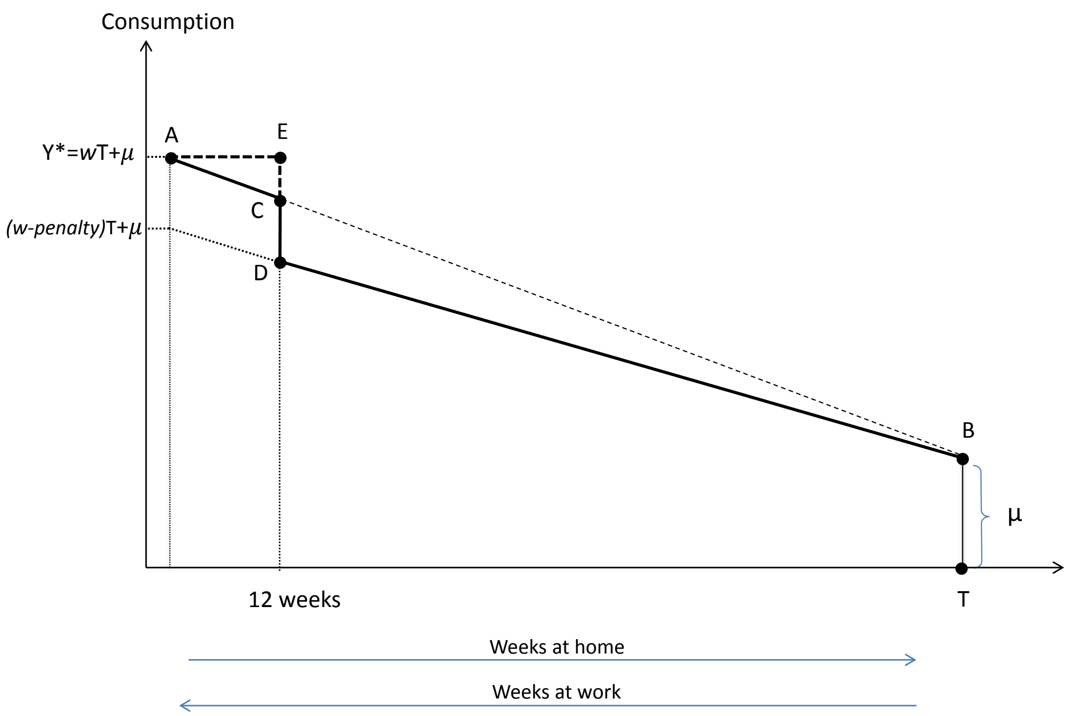


Figure 1.2.2. Budget Constraint Prior to Introduction of Paid Leave

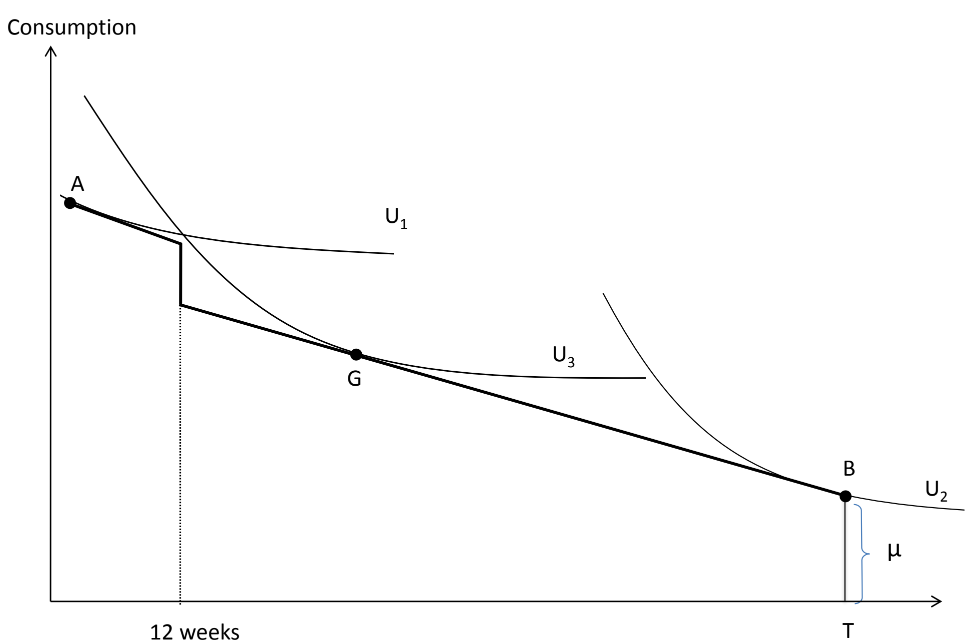


Figure 1.3. Labor-Supply Responses to the Introduction of Paid Parental Leave: Four Cases

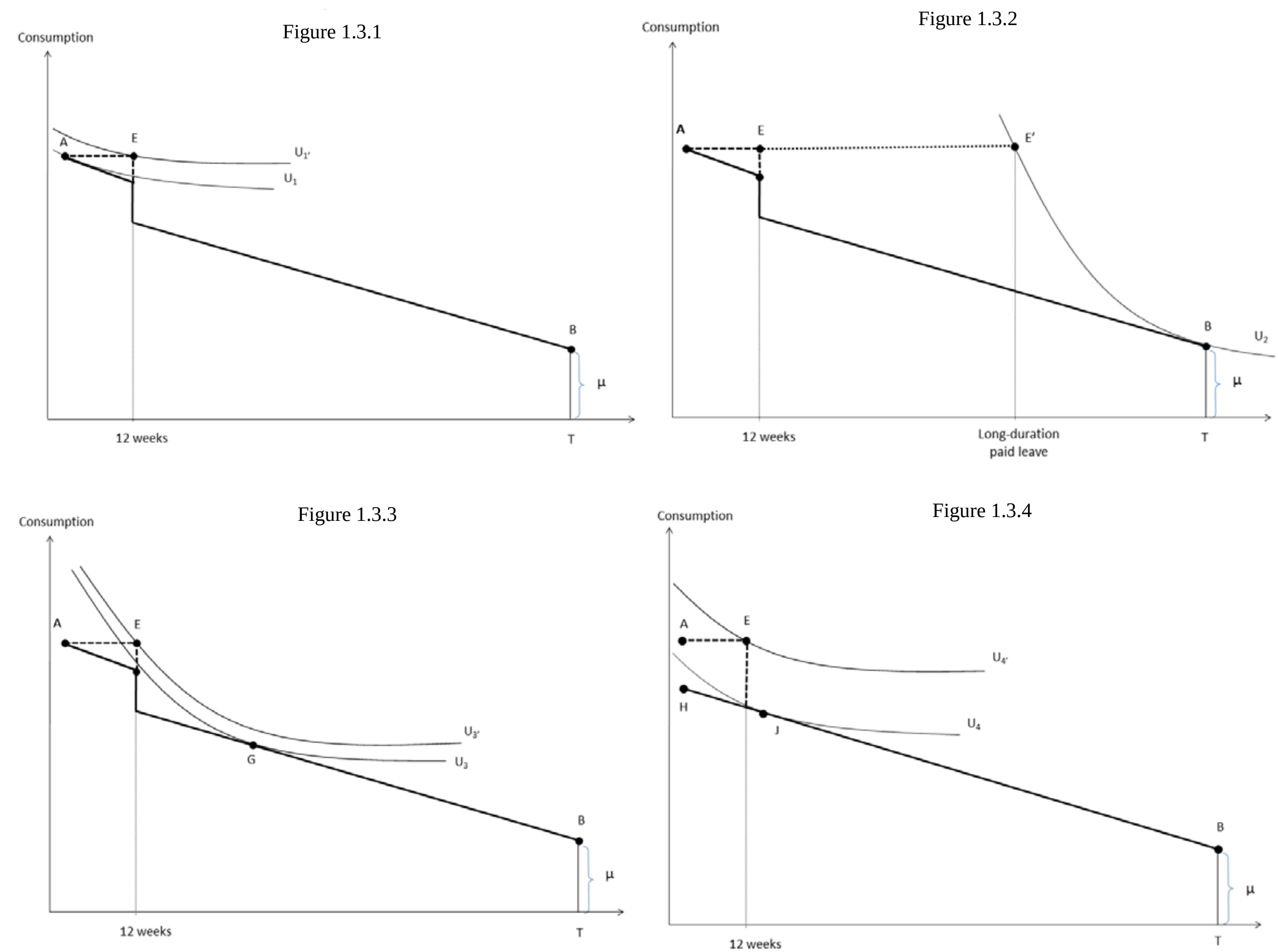
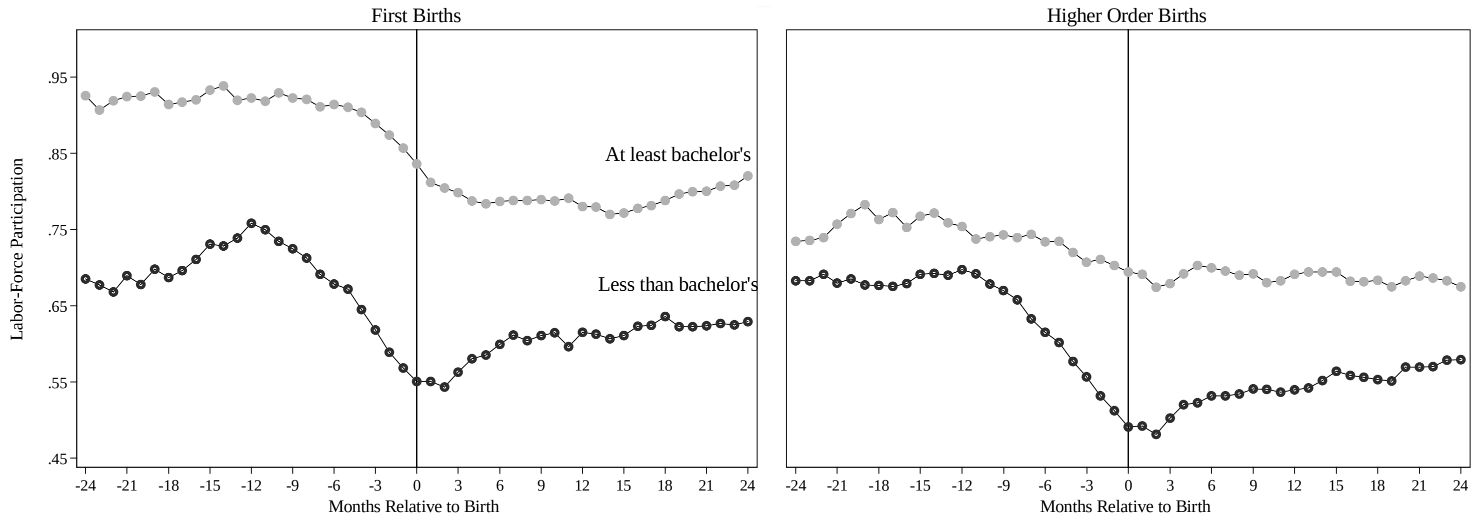
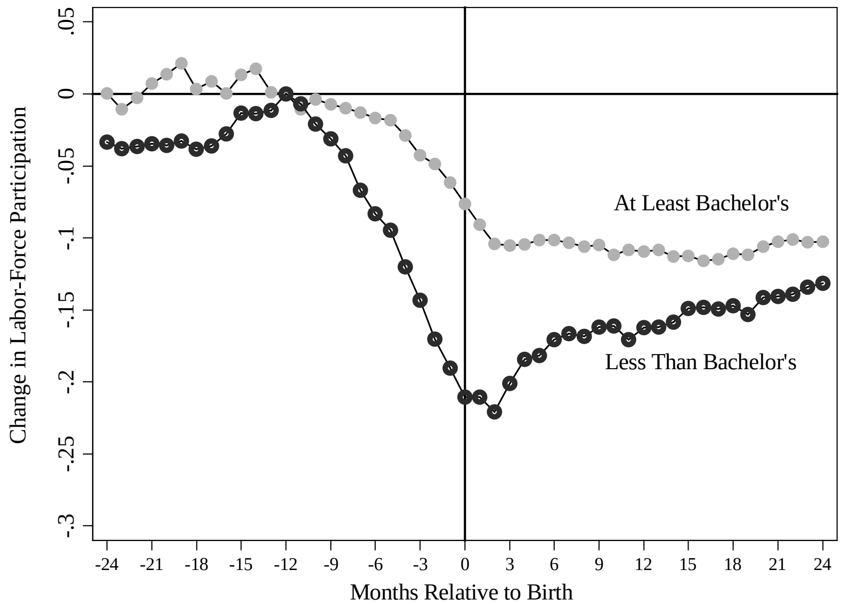


Figure 1.4.1. Patterns of Labor-Force Participation around Birth



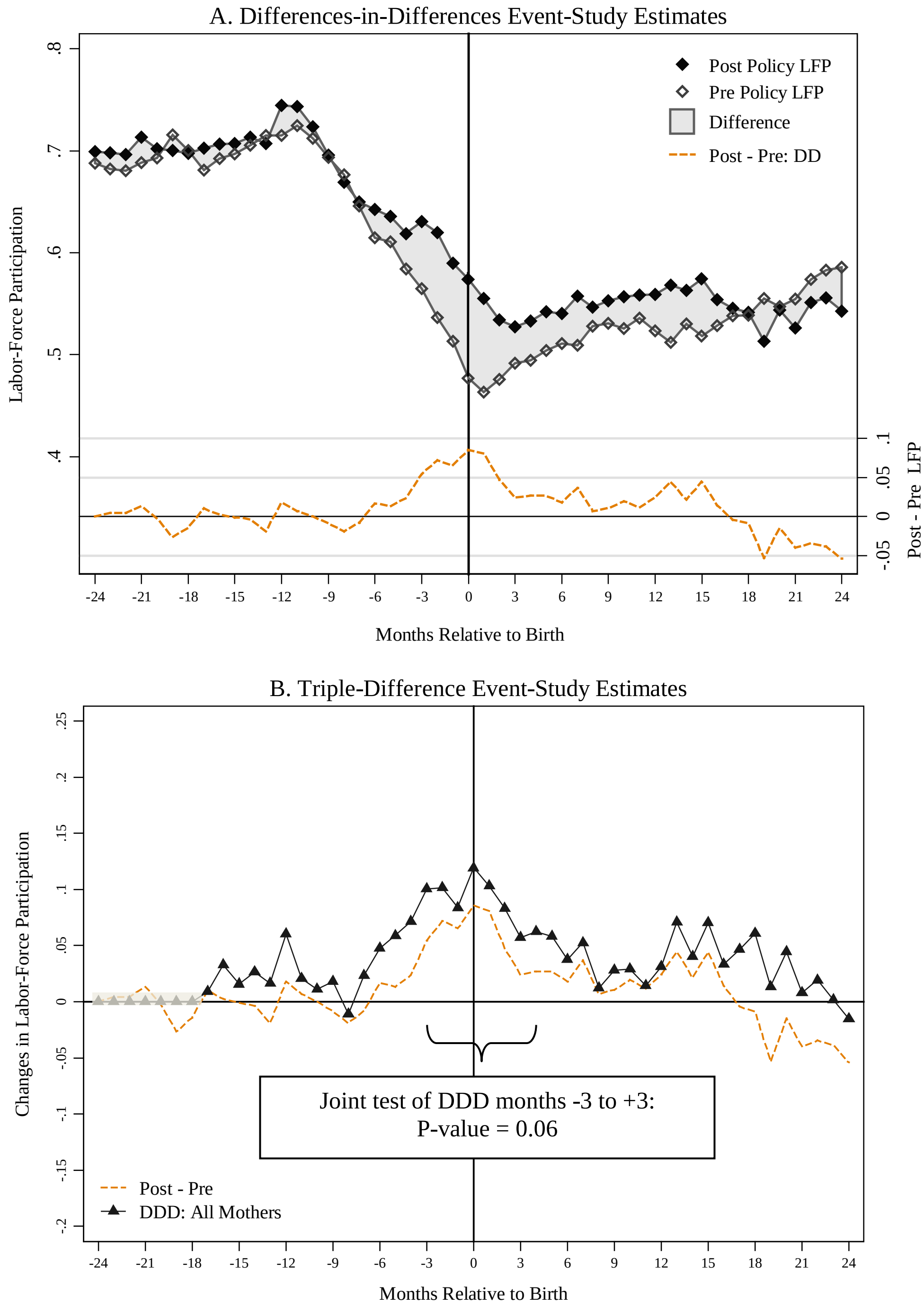
Notes: The figure shows the *level* of labor-force participation by parity and mother's education for women age 18-45 who gave birth in the 2004 and 2008 SIPP panels. Each line plots coefficients on month-relative-to-birth dummies added back to the average level of participation a year prior to birth—these are the δ_j coefficients (added back to the level of participation in the omitted period (m=-12)) from estimating equation (1) with a dependent variable an indicator for being in the labor force. SIPP sampling weights are used.

Figure 1.4.2. Changes in Labor-Force Participation around Birth



Notes: The figure shows *changes* in labor-force participation around birth by mother's education for women who gave birth in the 2004 and 2008 panels. In this figure, first- and higher-order births are pooled. By plotting the coefficients on month-relative-to-birth dummies with the month -12 omitted, the level of participation is normalized to zero at one year prior to birth. These are the δ_j coefficients from estimating equation (1) with a dependent variable and indicator for being in the labor force. Year fixed effects are included in this specification and I use SIPP sampling weights.

Figure 1.5. Impact of CA and NJ Paid Leave Laws on Labor-Force Participation around Birth

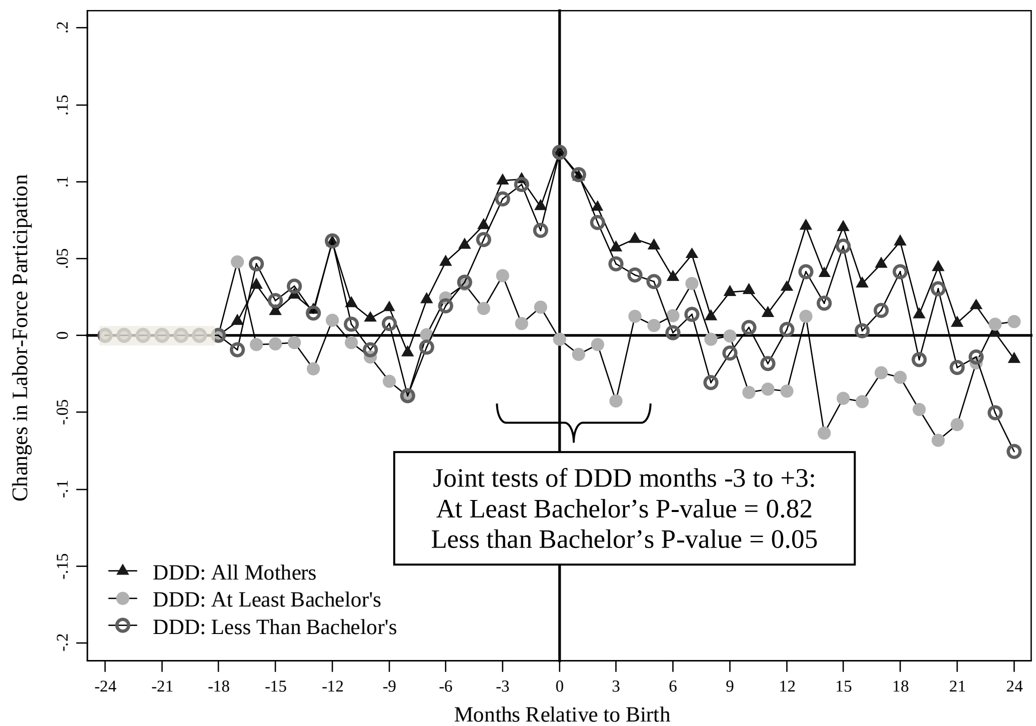


Notes: Panel A shows the level of labor-force participation in the months relative to birth for women giving birth pre- and post-policy in California and New Jersey. The dotted line plots the monthly differences between participation in the pre- and post-policy periods; this provides DD estimates of the impact of paid leave on labor-force

Notes to Figure 5 continued:

participation. Panel B repeats the DD estimates from Panel A, and plots DDD estimates, which are coefficients on the interaction between the month-relative-to-birth and an indicator for giving birth in a policy state after the law was enacted. These are the α_j coefficients from estimating equation (2) with the dependent variable an indicator for being in the labor force. The regression includes year, state, and state-by-year fixed effects as described in Section V.B. Months -24 to -18 are omitted as the pre-birth comparison period (hence the shaded markers for these months). SIPP sampling weights are used. The sample for DDD estimates is women giving birth in California and New Jersey and the control states as described in Table 2. Table 3 provides a summary of these results with average point estimates and significance levels in before- and after-birth periods

Figure 1.6. Heterogeneous Impacts of Paid Leave Laws by Education



0

Notes: The figure plots DDD estimates separately by mother's level of education, which are coefficients on the interaction between the month-relative-to-birth and an indicator for giving birth in a policy state after the law was enacted. These are the α_j coefficients from estimating equation (2) with the dependent variable an indicator for being in the labor force. The regression includes year, state, and state-by-year fixed effects as described in Section V.B. Months -24 to -18 are omitted in as the pre-birth comparison period (hence the shaded markers for these months). SIPP sampling weights are used. The sample for DDD estimates is women giving birth in the policy states—California and New Jersey—and the control states as described in Table 2. Table 3 provides a summary of these results with average point estimates and significance in before- and after-birth periods.

Table 1.1. Administrative Records on Claims for Paid Parental Leave

	California		New Jersey	
	2004/05	2011/2012	2010	2011
Bonding Claims	132,007	183,421	23,696	24,413
% of claims made by women	83%	71%	89%	87%
Average weekly benefit	$409	$497	$486	$489
Average weeks per claim	4.8	5.35	5.2	5.3
Total Paid (millions)	$263.45	$467.57	$60.40	$63.20
Number of births in State	544,843	502,120	106,922	105,883
% of births in the state represented by a claim	24%	37%	22%	23%

Notes: Source: Claims data: New Jersey: http://lwd.dol.state.nj.us/labor/fli/fliindex.html California: http://www.edd.ca.gov/disability/paid_family_leave.htm; birth data: National Center for Health Statistics, National Vital Statistics Reports 2004, 2010, 2011. "Total paid" includes a small number of claims made for other types of family leave.

Table 1.2. Summary of Paid Leave Birth Sample

	Pre	Post	Total	Control States
Full Birth Sample				
California	794	497	1,291	
New Jersey	273	46	319	
Total	1,067	543	1,610	2,126
Less than Bachelor's Degree				
California	614	362	976	
New Jersey	167	26	193	
Total	781	388	1,169	1,577
At Least a Bachelor's Degree				
California	180	135	315	
New Jersey	106	20	126	
Total	286	155	441	654

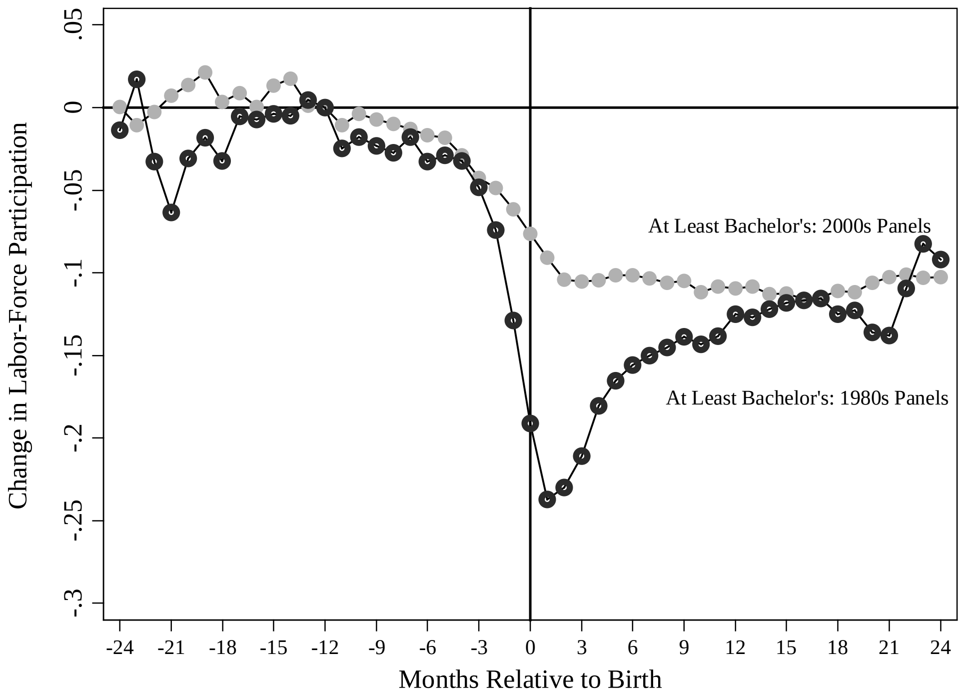
Notes: The sample includes all women age 18 to 45 who give birth in the 1996, 2001, 2004 and 2008 SIPP panels. Pre and Post refer to whether the birth occurred before or after the paid leave law was enacted in the respective state. Control states are New York, Texas and Florida. Pre and post are not listed for the control states as the laws are enacted on different dates in California and New Jersey.

Table 1.3. Summary of Triple-Difference Impacts of CA and NJ Paid Leave Laws on Labor-Force Participation

	Pre-pregnancy	Around birth	Longer-term post-birth
	-16 to -10	-3 to +3	+12 to +18
All Mothers			
Average DDD impact	0.03	0.09	0.05
P-value of joint test	0.21	0.06	0.24
P-value of difference vs -3 to +3	0.11	NA	0.33
Bachelor's Plus			
Average DDD impact	-0.01	0.00	-0.03
P-value of joint test	0.76	0.84	0.34
P-value of difference vs -3 to +3	0.90	NA	0.67
Less than Bachelor's			
Average DDD impact	0.03	0.09	0.03
P-value of joint test	0.33	0.05	0.36
P-value of difference vs -3 to +3	0.25	NA	0.28

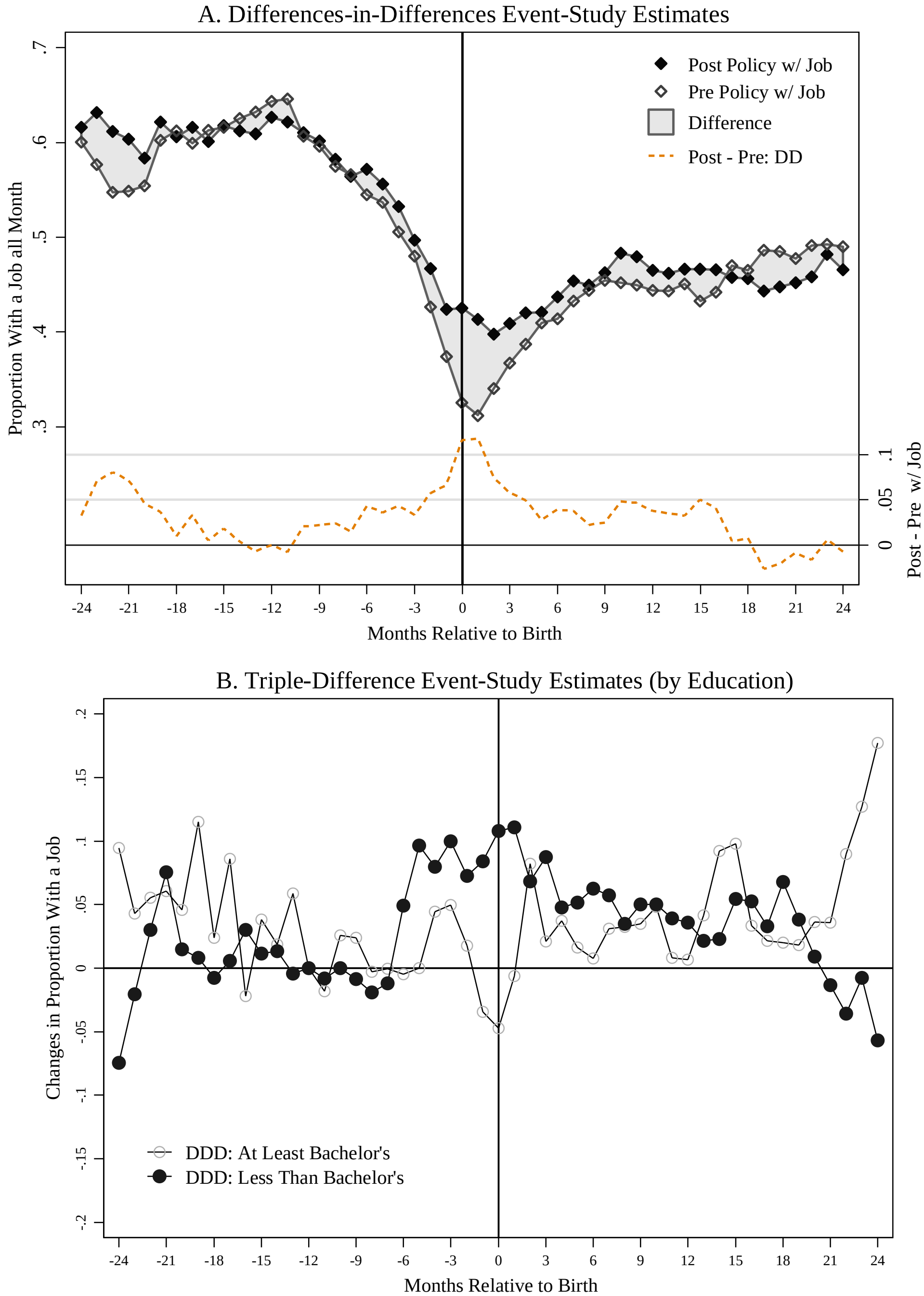
Notes: The table summarizes the monthly DDD results shown in Figures 1.5 and 1.6—see notes to Figures 1.5 and 1.6 for a description of the estimates. Each column summarizes the estimated impacts for a 7-month period: 1) pre-pregnancy, 2) directly around the month of birth, 3) longer-term post-birth. For each time period, rows give 1) the average of the monthly point estimates, 2) the p-value of a test of the joint significance of the monthly estimates in the period, and 3) a test of the significance of the difference of the estimates during the period compared to the impact in months -3 to +3.

Appendix Figure 1.1. Changes in Labor-Force Participation around Birth: Women with at Least a Bachelor's Degree, Comparing the 1980s to the 2000s



Notes: The figure compares birth-related interruptions in labor-force participation in the 1980s to the 2000s for women with at least a bachelor's degree. The lines are plots of δ_j coefficients from equation(1) with dependent variable an indicator for being in the labor force, estimated separately by decade for women with at least a bachelor's degree giving birth at ages 18-45 in the 1984, 1985, 1986, 2004 & 2008 SIPP panels. Year fixed effects are included in this specification and I use SIPP sampling weights. This figure is a version of figures appearing in Byker (2012).

Appendix Figure 1.2. Impact of CA and NJ Paid Leave Laws on Proportion "With a Job All Month" around Birth



Notes: The outcome in this figure is the whether a woman was "with a job entire month, working all weeks." In the SIPP, women who are on *paid* (but not unpaid) leave are given this code as well as women actually working.
Panel A: See notes to Figure 1.5 panel A. Panel B: See notes to Figure 1.6.

Appendix Table 1.1 Summary of Birth Sample from SIPP Panels

	2000s		1980s	
	Pooled 2004, 2008		Pooled 1984-86	
Dates	Feb 04 - Aug 12		Oct 83 - Apr 88	
Births to women age 18-45	6,284		3,670	
First births	2,621	41.7%	1,987	54.1%
Higer-order births	3,663	58.3%	1,683	45.9%
Panel B				
Race				
White	3,964	59.5%	3,035	84.3%
Black	760	13.1%	491	15.7%
Hispanic	1,039	19.3%	NA	
Other	521	8.1%		
Marital Status				
Married Sps present	4,218	69.4%	2,944	80.5%
Separated, Div, Wid	426	5.8%	249	6.6%
Never Married	1,640	24.8%	477	13.0%
Education				
Less than Bachelors	4,377	68.9%	3,054	83.1%
High School or less	2,363	37.1%	1,809	49.2%
Some College	2,014	31.8%	1,245	33.9%
Bachelors Only	1,305	21.2%	386	10.5%
Masters Plus	602	9.9%	230	6.4%
Masters	450	7.4%		
Professional	89	1.4%	NA	
PhD	63	1.0%		
At Least Bachelors	1,907	31.1%	616	16.9%

Notes: In Panel B numbers of observations are shown while percentages in each category are calculated using SIPP sampling weights. The Census Bureau changed survey question over time: the 1980's panels do not give information on Hispanic origin. Ambiguity in 1980s coding of education variables makes it impossible to make an exact distinction between some college, bachelor’s and graduate degree. Measurement error in defining levels of higher education and challenges in reconciling old and new census bureau education questions are well documented (Jaeger 1997, Black et al. 2003). Some college includes associates and vocational degrees.

REFERNCES

Abadie, Alberto; Alexis Diamond and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105(490), 493-505.

Appelbaum, Eileen and Ruth Milkman. 2011. "Leaves That Pay: Employer and Worker Experiences with Paid Family Leave in California," Center for Economic and Policy Research,

Baker, Michael and Kevin Milligan. 2008. "How Does Job-Protected Maternity Leave Affect Mothers' Employment?" *Journal of Labor Economics*, 26(4), 655-91.

Baum, Charles L. 2003a. "The Effects of Maternity Leave Legislation on Mothers' Labor Supply after Childbirth." *Southern Economic Journal*, 69(4), 772-99.

____. 2003b. "Does Early Maternal Employment Harm Child Development? An Analysis of the Potential Benefits of Leave Taking." *Journal of Labor Economics*, 21(2), 409-48.

Baum, Charles L. and Christopher J. Ruhm. 2013. "The Effects of Paid Family Leave in California on Labor Market Outcomes." *National Bureau of Economic Research Working Paper Series*, No. 19741.

Berger, Lawrence M.; Jennifer Hill and Jane Waldfogel. 2005. "Maternity Leave, Early Maternal Employment and Child Health and Development in the Us*." *The Economic Journal*, 115(501), F29-F47.

Bertrand, Marianne; Esther Duflo and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119(1), 249-75.

Bertrand, Marianne; Claudia Goldin and Lawrence F. Katz. 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics*, 2(3), 228-55.

Black, Dan; Seth Sanders and Lowell Taylor. 2003. "Measurement of Higher Education in the Census and Current Population Survey." *Journal of the American Statistical Association*, 98(463), 545-54.

Blau, Francine D. and Lawrence M. Kahn. 2013. "Female Labor Supply: Why Is the Us Falling Behind?" *National Bureau of Economic Research Working Paper Series*, No. 18702.

____. 1997. "Swimming Upstream: Trends in the Gender Wage Differential in the 1980s." *Journal of Labor Economics*, 15(1), 1-42.

____. 2004. "The Us Gender Pay Gap in the 1990s: Slowing Convergence." *National Bureau of Economic Research Working Paper Series*, No. 10853.

Buono, Barbara (Senator Buono). 2008. "Senate Approves Paid Familiy Leave Legislation," Trenton, New Jersey: New Jersey Senate Democrats (njsendems.org),

Bureau of Labor Statistics. 1992 and 1993. Employee Benefits Survey. Washington, D.C.: U.S. Dept. of Labor. http://www.bls.gov/ncs/#datahttp://www.bls.gov/ncs/ebs/.

_____. 2012. National Compensation Survey: Employee Benefits in the United States. Washington, D.C.: U.S. Dept. of Labor. http://www.bls.gov/ncs/#data.

Byker, Tanya. 2012. "The Opt-out Continuation: Education, Work and Motherhood from 1984 to 2008,"

Cameron, A. Colin; Jonah B. Gelbach and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics*, 90(3), 414-27.

Carneiro, Pedro; Katrine Vellesen Loken and Kjell G. Salvanes. 2011. "A Flying Start? Maternity Leave Benefits and Long Run Outcomes of Children," Institute for the Study of Labor (IZA),

Clinton, William. 2013. "Why I signed the Family and Medical Leave Act." Politico, February 5**.** http://www.politico.com/story/2013/02/the-family-and-medical-leave-act-20-years-later-87157.html#ixzz2jmyxWkXn

Curtis, E. Mark; Barry T. Hirsch and Mary C. Schroeder. 2013. "Evaluating Workplace Mandates with Flows Versus Stocks: An Application to California Paid Leave,"

Dahl, Gordon B.; Katrine V. Løken; Magne Mogstad and Kari Vea Salvanes. 2013. "What Is the Case for Paid Maternity Leave?" *National Bureau of Economic Research Working Paper Series*, No. 19595.

Das, Tirthatanmoy and Solomon Polachek. 2014. "Unanticipated Effects of California's Paid Family Leave Program " *The Institute for the Study of Labor (IZA) Discussion Paper Series*, No. 8023.

Deak, Michael. 2008. "New Jersey Workers to Be Notified by Monday, Dec. 15 of Changes in Family Leave Act," *Home News Tribune, Dec 14, 2008.*

Dunifon, Rachel; Anne Toft Hansen; Sean Nicholson and Lisbeth Palmhøj Nielsen. 2013. "The Effect of Maternal Employment on Children's Academic Performance." *National Bureau of Economic Research Working Paper Series*, No. 19364.

Espinola-Arrendondo, Ana and Sunita Mondal. 2010. "The Effect of Parental Leave on Female Employment: Evidence from State Policies." *Washington State University School of Economic Sciences Working Paper Series*, No 2008-15.

Gruber, Jonathan. 1994. "The Incidence of Mandated Maternity Benefits." *The American Economic Review*, 84(3), 622-41.

Han, Wen-Jui; Christopher Ruhm and Jane Waldfogel. 2009. "Parental Leave Policies and Parents' Employment and Leave-Taking." *Journal of Policy Analysis and Management*, 28(1), 29-54.

Hotchkiss, Julie L. and M. Melinda Pitts. 2007. "The Role of Labor Market Intermittency in Explaining Gender Wage Differentials." *The American Economic Review*, 97(2), 417-21.

Jacobson, Louis S.; Robert J. LaLonde and Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *The American Economic Review*, 83(4), 685-709.

Jaeger, David A. 1997. "Reconciling the Old and New Census Bureau Education Questions: Recommendations for Researchers." *Journal of Business & Economic Statistics*, 15(3), 300-09.

Klerman, Jacob Alex; Kelly Daley and Alyssa Pozniak. 2012. "Family and Medical Leave in 2012: Executive Summary," Abt Associates Inc.,

Klerman, Jacob Alex and Arleen Leibowitz. 1994. "The Work-Employment Distinction among New Mothers." *The Journal of Human Resources*, 29(2), 277-303.

Lalive, Rafael and Josef Zweimüller. 2009. "How Does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments." *The Quarterly Journal of Economics*, 124(3), 1363-402.

Martin JA, Hamilton BE, Sutton PD, et al. 2006. "Births: Final data for 2004." National Vital Statistics Reports; 55(1). Hyattsville, MD: National Center for Health Statistics.

Martin JA, Hamilton BE, Ventura SJ, et al. 2012. "Births: Final data for 2010." National Vital Statistics Reports; 61(1). Hyattsville, MD: National Center for Health Statistics.

____. 2013. "Births: Final data for 2011." National Vital Statistics Reports; 62(1). Hyattsville, MD: National Center for Health Statistics.

Mincer, Jacob and Solomon Polachek. 1974. "Family Investments in Human Capital: Earnings of Women." *Journal of Political Economy*, 82(2), S76-S108.

National Partnership for Women and Families. 2012. "Expecting Better: A State-by-State Analysis of Laws That Help New Parent," http://www.nationalpartnership.org/research-library/work-family/expecting-better.pdf

OECD. 2012. OECD Family Database, OECD, Paris. www.oecd.org/social/family/database

O'Neill, June; and Solomon; Polachek. 1993. "Why the Gender Gap in Wages Narrowed in the 1980s." *Journal of Labor Economics*, 11(1), 205-28.

Rodriguez, Robert. 2004. "California's Paid Family Leave Begins," *Knight Ridder Tribune Business News, Jul 01, 2004.*

Rossin-Slater, Maya; Christopher J. Ruhm and Jane Waldfogel. 2013. "The Effects of California's Paid Family Leave Program on Mothers' Leave-Taking and Subsequent Labor Market Outcomes." *Journal of Policy Analysis and Management*, 32(2), 224-45.

Rossin, Maya. 2011. "The Effects of Maternity Leave on Children's Birth and Infant Health Outcomes in the United States." *Journal of Health Economics*, 30(2), 221-39.

Ruhm, Christopher J. 1998. "The Economic Consequences of Parental Leave Mandates:

Lessons from Europe." *The Quarterly Journal of Economics*, 113(1), 285-317.

____. 2004. "Parental Employment and Child Cognitive Development." *The Journal of Human Resources*, 39(1), 155-92.

U.S. Census Bureau. 1995 and 2012. Statistical Abstract of the United States. Washington, DC; http://www.census.gov/compendia/statab/

Van Giezen, Robert W. 2013. "Paid Leave in Private Industry over the Past 20 Years." *Beyond the Numbers: Pay & Benefits (U.S. Bureau of Labor Statistics)*, 2(18).

Waldfogel, Jane. 1999. "The Impact of the Family and Medical Leave Act." *Journal of Policy Analysis and Management*, 18(2), 281-302.

Chapter 2

Fertility and Family Well-being Effects of an Aggressive Family Planning Policy in Peru in the 1990's: A Reweighting Estimator with a Contaminated Treatment Group Approach

With Italo Gutierrez

Abstract: In the mid-1990's President Fujimori of Peru initiated an aggressive family planning program with the stated purpose of addressing widespread poverty. While female sterilization was an official element of the program, anecdotal evidence suggests that health workers were secretly given large sterilization quotas and reportedly used bribes, coercion, and even force to meet them. While the details of the program were not public, the Peruvian Demographic and Health Surveys (DHS) provide evidence of a large increase in sterilizations during the suspected program window. We address three research questions: First, who was affected by the sterilization program? Second, what was the impact of the program on fertility? Third, what, if any, impact did the program have on household well-being? We use a rich set of controls from the DHS with a reweighting procedure modified to account for a "contaminated" treatment group in order to estimate the effects of the sterilization program. We find substantial impacts of the program on fertility, but small or insignificant impacts on other household outcomes. Thus, our results suggest that the mere reduction of fertility may not be associated with improvements in households' welfare in the context of coerced sterilizations.

Acknowledgements:
We are grateful to Martha Bailey, Robert Garlick, Andrew Goodman-Bacon, David Lam, and Jeff Smith for helpful comments and guidance.
.

2.1 Introduction

In the mid-1990's President Fujimori of Peru initiated an aggressive family planning program with the stated purpose of addressing widespread poverty in the country. The 1991-1992 Peruvian Demographic and Health Survey (DHS II) provided evidence that seemed to bolster Fujimori's claim that there was a "the vicious circle [of] poverty--unwanted child--poverty" in Peru.[1] Table 2.1, based on data from DHS II, shows the strong negative correlation between wealth (and education) and fertility in Peru. The Peruvian DHS II also indicated an unmet need for contraception with 35 percent of all women who gave birth within the last five years responding that their latest birth was not wanted; this percentage of unwanted last births increases to 65 percent among women with three or more children. Against this backdrop, Fujimori's plan initially had support from the United Nations (UNFPA), USAID and NGOs, if not from powerful conservative religious forces within Peru. By early 1998, however, claims of sterilizations performed on poor rural women without consent had caused a political uproar in Peru and the controversy spread to the international community. Tubal ligation, a form of female sterilization, was a publicly stated element of the program, but anecdotal evidence suggests that health workers were given large sterilization quotas and often used "bribes," coercion, and even physical force to meet them. [2]

Sterilization quotas were not officially reported by the Fujimori administration and there were no publicly stated guidelines about which populations were targeted by the sterilization campaign. However, Peruvian Demographic and Health Surveys collected in 2000 (DHS IV) and

[1] Address by President Alberto Fujimori at the United Nations, New York, 1999.
[2] A report by Guilia Tamayo from the NGO Flora Tristan published in 1999 provides evidence based on interviews with sterilized women and investigations of rural "health festivals." The post-Fujimori government of Alejandro Toledo also produced reports documenting human rights violations under the Fujimori sterilization program. The Toledo government was, however, reported to be opposed to birth control in general on religious grounds (Boesten, 2007, Vasquez del Aguila, 2006).

2004-2008 (DHS V) asked respondents about their current form of contraception and the date they initiated use.[3] Figure 2.1 shows a dramatic spike in female sterilizations in 1996 and 1997 and an equally dramatic fall by 1998 when the controversy erupted. We will consider 1996-1997 to be the policy window for our analysis.[4] Based on United Nations age-and gender-specific population tables, we estimate that the DHS reports of sterilization from 1996 to 1997 imply that nearly 172,000 women were sterilized in those two years--close to 5 percent of Peruvian women aged 25-49. If we consider the relevant population to be poor women, as reported, the proportion sterilized is much higher.

In this paper we will address three main research questions. Our first goal is to understand who was affected by the Fujimori sterilization policy using the nationally representative random sample of women in the DHS. The second aim of the paper is to estimate the causal impact of the policy on fertility: How many fewer children were born due to the policy? Third, we attempt to understand what, if any, impact a reduction in the counterfactual number of children had on women's employment and on household well-being for those affected by the policy using DHS IV to measure outcomes three years after the policy and DHS V to measure outcomes seven to eleven years after the policy. We tackle these questions sequentially, with each stage feeding into the next. We carefully outline the assumptions behind causal identification at each stage. We also attempt to explore and sign any potential bias in our estimates. We continue to conduct robustness checks that test the assumptions and credibility of our causal claims.

[3] The Demographic and Health Surveys are cross-sectional surveys. So that DHS IV and DHS V do not represent waves of a panel, but rather repeated cross sections.

[4] This evidence from the DHS on the timing of the policy is corroborated by Tamayo (1999) and post-Fujimori Health Ministry reports.

There is considerable debate about the causal direction of the correlation between poverty and family size in developing countries like Peru. Fujimori's own claim of a "vicious circle" points directly to the simultaneity—endogeneity— inherent in the study of the link between family planning and economic development. However, credible evidence is vitally important since Fujimori's logic that Peru needed to reduce family size in order to eliminate poverty was the driving force behind a policy that lead to serious human rights violations. The challenges to identification in evaluating population programs are elaborated in recent papers by (Schultz, 2005) and (Moffitt, 2005). Both of these papers highlight the difficulties of establishing causation in population research, but also the great policy importance of accepting these challenges, being honest and clear about assumptions, and seeking out mechanisms that can explain observed behaviors.

There are at least two major challenges specific to the Peruvian sterilization campaign that we must tackle in order to understand who was affected by the policy and then take the next step of identifying the policy's impacts. First, the details of the policy were secret. Second, there was a non-trivial and slightly increasing rate of female sterilization prior to the advent of the 1996-1997 Fujimori sterilization policy as can be seen in Figure 2.1. This underlying rate of sterilization likely continued during the policy, but we are unable to distinguish directly in the data which women would have been sterilized anyway, and which were sterilized because of the policy. We suspect that sterilizations that were not caused by the policy were voluntary. If some underlying level of sterilization continued during 1996-1997, simply looking at all sterilizations that occurred during the policy window will conflate the impact of potentially voluntary and potentially coerced sterilizations—impacts that we suspect may be quite different. Our methodology aims to tackle both of these challenges using the rich information in the DHS to forensically uncover the characteristics of the population that was targeted by the policy. We use the complete birth histories and detailed geographic information available in the DHS along with timing of sterilization to construct a reweighting estimator along the lines of DiNardo et al. (1996). Our estimator of the treatment effect of the policy is modified, however, to account for the fact that the group of all women who were sterilized during the policy make up a “contaminated” treatment group--in the data we know who was sterilized during the policy period, but among these women we do not know who was treated by the policy.

There is evidence based on hundreds of interviews that women were tricked, pressured, and even physically forced into sterilization procedures in 1996 and 1997 (Tamayo 1999). However, in our data we cannot determine any level of coercion or force during the policy. Furthermore, we cannot confirm that sterilizations that occurred outside of the policy were voluntary. Therefore, going forward, we refrain from using the terms “voluntary” and “coerced” or “forced,” and we distinguish, rather, between sterilizations that we predict would have occurred even in the absence of the program, and those that were caused by the 1996-1997 policy. Given that our methodology is based on predictions of which women were in each category, we are further able to tackle the question of whether the impact of sterilization was different among women targeted by the policy compared to women whose sterilizations were not caused by the policy.

We find that women targeted by the Fujimori sterilization policy were on average 31 years old, had four children at the time of sterilization, and 5.6 years of schooling. We estimate that roughly half of the women treated by the policy lived in rural areas and a quarter were from rural mountain regions, but we also find that a significant proportion of treated women came from urban coastal areas like Lima. We estimate that being sterilized by the policy led these women to have 0.33 fewer children by 2000, and 0.85 fewer children by 2004. We find small and marginally significant impacts of the policy on women's and children's outcomes, with the exception of statistically significant improvements in the height for age and school enrollment of daughters of treated women.

The counterfactual comparisons we use in our estimation procedure rely on the assumption that all of the factors that lead women to be sterilized by the policy are observed and that we have properly controlled for them. This is a strong assumption, one that we continue to examine. For example, we might be concerned that women who were observationally the same as women sterilized by the policy, but who were not sterilized were different in unobserved ways that are correlated with fertility. In particular, we may worry that they are women who had a greater desire for additional children than those who succumbed to the policy. The DHS surveys asked women about the wantedness of all pregnancies in the past five years. We compare responses to these questions in DHS IV between the treatment and control groups created by reweighting and find that the percentages of women who wanted (or did not want) their last pregnancy are not identical, but the reweighting improves the match considerably.[5] We are encouraged by these results and think this is suggestive evidence that while we are matching on observed characteristics, our treatment and control group may also match on unobserved characteristics.

2.2 Methodology to Forensically Estimate the Impact of the Fujimori Sterilization Policy

Our goal is to estimate the effect of being sterilized by the Fujimori Sterilization Policy (FSP) on fertility and on measures of family well-being. Estimating who was treated by the

[5] We only make this comparison for women who had pregnancies in the last five years but before the policy. This restriction is necessitated by the range of data available and our desire to make a proper counterfactual comparison, but limits the sample size. Details can be found in Table 2.

policy is a crucial first step in accomplishing this goal. This is not the usual first step in treatment effects estimation, but it is required in this case because of the secrecy of the policy and the nature of the information we have about sterilized women. Recall, that we know if a woman was sterilized and when she was sterilized. However, we suspect that some of the women sterilized during the policy period were not treated by the policy--they would have been sterilized anyway. In other words, our information on who was sterilized during the policy period is contaminated information on treatment status. Finally, treatment was assigned based on criteria that are not publicly available and those criteria were likely far from random assignment. To motivate the modifications we make to the standard treatment effects estimation, we will begin by outlining the methods we would use if we had either "ideal" data or at least a more typical amount of information about a policy.

We will use notation standard in the treatment effects literature. We define an indicator S to denote whether a woman is sterilized, and an indicator D to denote if a woman is sterilized (treated) by a sterilization policy. We assume that each woman has two potential outcomes, Y_0 if she is not treated and Y_1 if she is treated. We only ever observe one of these potential outcomes, but we are able estimate the average treatment effect on the treated, $E[Y_{i1} - Y_{i0}|D = 1] = \mathrm{E}[\mathrm{Y}_1 |\mathrm{D} = 1] - \mathrm{E}[\mathrm{Y}_{\mathrm{i}0}|\mathrm{D} = 1]$, under different assumptions given the type of data available and the way treatment was assigned.

Random assignment allows the most straightforward estimation strategy. If the only sterilizations that took place during the policy period were those caused by the policy and sterilizations were randomly assigned (and furthermore we had data on who was sterilized), we would simply compare the outcomes of sterilized women to those of non-sterilized women. In this case $S = D$ and $E[Y_{i1}|D = 1]$ is observed--the average outcome among sterilized women.

The average outcome among non-sterilized women $E[Y_i|D=0]$ would be an unbiased estimate of $E[Y_{i0}\}|D=1]$because of random assignment. If there were other sterilizations occurring but we knew who was sterilized by the policy, we would not always have $S=D$ but we would apply the analysis to all women not previously sterilized.

Now consider a scenario where sterilizations were not randomly assigned, but that the only sterilizations taking place were caused by the policy (and we know who was sterilized). Again in this case $E[Y_{i1}|D=1]$ is observed. But now without random assignment of the treatment, the average outcome of non-sterilized women is no longer an acceptable counterfactual. However, if we believe that we observe all of the characteristics that lead to selection into treatment, we can use those factors to estimate $E[Y_{i0}|D=1]$ using the non-sterilized population. This could be done with ordinary least squares using a treatment dummy and the necessary controls. Propensity score methods like matching and reweighting would rely on estimating the probability of treatment based on observable characteristics of treated women. Both methods rely on the assumption of selection into treatment on observables and both use the characteristics of the treated group to create a counterfactual comparison group among the non-treated that resembles the treated group. If there were other sterilizations occurring during the policy period, but we knew which women were sterilized by the policy, the same analysis would be conducted by simply removing women sterilized outside of the program from the sample.[6]

In all of the scenarios described above, $E[Y_{i1}|D=1]$ is observed and the researcher only needs to think of how to find an unbiased estimate of $E[Y_{i0}|D=1]$. In our case, we do not directly observe $E[Y_{i1}|D=1]$ because we do not observe D. So we must estimate both

[6] We focus on a propensity score based reweighting method as it makes the intuition of our process more clear and the weights we create allow us to show the characteristics of the women we hypothesize were in the different categories. In this way we can infer from our analysis who was affected by the policy--the first of our research goals.

$E[Y_{i1}|D=1]$ and $E[Y_{i0}|D=1]$. We will use the rich information available in the DHS on birth and martial histories, geographic and demographic characteristics of women sterilized both before and during the policy to separately identify treated women from women who would have been sterilized in the absence of the program. We will use the information in the DHS to estimate propensity scores for probability of sterilization and proceed with a reweighting strategy. The next subsection describes the assumptions and modifications to standard reweighting techniques that allow us to estimate both $E[Y_{i1}|D=1]$and $E[Y_{i0}|D=1]$--the elements necessary to find the average treatment effect on the treated.

2.2.A A Modified Reweighting Strategy

2.2.A.i Notation and relevant probabilities

We begin by modifying the notation outlined above to accommodate the unique features of our identification strategy. Because we will distinguish between sterilizations that occurred before and during the policy, we now define an indicator variable S_t to denote whether a woman is sterilized during a given period. A woman who is sterilized in period t will have a value of one for S_t and a value of zero otherwise i.e. $s_t \in \{0,1\}$.[7] The index $t \in \{1,2\}$equals one if the time period is prior to the FSP time period (1990-1994), and equals two if the time period coincides with the FSP time window (1996-1997).[8] Since sterilization is a permanent one-time procedure, a woman can only have $S_t = 1$ in one of the periods.[9] We are mute regarding sterilizations that happened after the FSP was dismantled. In other words, we assume that they would have happened regardless of the FSP and thus are included in our control groups.

---

[7] As is conventional, capital letters denote random variables and small letters denote specific realizations of those random variables.
[8] We leave out the calendar year 1995 because it is possible that the FSP might have started in the later part of that year.
[9] In the data we do not find any sterilized women who give birth after the date they report being sterilized.

Our treatment of interest is sterilization by the FSP, denoted by $D, d_t \in \{0,1\}$. Women who were sterilized because of the FSP have a value of one for D, and women who were not sterilized or whose sterilizations were not caused by the FSP have a value of zero. Thus, a woman who was sterilized in 1997 by the FSP would have $S_1 = 0, S_2 = 1, D = 1$ and a woman who was sterilized during the FSP window but would have been sterilized regardless of the policy would have $S_1 = 0, S_2 = 1, D = 0$.[10] Finally, we denote other observed variables by X.

Now we define a series of probabilities we will use in our estimation strategy and the assumptions required to estimate them given our data. Equation gives the probability of a woman with observed characteristics x, becoming sterilized during the FSP time period (given that she was not sterilized before),

$$\begin{aligned} \boldsymbol{P(S_2 = 1|X = x)} &= \boldsymbol{P(S_2 = 1, D = 0|X = x) + P(S_2 = 1, D = 1|X = x)} \\ &= \boldsymbol{P(S_2 = 1, D = 0|X = x) + P(D = 1|X = x)} \qquad (2.1) \end{aligned}$$

The first equality holds because $D = 0$ and $D = 1$ are mutually exclusive and collectively exhaustive events. The second equality holds because $S_2 = 1$ for all cases where $D = 1$. Assumption 1 below allows us to exploit the information we have about women prior to the FSP.

Assumption 1: The probability of a woman being sterilized during the policy period who would have been sterilized even in the absence of the policy is the same as the probability of sterilization before the FSP was implemented for a woman with similar observable characteristics (X).[11] In other words we assume that

[10] Other relevant categories are women who were sterilized prior to the FSP who would have $S\_1 = 1, S_2 = 0, D = 0$ and women who were never sterilized by the end of the FSP would have $S_1 = 0, S_2 = 0$ and obviously $D = 0$.
[11] We allow for a time trend in our estimation to capture the underlying national trend in sterilization take-up prior to the implementation of the FSP.

$$\boldsymbol{P(S_2 = 1, D = 0|X = x) = P(S_1 = 1|X = x)} \tag{2.2}$$

Under Assumption 1 we can re-write Equation (2.1) to express the probability of being treated as:

$$\boldsymbol{P(D = 1|X = x) = P(S_2 = 1|X = x) - P(S_1 = 1, |X = x)} \tag{2.3}$$

To simplify notation we re-write the probability of sterilization in the pre-policy period for a woman with observed characteristics x as $P_1(x)$ and the probability of sterilization during the FSP as $P_2(x)$.[12] Thus Equation (2.3) becomes

$$\begin{aligned} \boldsymbol{P(D = 1|X = x)} &= \boldsymbol{P_2(x) - P_1(x)} \\ &= \Delta \boldsymbol{P(x)} \end{aligned} \tag{2.4}$$

We can also define the probability of being treated conditional on being sterilized during the FSP time period

$$\boldsymbol{P(D = 1|S_2 = 1, X = x) = \frac{P(D = 1|X = x)}{P(S_2 = 1|X = x)} = \frac{\Delta P(x)}{P_2(x)}} \tag{2.5}$$

Finally, we modify our notation regarding the potential outcome of interest Y_d, where d indexes the state of the treatment variable D. The outcome that is realized (and observed) is Y, which is not indexed by d. Since we do not know who was sterilized by the FSP (treated) and who would have been sterilized even in the absence of the FSP (non-treated) among women who were sterilized during the years 1996-1997 (i.e. $S_2 = 1$) the observational rule is modified from the standard case. Equation (2.6) gives the modified expression for the observed outcome in terms of the relevant potential outcomes

[12] $P_2(x)$ and $P_1(x)$ can be estimated in the data using a probit or a logit, and can be thought of as the propensity scores used in matching and weighting estimators.

$$\boldsymbol{Y = s_2[dY_1 + (1 - d)Y_0] + (1 - s_2)Y_0}$$
$$\boldsymbol{Y = s_2\tilde{Y} + (1 - s_2)Y_0} \qquad (2.6)$$

The term $\tilde{Y} = [dY_1 + (1 - d)Y_0]$ highlights the fact that not all women sterilized from 1996-1997 were treated by the policy. In other words, $\tilde{Y}$ represents the outcomes of the contaminated treatment group.

As discussed above, our goal is to estimate average treatment effect on the treated, i.e. the impact of the sterilization on those women who were sterilized by the FSP:

$$\begin{aligned} ATT &= E[Y_1 -_0 | D = 1] \\ &= \mathrm{E}[\mathrm{Y}_1 | \mathrm{D} = 1] - \mathrm{E}[\mathrm{Y}_0 | \mathrm{D} = 1] \end{aligned} \qquad (2.7)$$

2.2.A.ii Reweighting approach to deal with a contaminated treatment group

Given our observational rule in Equation (2.6), we cannot directly estimate the first term of Equation (2.7), $E[Y_1|D = 1]$. However, note that

$$\boldsymbol{E[\tilde{Y}|D = 1] = E[Y_1|D = 1]} \qquad (2.8)$$

We observe $\tilde{Y}$ for women sterilized during the FSP period ($S_2 = 1$) and we can use the probabilities derived above to estimate $E[\tilde{Y}|D = 1]$, under certain assumptions. Note that

$$\boldsymbol{\begin{aligned} E[\tilde{Y}|D = 1] &= \int \tilde{y} f(\tilde{y}|D = 1) d\tilde{y} \\ &= \iint \tilde{y} f(\tilde{y}, x|D = 1) dx d\tilde{y} \end{aligned}}$$

If we multiply and divide the integrand by $f(\tilde{y}, x|S_2 = 1)$ and then apply Bayes Rule to the numerator and denominator we get

$$\boldsymbol{E\left[\tilde{Y}\middle|D=1\right]=\iint \tilde{y}\frac{f(\tilde{y},x|D=1)}{f(\tilde{y},x|S_2=1)}f(\tilde{y},x|S_2=1)dxd\tilde{y}} \tag{2.9}$$
$$\boldsymbol{=\iint \tilde{y}\frac{f(\tilde{y},x|D=1)f(\tilde{y},x)f(S_2=1)}{f(S_2=1|\tilde{y},x)f(\tilde{y},x)f(D=1)}f(\tilde{y},x|S_2=1)dxd\tilde{y}}$$

Next we introduce the standard assumption in matching estimators:

Assumption 2: Strong Ignorability Assumption. We assume that after conditioning on X the probability of being treated and of being sterilized during the years of the FSP are independent of the potential outcomes $\{Y_0, Y_1\}$ and, thus, they are also independent of $\tilde{Y}$. In other words, and invoking equation (2.4), we assume that

$$\boldsymbol{f(D=1|y_0,y_1,x)=f(D=1|\tilde{y},x)=f(D=1,|x)=\Delta\, P(x)}$$
$$\boldsymbol{f(S_2=1|y_0,y_1,x)=f(S_2=1|\tilde{y},x)=f(S_2=1|x)=P_2(x)} \tag{2.10}$$

Using Assumption 2 we can re-express equation (2.9) as

$$E[\tilde{Y}|D=1]=\frac{f(S_2=1)}{f(D=1)}\iint \tilde{y}\frac{f(D=1|x)}{f(S_2=1|x)}f(\tilde{y},x|S_2=1)dxd\tilde{y}$$
$$=\frac{P(S_2=1)}{P(D=1)}\iint \tilde{y}\frac{\Delta P(x)}{P_2(x)}\, f(\tilde{y},x|S_2=1)dxd\tilde{y} \tag{2.11}$$

Since we have a sample from $f(\tilde{y},x|S_2=1)$ we can estimate the expected value in Equation (2.11) with the finite sample estimator: [13]

[13] The finite sample estimator of Equation (2.11) is given by:

$$E[\tilde{Y}a|D=1]=\frac{\widehat{\hat{P}(S_2=1)}}{\hat{P}(D=1)}\frac{\sum_{i=1}^{N}y_i\omega_i s_{2i}}{\sum_{i=1}^{N}s_{2i}}$$

However, note that the population value $\frac{\widehat{\hat{P}(S_2=1)}}{\hat{P}(D=1)}$ can be approximated in finite samples by $\left[\frac{\sum_{i=1}^{N}y_i\omega_i s_{2i}}{\sum_{i=1}^{N}s_{2i}}\right]^{-1}$, which gives the expression in Equation (12).

$$E[\tilde{Y}|\widehat{D=1}] = E\left[Y_1|\widehat{D=1}\right] = \frac{\sum_{i=1}^{N} y_i \omega_{1i} s_{2i}}{\sum_{i=1}^{N} \omega_{1i} s_{2i}} \tag{2.12}$$

$$\omega_{1i} = \frac{\Delta \widehat{P(x_\iota)}}{\widehat{P_2(X_\iota)}} \times \phi_i \tag{2.13}$$

Where ϕ_i is the DHS sampling weight of woman i. Thus, the expected value $E[Y_1|D=1]$ is a weighted average of the observed outcome, Y, for women who were sterilized during the FSP time period. The weights are proportional to the probability that, conditional on being sterilized during that period, the woman was a induced to be sterilized by the policy (see Equation (2.5)).

Invoking the strong ignorability assumption described by Assumption 2 and following similar steps as before, it can be shown that $E[Y_0|D=1]$ is given by:

$$\begin{aligned}
E[Y_0|D=1] &= \int y_0 f(y_0|D=1) dy_0 \\
&= \iint y_0 f(y_0, x|D=1) dx dy_0 \qquad (2.14) \\
&= \iint y_0 \frac{f(y_0, x|D=1)}{f(y_0, x|S_2=0)} f(y_0, x|S_2=0) dx dy_0 \\
&= \iint y_0 \frac{f(D=1|y_0, x) f(y_0, x) f(S_2=0)}{f(T=0, S=0|y_0, x) f(y_0, x) f(D=1)} f(y_0, x|S_2=0) dx dy_0 \\
&= \frac{P(S_2=0)}{P(D=1)} \iint y_0 \frac{\Delta P(x)}{1-P_2(x)} \mathrm{f(y_0, x|S_2=0)dxdy_0}
\end{aligned}$$

And the sample estimator is given by:

$$\boldsymbol{E\left[Y_0|\widehat{D=1}\right] = \frac{\sum_{i=1}^{N} y_i \omega_{0i}(1-s_{2i})}{\sum_{i=1}^{N} \omega_{0i}(1-s_{2i})}} \tag{2.15}$$

$$\boldsymbol{\omega_{0i} = \frac{\Delta \widehat{P(x_\iota)}}{1-\widehat{P_2(x_\iota)}} \times \phi_i} \tag{2.16}$$

Thus, the expected value $E[Y_0|D=1]$ is a weighted average of the observed outcome, Y, for women who were not sterilized during the FSP time period, where the weights allow us to construct a counterfactual control group for the treated group. We can re-write the weights to give them a clearer interpretation:

$$\boldsymbol{\omega_{0i}} = \frac{\widehat{\boldsymbol{P_2(x_\iota)}}}{\boldsymbol{1 - \widehat{P_2(x_\iota)}}} \times \frac{\Delta\widehat{\boldsymbol{P(x_\iota)}}}{\widehat{\boldsymbol{P_2(x_\iota)}}} \times \boldsymbol{\phi_i} \tag{2.17}$$

$$= \frac{\widehat{\boldsymbol{P_2(x_\iota)}}}{\boldsymbol{1 - \widehat{P_2(x_\iota)}}} \times \boldsymbol{\omega_{1i}} \tag{2.18}$$

In this form, it becomes evident that the weights are the result of a two-step (matching) procedure. In the first step, described by the term $\frac{\widehat{P_2(x_\iota)}}{1-\widehat{P_2(x_\iota)}}$, we reweigh the outcomes of women who were not sterilized during the FSP time period by giving higher weights to the outcomes of those women who are observationally more similar to women that were sterilized during the FSP time period. The second step is essentially the same as the reweighting performed before on the $S_2 = 1$ group and thus is described by the term ω_{1i}. In other words, in the second step we give more weight to the outcomes of women with a higher counter-factual probability of being treated by the FSP (given the counterfactual of being sterilized at all during the FSP time period).

By joining the results of Equations (12) and (15), we can estimate the ATT using:

$$\widehat{\boldsymbol{ATT}} = \frac{\sum_{i=1}^{N} \boldsymbol{y_i \omega_{1i} s_{2i}}}{\sum_{i=1}^{N} \boldsymbol{\omega_{1i} s_{2i}}} - \frac{\sum_{i=1}^{N} \boldsymbol{y_i \omega_{0i} (1 - s_{2i})}}{\sum_{i=1}^{N} \boldsymbol{\omega_{0i} (1 - s_{2i})}} \tag{2.19}$$

We have focused so far on the impact of sterilizations that were caused by the FSP, but we could also, for comparison purposes, be interested in the impact of sterilizations that occurred outside of the policy. Using the same procedures as before, it can be shown that the ATT for sterilizations that occurred during 1996-1997 but would have occurred even in the absence of the policy can be estimated using equation ():

$$\boldsymbol{E\left[Y_1 - Y_0 \middle| \widehat{S_2 = 1}, D = 0\right] = \frac{\sum_{i=1}^{N} y_i \widetilde{\omega}_{1i} s_{2i}}{\sum_{i=1}^{N} \widetilde{\omega}_{1i} s_{2i}} - \frac{\sum_{i=1}^{N} y_i \widetilde{\omega}_{0i}(1 - s_{2i})}{\sum_{i=1}^{N} \widetilde{\omega}_{0i}(1 - s_{2i})}} \qquad (2.20)$$

where $\widetilde{\omega}_{1i} = \frac{\widehat{P_1(x_i)}}{\widehat{P_2(x_i)}} \times \phi_i$ and $\widetilde{\omega}_{0i} = \frac{\widehat{P_1(x_i)}}{1-\widehat{P_2(x_i)}} \times \phi_i$.

2.3 Data: Peruvian Demographic and Health Surveys

We investigate the Peruvian sterilization policy using the fourth, and fifth waves of the Demographic and Health Surveys for Peru (hereafter DHSIV, and DHSV.) The Demographic and Health Surveys are nationally representative cross sectional surveys. Both DHSIV and DHSV were conducted after the policy had ended and thus allow us to look at potential impacts on fertility and other household outcomes. DHS IV was conducted in 2000 and has a sample size of 27,843 women aged 15-49; and DHS V was collected continuously over the course of 2004 to 2008 and has a sample size of 41,648 women. The primary advantage of the survey for addressing our research questions is the information collected on birth control methods including sterilization and the date when the sterilization occurred. The surveys also include detailed birth histories and information on place of residence. Our analysis sample includes all women who were eligible to be sterilized during the policy period 1996 to 1997-- ever-married women who had at least one child and who were not previously sterilized[14]--giving us a sample size of 14,430 eligible women in DHSIV, 707 of whom were sterilized during the policy period; and 16,673 eligible women in DHS V, 735 of whom were sterilized during the policy.

[14] There was an existing law prior to the policy requiring spousal consent for sterilization (Coe, 2004) and all women who report sterilizations in both DHS were married (or had previously been married) and had at least one child at the time of sterilization.

We estimate the impact of the FSP on fertility as well as on household outcomes. To measure fertility we use the number of surviving children in a given year. We can also use the birth histories to measure number of children ever born and child mortality, which are alternative outcomes we plan to pursue. Next we look at the impact of the policy on women's outcomes. The DHS has limited information on labor force outcomes, but we make use of a question asking whether the respondent is currently working and we use this as a proxy for labor force participation. We also estimate the impact on reports of domestic violence in the last 12 months as sterilization could impact a woman's bargaining power relative to her spouse.

We examine several outcomes of household children to test whether the policy impacted well-being as measured through health and education. We want to compare children whose mothers were sterilized--and therefore had no more siblings--to counterfactual children whose mothers were not sterilized and therefore likely had younger siblings. This kind of comparison would allow us to test a quality/quantity trade-off. Therefore, we only look at outcomes for children born prior to the policy. Weight for height and height for age was collected for all children age four and under, so we are restricted to DHS IV for this outcome since all children under four were born after the policy by the time the DHS V survey began in 2004. In both DHS IV and DHS V we measure years of schooling and current enrollment (controlling for age) of household children under 15. In the DHS V we can also examine the education level of girls over 15 who are old enough to be survey respondents and but were children at the time of the policy. Having fewer younger siblings to help care for could have allowed girls to stay in school.

One limitation of the Peruvian DHS is that it does not contain accurate information on respondents' ethnic group. One of the claims of human rights activists is that the Fujimori policy targeted indigenous women from the Quechua or Aymara groups. DHS IV and V ask respondents their language among which Quechua or Aymara are choices. But only 15% of eligible women DHS V responded that they spoke either of these languages. This variable clearly does not accurately measure ethnicity, as the Amerindian population is closer to 40% of the Peruvian population.

2.3.A Descriptive Analysis of Sterilization in Peru 1990-1998

Table 2.1 presents summary statistics from DHS IV and DHS V (and from DHS II (1990-1002) prior to the policy) relating to fertility highlighting the strong negative correlation between family size and income (proxied by education) or wealth. We see that going from the highest to the lowest levels of mother's education doubles the number of living children for mothers over 40 from 2.2 children to more than 5.4 in 2000. This contrast is similarly strong across the wealth index. Rural households have substantially more children than urban households. Comparing number of children across the two surveys we see that fertility decreased at all education and wealth levels from 2000 to the 2004-2008 period.

Figure 2.1shows the number of sterilizations reported in DHS IV and DHS V by year, and confirms the sharp increase in sterilizations during the policy period. In the analysis that follows we will consider 1996 and 1997 to be the aggressive sterilization "policy period." Using age- and gender-specific population estimates for Peru from the United Nations Population Division, we can estimate the number of actual sterilizations implied by the self-reported sterilizations in the nationally representative DHS surveys. Based on the UN estimated Peruvian population of women age 15 to 49 in 2000, the DHS IV is a 0.41 percent sample of the relevant population. Based on this sampling scale, the 417 sterilization reported in 1997 (representing 408 women when weighted) imply that 99,430 women were sterilized in that year, which is remarkably similar to the numbers reported by Fujimori's opponents. If we sum together all of the DHS IV reported sterilizations that occurred during the supposed policy period from 1996 to 1998, we estimate that roughly 218,626 women were sterilized. This implies that approximately 3.4 percent of women age 15 to 49 were sterilized, or 4.5 percent of women age 20 to 45 who were in their prime fertility years. If certain regional or demographic characteristics were specifically targeted the percentage of the relevant population that was sterilized could be much higher. There are some notable differences between the DHS IV and DHS V surveys in the reporting on female sterilization that occurred during the 1990s. Looking at Figure 1 we see that the increase in sterilizations from the pre-policy period to the policy period was more gradual in DHS V and less abrupt than in DHS IV, and similarly more gradual for the decrease in sterilizations after the policy ended. Part of the difference could be recall bias given that the DHS V survey took place eight to 12 years after the policy, while DHS IV was conducted only three years from the peak of the policy. DHS V collected a smaller nationally representative sample in each of the five years of the survey, which ranged from a 0.07 percent sample of the age-/gender-specific population in 2004, to a 0.2 percent sample in 2008. The implied number of sterilizations in 1997 based on DHS V reports is 79,752 and the total number of sterilizations over from 1996 to 1998 is 179,352, or 2.8 percent of women aged 15 to 49. It is possible that some of the sterilizations that actually occurred during the policy were mistakenly reported to occur in the year before or after the policy ended in DHS V. If this were the case, we would expect to find muted treatment effects using DHS V data. These discrepancies deserve further investigation.

2.3.B Estimating the probability of treatment by the Fujimori sterilization policy using DHS IV and DHS V

We estimate the probability of sterilization in the pre-policy and policy periods using a pseudo panel constructed from the cross-sectional data in the DHS surveys. These probabilities, conditional on observable characteristics, are the propensity scores $P_1(x)$ and $P_2(x)$ described in the methodology section. We use the date of sterilization and other variables to construct a longitudinal history for each woman describing her fertility and marital time path from the beginning of what we consider the pre-policy period, 1990 to the end of the policy period in 1998. Each woman has one observation for each year and dummy indicating whether she is sterilized in each year. Once she is sterilized she has no further observations. Using this type of quasi panel allows us to estimate the conditional probability of being sterilized in each year—the annual hazard of being sterilized given that one has not been sterilized up that point.[15] This approach takes account of the fact that a woman sterilized in 1997 was at risk of being sterilized in all previous periods and as such should be included in calculating the probability of being sterilized in 1994, for example. Furthermore, because of detailed birth and (somewhat) detailed marital histories provided in the DHS surveys we can use richer information about spacing of children in the quasi panel than in a cross sectional estimation of probability of sterilization by year. We also include 56 regional categories starting with Peru's 25 departments and further differentiating by geography (jungle, mountain, coastal) and by urban and rural status. Other covariates are age, number and age of children, number of boys, infant mortality, age at first birth, and education. Finally, using the pseudo panel we are able to include a time trend in the logit to account for secular changes in fertility and sterilization that could be occurring within each period. We estimate the probability of sterilization in each period using a logit.

When we calculate $\Delta\hat{P}(x) = \widehat{P_2}(x) - \widehat{P_1}(x)$ as in Equation (2.4) in some cases the value is negative leading $\hat{P}(x)$ $(D = 1|S_2 = 1, X = x$ (Equation(2.5)) to be less than zero. Since this

[15]This is based on a extension of proportional hazard models to discrete time proposed by Cox (1972). Estimating a logit regression on a set of pseudo observations generated from a cross-section amounts to fitting a discrete-time proportional-hazards model. See Allison (1982) and notes on this by (Rodriguez, 2007): http://data.princeton.edu/wws509/notes/c7.pdf.

object is a probability, negative values are not defined and we set these values to zero. The rationale is that such women, if sterilized, had a zero probability of being treated by the policy. The number of observations for which we make this adjustment is noted in the results tables.

2.4 Estimating the Impact of the Fujimori Sterilization Policy

2.4.A Characteristics of women targeted by the FSP

Table 2.2, based on DHS IV, and

Table 2.3, based on DHS V, show the characteristics of the sample of eligible women before and after reweighting. The first three columns of Table 2.2 and 2.3 give the characteristics of the non-reweighted sample of women eligible to be sterilized during the policy—ever married women with at least one child who were not previously sterilized--separated by whether they were sterilized during the policy period. The findings based on DHS IV and DHS V are similar, so we will summarize them jointly. Column 2 gives the characteristics of what we have called the contaminated treatment group which includes both women treated by the FSP and women who were not treated by the policy and would have been sterilized even the absence of the program. We see that sterilized women are older, have more children, slightly less education than the average eligible woman, but that a roughly similar proportion of women sterilized during the policy live in rural areas.

In columns 4 and 5, we apply the weights described in the methodology section to create the effective treatment and control groups we use to estimate the impacts of the policy. The group of women who were sterilized during 1996-1997 are reweighted to represent only the group of women who were treated by the policy (i.e. we use weightω_0). The non-sterilized women are reweighted to match the observable characteristics of the treated women (i.e. we use weight ω_1). The differences compared to column 2 are striking. The women we estimate to be in the treatment group are younger, considerably less educated, and much more likely to live in rural areas than the contaminated treatment group suggested. The last two columns of Table 2.2 and 2.3, specifically highlight the differences between women who were sterilized by the policy and sterilized women who we estimate would have been sterilized even in the absence of the policy. Women sterilized outside of the policy are considerably more educated and more likely to live in urban areas, though they do not have substantially fewer children. If we suspect there may be heterogeneous treatment effects of sterilization by these characteristics, then these columns confirm the benefit of our method in separating these two types of sterilized women. Column 6 gives the demographic characteristics of the women we estimate were affected by the Fujimori Sterilization Policy and thus provide an answer to our first research question: who was targeted by the Fujimori Sterilization Policy? We estimate that women targeted by the policy were on average 31 years old, had four children at the time of sterilization, and 5.6 years of schooling. Their average age at first birth was 19. Roughly half of these women lived in rural areas and a quarter were from rural mountain regions, but we also find that a significant proportion of treated women came from urban coastal areas like Lima.

2.4.B Impact of the FSP on fertility

Table 2.4 shows the estimated impact of the FSP on fertility. The following pattern will be used in all of the subsequent results tables (unless otherwise noted): The first three columns give results based on DHS IV which was collected three years after the policy. The next section of three columns are based on DHS V which was collected seven to eleven years after the policy. The first column in each section gives the results of the standard reweighing estimation that only reweights observations of women not sterilized during the FSP. In other words these standard reweighting estimates do not account for the contamination of the treatment group and conflate the impact of sterilization on women treated by FSP and those who would have been sterilized even in the absence of the FSP. The second column gives our preferred specification-- the estimated ATT for women sterilized by the FSP based on the reweighting technique described in the methodology section (in notation these women have $S_2 = 1, D = 1$). Finally the third column is the estimated impact of sterilization on outcomes for women who were sterilized during the policy period but were not treated by the FSP (in notation these women have $S_2 = 1, D = 0$).

In column 2, based on DHS IV, we estimate that by 2000, women treated by the FSP had 0.33 fewer children than the non-sterilized control group , and in column 5, based on DHSV, we estimate that by 2004, treated women had 0.85 fewer children. These estimates of the impact on fertility are larger than the standard reweighting estimates in columns 1 and 4, suggesting that the policy had a stronger impact on fertility among treated women than among women who would have been sterilized anyway. This is shown by the estimates in columns 3 and 6 which give the estimated impact of sterilization on women sterilized outside of the FSP. We find that women sterilized outside of the FSP had 0.22 fewer children by 2000, and 0.58 fewer children by 2004 than relevant counterfactual women. All of the coefficients in Table 2.4 are negative, but could also be expressed positively as the number of additional children born to women in the control group(s). We think these results are large but plausible given the age and existing fertility of the treated women, the amount of time since the policy, and limited access to contraception available in Peru.

2.4.C Impact of the FSP on household outcomes

The remaining tables provide estimates of the impact of the FSP on women's, and children's outcomes. Since in the previous section we estimate that the policy led to a substantial decrease in fertility, we can hypothesize that any impacts on other outcomes were the result of lowered fertility. However, at this point we cannot rule out impacts of the nature of the policy itself, for example the trauma of a coercive act, on outcomes. In this summary of results, we will focus on the estimated impact of the FSP in columns 2 and 5 of the table. Table 2.5 provides estimates of the impact of the policy on women's labor force participation. Column 2 shows an increase in probability of working of five percent in 2000 based on reweighting which is significant at a 10 percent level. However, there is no significant impact on working by DHS V in 2004-2008 as seen in column 6. Table 2.6 shows estimates for the binary outcome of experiencing domestic violence (either physical or sexual) in the last 12 months (this information is only available in DHS V). We estimate that being sterilized by the FSP increased the likelihood of experiencing domestic violence by 5 percentage points.

Table 2.3 shows that 13 percent of eligible women report domestic violence in the last 12 months. We need to further explore the mechanisms of this estimated impact, but changes in ability to bear children may impact women's bargaining power within the household. We view our estimated impact on domestic violence with caution, however, as it could be the case that women susceptible to domestic violence could also be those more susceptible to a coercive government policy.

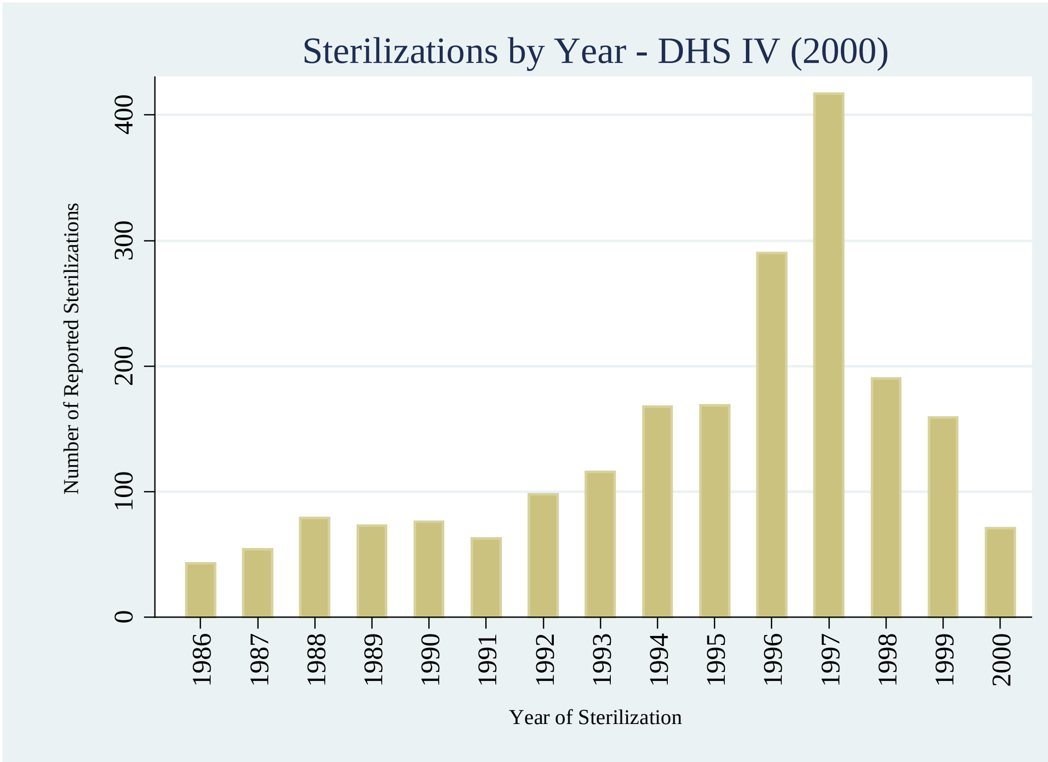
Turning to the impact of the FSP on the children of sterilized women, in summary we find small and mostly non- or marginally significant impacts when we combine girls and boys. Table 2.7 gives the estimated impact on biometric measures of weight for height and height for age (in standard deviations from the reference median) based on DHS IV among children under four who were born prior to the policy. This table only shows results for DHS IV since children born prior to the policy in DHS V are over four years old, and the surveys only record biometric information for children under four. Point estimates on weight for height in column 5 are small and none are statistically significant. Impacts on height for age in column 2, which is a longer term measure of health, are somewhat larger, but again, for the most part, not significant. When we examine girls separately in Table 2.8, however, we find that daughters of women treated by the FSP had height for age that was 0.29 standard deviations greater than counterfactual girls. This estimate is significant at a one percent level. Looking at column 3, we see that there is a similar, though not significant positive impact on daughters of women sterilized outside of the FSP.

In Table 2.9 and Table 2.10, we find a small but significant positive impact on years of schooling and enrollment for children under age 15 in DHS IV three years after the policy.. While the magnitude of the impact is similar for DHS V, seven to 11 years after the policy, the estimates are not statistically significant. When we look just at girls' enrollment in Table 2.11, we find that there is a 2.3 percentage point increase in school enrollment for girls of women sterilized by the FSP about double the impact found when we combined boys and girls. These small impacts are likely due to the high levels of enrollment of primary school children in Peru, even in rural areas, leaving little margin to increase schooling for these ages. However, we find no impact on education of older girls in Table 2.12 based on DHS V. These are girls who are aged 15 to 22 and are respondents to DHS V as adults, but were children at the time of the policy. The fact that their mothers were sterilized could mean they had fewer young siblings to help care for than counterfactual girls and were able to stay in school longer. We do not find evidence of this kind of impact. However, if this impact accrued to girls who then moved out of the house younger, we will not be able to measure the effect.

2.5 Conclusion

There is a continuing debate about the causal link between access to family planning and reductions in fertility in both the developed and developing world. Beyond any direct impact on the level of fertility, access to contraception clearly allows women to control the timing of fertility, which reduces constraints on choices about work and caring for existing children. Recent research in both the United States (Bailey, 2006) and Columbia (Miller, 2010) uses plausibly exogenous variation in access to show that contraception significantly increases female educational attainment and labor force participation by allowing women to delay first births. Our preliminary findings in Peru seem to confirm that the mere reduction of fertility that is not necessarily associated with substantial improvements in welfare in the context of potentially coerced sterilizations. We are finding that when birth control is imposed, the benefits of making choices about fertility may not accrue to women and their households. While we do find small improvements in height for age and school enrollment for girls whose mothers were sterilized by the Fujimori sterilization policy, in general the substantial decrease in fertility caused by the policy does not seem to be associated with substantial improvements in family well-being.

Figure 2.1. Number of Reported Sterilizations by Year
Peruvian Demographic and Health Surveys



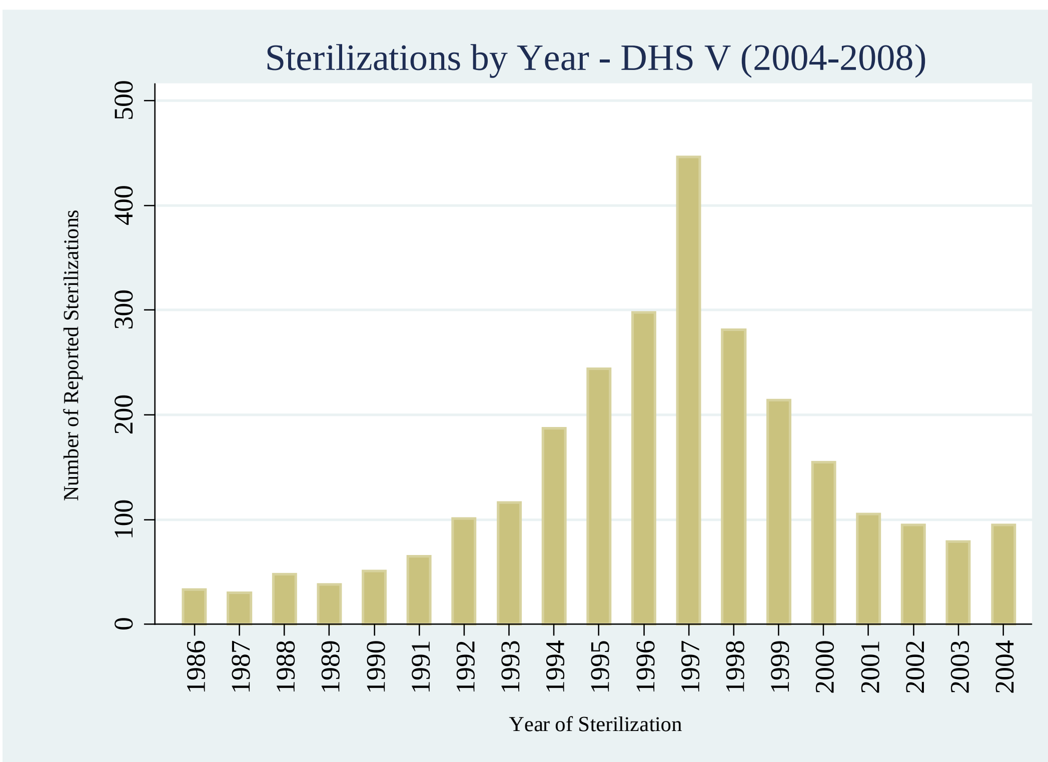


Table 2.1. Number of Living Children for Mothers over 40 by Demographic Characteristics

	DHS II (1991-1992)		DHS IV (2000)	
	Mean	Std Dev	Mean	Std Dev
By Mother's Education				
no education	5.64	2.60	5.43	2.48
primary	4.42	2.34	4.80	2.39
secondary	3.24	1.79	3.22	1.80
higher	2.61	1.60	2.16	1.35
Total	4.39	2.55	3.91	2.37
By Qntls of HH Wealth Index				
lowest q	6.08	2.75	5.73	2.56
second q	5.43	2.54	5.05	2.33
middle q	4.80	2.38	4.06	2.25
fourth q	3.87	2.22	3.28	1.89
highest q	2.94	1.72	2.59	1.62
Total	4.39	2.55	3.84	2.35

Note: Using Sampling Weight Provided by DHS

Table 2.2. DHS IV : Characteristics of Eligible Women - Reweighted Estimates of Characteristics of Treated Women

	Among All Eligible Women					Among Women Sterilized 1996-1997 $S_2 = 1$	
	Not Reweighted		Reweighted to represent D=1			Reweighted	
	(1)	(2)	(3)	(4)	p-value for zero diff	(5)	(6)
	$S_2 = 1$	$S_2 = 0$	$S_2 = 1$	$S_2 = 0$		$D = 1$	$D = 0$
Pre-Policy Characteristics							
Age in 1996	32.36	31.44	31.24	31.05	0.53	31.24	32.62
# Kids in 1996	3.91	2.81	4.08	3.91	0.08	4.08	3.81
Years of education	7.12	7.89	5.65	5.88	0.33	5.65	7.95
Age at first birth	20.08	20.72	19.18	19.28	0.60	19.18	20.52
rural	0.35	0.36	0.49	0.48	0.84	0.49	0.25
coast	0.58	0.51	0.48	0.47	0.77	0.48	0.72
mountain	0.26	0.36	0.33	0.34	0.85	0.33	0.16
jungle	0.15	0.13	0.18	0.19	0.85	0.18	0.13
urban coast	0.48	0.45	0.37	0.38	0.87	0.37	0.62
rural coast	0.10	0.06	0.11	0.10	0.46	0.11	0.10
urban mountain	0.09	0.13	0.08	0.08	0.85	0.08	0.05
rural mountain	0.17	0.24	0.25	0.26	0.76	0.25	0.10
urban jungle	0.07	0.06	0.06	0.07	0.77	0.06	0.08
rural jungle	0.08	0.07	0.12	0.12	0.97	0.12	0.05
Outcomes and other variables							
Wanted last Prenancy[1]	0.30	0.47	0.28	0.35	0.04	0.28	0.30
Wanted last Prenancy Later	0.14	0.22	0.16	0.15	0.71	0.16	0.14
Did not Want last Prenancy	0.55	0.31	0.57	0.51	0.12	0.57	0.56

Notes: Columns 1 and 2 are the non-reweighted sample of women eligible to be sterilized during the policy window. $S_2 = 1$were sterilized (contaminated treatment group) and $S_2 = 0$ were not. In Columns 4 and 5, we apply the weights described in the methodology section to create our preferred treatment and control groups. Women who were sterilized during 1996-1997 are reweighted to represent only the women who were treated by the policy using weight ω_0. Non-sterilized women are reweighted using weight ω_1.Columns 5 and 6 reweight sterilized women to compare those treated by the policy to those who we estimate would have been sterilized even in the absence of the FSP. [1]Among pregnancies that occurred prior to 1997.

Table 2.3. DHS V : Characteristics of Eligible Women - Reweighted Estimates of Characteristics of Treated Women

	Among All Eligible Women					Among Sterilized women	
	Not Reweighted			Reweighted to represent D=1		Reweighted	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	all	$S_2 = 1$	$S_2 = 0$	$S_2 = 1$	$S_2 = 0$	D = 1	D = 0
Pre-Policy Characteristics							
Age in 1996	29.09	31.11	29.00	29.75	29.68	29.75	30.89
# Kids in 1996	2.51	3.71	2.45	3.70	3.73	3.70	3.86
Years of education	7.96	6.88	8.01	5.23	5.41	5.23	6.81
Age at first birth	20.53	20.31	20.53	19.97	19.82	19.97	20.11
rural	0.34	0.36	0.34	0.56	0.53	0.56	0.36
coast	0.50	0.54	0.50	0.38	0.37	0.38	0.55
mountain	0.37	0.32	0.37	0.42	0.43	0.42	0.30
jungle	0.13	0.15	0.13	0.20	0.20	0.20	0.15
urban coast	0.45	0.44	0.45	0.25	0.26	0.25	0.42
rural coast	0.05	0.10	0.05	0.13	0.11	0.13	0.13
urban mountain	0.14	0.11	0.15	0.10	0.11	0.10	0.12
rural mountain	0.23	0.20	0.23	0.32	0.32	0.32	0.18
urban jungle	0.07	0.08	0.06	0.09	0.09	0.09	0.10
rural jungle	0.06	0.06	0.06	0.11	0.11	0.11	0.05
Outcomes and other variables							
Domestic Violence in last 12 months	0.13	0.15	0.13	0.18	0.14	0.18	0.13
Labor Force Participation	0.73	0.72	0.73	0.77	0.75	0.77	0.69
Wealth Index	0.49	0.44	0.49	0.08	0.00	0.08	0.39

Notes: See notes to Table 2.2.

Table 2.4. Fertility Impact - Number of Children

	DHS IV (2000)			DHS V (2004-2008) - # of Kids in 2004		
	Standard Reweighting	Modified Reweighting		Standard Reweighting	Modified Reweighting	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0
S_2=1	-0.275***	-0.327***	-0.218***	-0.693***	-0.846***	-0.582***
	(0.013)	(0.017)	(0.013)	(0.021)	(0.025)	(0.023)
Age	-0.011***	-0.011***	-0.012***	-0.029***	-0.026***	-0.028***
	(0.001)	(0.002)	(0.001)	(0.003)	(0.004)	(0.003)
Years of Schooling	0.003	0.005	0.000	-0.031***	-0.026**	-0.030***
	(0.007)	(0.008)	(0.006)	(0.008)	(0.010)	(0.009)
# of children in 1997	1.007***	1.004***	1.009***	1.020***	1.010***	1.023***
	(0.005)	(0.006)	(0.006)	(0.010)	(0.011)	(0.012)
Age at first birth	0.003	0.001	0.004**	0.007	0.006	0.007
	(0.002)	(0.003)	(0.002)	(0.004)	(0.005)	(0.005)
Geographic controls	yes	yes	yes	yes	yes	yes
Constant	0.543***	0.601***	0.482***	1.534***	1.531***	1.490***
	(0.084)	(0.128)	(0.080)	(0.188)	(0.274)	(0.178)
R-squared	0.958	0.948	0.969	0.893	0.892	0.905
Observations	14430	14430	14430	16673	16673	16673
Observations with positive $\Delta P(x)$	NA	10607	10607	NA	14509	14509

Notes: The first three columns give results based on DHS IV the next three columns are based on DHS V. The first column in each section gives the results of the standard reweighing estimator that only weights the non-sterilized women. The second column gives our preferred specification based on the reweighting technique described in the methodology section. The third column is the estimated impact of sterilization on outcomes for women who were sterilized during the policy period but were not treated by the FSP. Robust standard errors in parentheses. ***p<0.01. **p<.05, *p<0.1

Table 2.5. Impact on Women's Labor Force Participation

	DHS IV (2000)			DHS V (2004-2008)		
	Standard Reweighting	Modified Reweighting		Standard Reweighting	Modified Reweighting	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0
S_2=1	0.022	0.049*	-0.009	-0.007	0.013	-0.023
	(0.025)	(0.027)	(0.030)	(0.026)	(0.026)	(0.033)
Age	0.011***	0.012***	0.010***	0.000	0.004	-0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Years of Schooling	0.016	0.002	0.031**	0.007	-0.001	0.007
	(0.010)	(0.012)	(0.012)	(0.011)	(0.012)	(0.013)
# of children in 1997	-0.022**	-0.034***	-0.008	0.005	-0.006	0.009
	(0.010)	(0.012)	(0.011)	(0.011)	(0.013)	(0.013)
Age at first birth	-0.012***	-0.011**	-0.014***	-0.003	-0.005	-0.004
	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Geographic controls	yes	yes	yes	yes	yes	yes
Constant	0.658***	0.370**	0.760***	1.005***	0.913***	1.037***
	(0.172)	(0.187)	(0.169)	(0.115)	(0.127)	(0.149)
R-squared	0.120	0.146	0.113	0.075	0.112	0.072
Observations	14430	14430	14430	16673	16673	16673
Observations with positive $\Delta P(x)$	NA	10607	10607	NA	14509	14509

Notes: See notes to Table 2.4. Robust standard errors in parentheses. ***p<0.01. **p<.05, *p<0.1

Table 2.6. Impact on Domestic Violence (reported in the last 12 months)

	DHS V (2004-2008)		
	Standard Reweighting	Modified Reweighting	
	(1)	(2)	(3)
Treatment	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0
S_2=1	0.014	0.050*	-0.010
	(0.021)	(0.027)	(0.023)
Age	-0.003	-0.003	-0.003
	(0.003)	(0.004)	(0.003)
Years of Schooling	-0.002	-0.006	-0.001
	(0.008)	(0.011)	(0.009)
# of children in 1997	0.013	0.020	0.009
	(0.011)	(0.015)	(0.013)
Age at first birth	-0.004	-0.004	-0.004
	(0.003)	(0.005)	(0.004)
Geographic controls	yes	yes	yes
Constant	0.120*	0.094	0.167**
	(0.072)	(0.092)	(0.082)
R-squared	0.054	0.090	0.056
Observations	13381	13381	13381
Observations with positive $\Delta P(x)$	NA	11825	11825

Notes: The first column gives the results of the standard reweighing estimator that only weights the non-sterilized women. The second column gives our preferred specification based on the reweighting technique described in the methodology section. The third column is the estimated impact of sterilization on outcomes for women who were sterilized during the policy period but were not treated by the FSP. Robust standard errors in parentheses. ***p<0.01. **p<.05, *p<0.1

Table 2.7. Impact on Children's Biometrics - DHS IV (Kids=<4 born prior to policy)

	Height for Age (in sd from reference median)			Weight for Height (in sd from reference median)		
	Standard Reweighting	Modified Reweighting		Standard Reweighting	Modified Reweighting	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0
S_2=1	0.136	0.118	0.141	-0.067	-0.048	-0.080
	(0.099)	(0.110)	(0.108)	(0.080)	(0.091)	(0.092)
Age	0.043***	0.037**	0.045***	-0.010	-0.011	-0.009
	(0.014)	(0.017)	(0.015)	(0.011)	(0.015)	(0.011)
Years of Schooling	0.062	0.052	0.057	-0.025	-0.045	-0.017
	(0.040)	(0.049)	(0.042)	(0.031)	(0.041)	(0.030)
# of children in 1997	-0.154***	-0.145**	-0.164***	0.038	0.052	0.023
	(0.050)	(0.061)	(0.050)	(0.047)	(0.058)	(0.041)
Age at first birth	-0.027	-0.013	-0.030*	0.004	-0.012	0.009
	(0.019)	(0.024)	(0.018)	(0.014)	(0.018)	(0.013)
Geographic controls	yes	yes	yes	yes	yes	yes
Controls for child's age	yes	yes	yes	yes	yes	yes
Constant	-1.006*	-0.927	-0.547	1.165**	0.523	0.995**
	(0.568)	(0.599)	(0.522)	(0.484)	(0.335)	(0.427)
R-squared	0.318	0.315	0.316	0.143	0.146	0.164
Observations	2899	2899	2898	2899	2899	2898
Observations with positive $\Delta P(x)$	NA	2160	2160	NA	2160	2160

Notes: The first column in each section gives the results of the standard reweighing estimator that only weights the non-sterilized women. The second column gives our preferred specification based on the reweighting technique described in the methodology section. The third column is the estimated impact of sterilization on outcomes for women who were sterilized during the policy period but were not treated by the FSP. Robust standard errors in parentheses. ***p<0.01. **p<.05, *p<0.1

Table 2.8. Impact on Girls Biometrics - DHS IV (Girls=<4 born prior to policy)

	Height for Age (in sd from reference median)			Weight for Height (in sd from reference median)		
	Standard Reweighting	Modified Reweighting		Standard Reweighting	Modified Reweighting	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0
S_2=1	0.276**	0.285**	0.228	-0.026	-0.047	-0.029
	(0.129)	(0.131)	(0.144)	(0.106)	(0.119)	(0.124)
Age	0.040**	0.045**	0.036	0.006	0.021	-0.001
	(0.020)	(0.020)	(0.024)	(0.015)	(0.021)	(0.015)
Years of Schooling	0.052	0.002	0.074	-0.038	-0.035	-0.035
	(0.052)	(0.058)	(0.062)	(0.043)	(0.051)	(0.046)
# of children in 1997	-0.174***	-0.162***	-0.195***	-0.001	-0.034	0.003
	(0.056)	(0.058)	(0.070)	(0.061)	(0.076)	(0.054)
Age at first birth	-0.032	-0.026	-0.034	0.007	-0.021	0.014
	(0.022)	(0.026)	(0.027)	(0.019)	(0.028)	(0.017)
Geographic controls	yes	yes	yes	yes	yes	yes
Controls for child's age	yes	yes	yes	yes	yes	yes
Constant	-0.613	-1.167**	-0.048	-0.058	0.077	0.241
	(0.421)	(0.554)	(0.550)	(0.382)	(0.421)	(0.477)
R-squared	0.352	0.376	0.345	0.212	0.267	0.219
Observations	1457	1457	1457	1457	1457	1457
Observations with positive $\Delta P(x)$	NA	1084	1084	NA	1084	1084

Notes: See notes to Table 2.7. Robust standard errors in parentheses. ***p<0.01. **p<.05, *p<0.1

Table 2.9. Impact on Years of Schooling for Own Children ages 5-14 (born prior to policy)

	DHS IV (2000)			DHS V (2004-2008)		
	Standard Reweighting	Modified Reweighting		Standard Reweighting	Modified Reweighting	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0
S_2=1	0.084**	0.093**	0.077	0.032	0.108	-0.034
	(0.039)	(0.041)	(0.050)	(0.085)	(0.094)	(0.103)
Age	0.042***	0.040***	0.046***	0.026**	0.043**	0.016
	(0.006)	(0.008)	(0.006)	(0.013)	(0.018)	(0.015)
Years of Schooling	0.046***	0.060***	0.024	0.122***	0.138***	0.110***
	(0.016)	(0.018)	(0.021)	(0.036)	(0.048)	(0.040)
# of children in 1997	-0.128***	-0.116***	-0.143***	-0.179***	-0.204***	-0.161***
	(0.018)	(0.022)	(0.019)	(0.038)	(0.052)	(0.042)
Age at first birth	-0.019***	-0.014	-0.023***	-0.009	-0.037*	0.013
	(0.007)	(0.010)	(0.007)	(0.016)	(0.021)	(0.017)
Geographic controls	yes	yes	yes	yes	yes	yes
Controls for child's age	yes	yes	yes	yes	yes	yes
Constant	-0.169	-0.383	-0.291	0.230	1.930***	0.294
	(0.383)	(0.265)	(0.216)	(0.820)	(0.508)	(0.569)
R-squared	0.807	0.792	0.826	0.366	0.329	0.431
Observations	22520	22513	22513	14021	14016	14016
Observations with positive $\Delta P(x)$	NA	18886	18886	NA	12329	12329

Notes: See notes to Table 2.4. Robust standard errors in parentheses. ***p<0.01. **p<.05, *p<0.1

Table 2.10. Impact on School Enrollment for Children ages 5-14 (born prior to policy)

	DHS IV (2000)			DHS V (2004-2008)		
	Standard Reweighting	Modified Reweighting		Standard Reweighting	Modified Reweighting	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0
S_2=1	0.005	0.016**	-0.009	0.001	0.009	-0.005
	(0.007)	(0.008)	(0.008)	(0.009)	(0.009)	(0.011)
Age	-0.000	-0.000	-0.001	0.002**	0.002	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years of Schooling	0.009***	0.011***	0.007**	0.008**	0.007	0.007*
	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
# of children in 1997	-0.004	-0.005	-0.003	-0.009***	-0.006	-0.011***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Age at first birth	0.002	0.001	0.002	-0.000	0.000	-0.000
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Geographic controls	yes	yes	yes	yes	yes	yes
Controls for child's age	yes	yes	yes	yes	yes	yes
Constant	-0.032	-0.099	-0.037	0.964***	0.958***	0.962***
	(0.071)	(0.103)	(0.058)	(0.051)	(0.036)	(0.052)
R-squared	0.649	0.632	0.681	0.126	0.128	0.150
Observations	22554	22554	22547	14021	14021	14016
Observations with positive $\Delta P(x)$	NA	18917	18917	NA	12329	12329

Notes: See notes to Table 2.4. Robust standard errors in parentheses. ***p<0.01. **p<.05, *p<0.1

Table 2.11. Impact on School Enrollment for Daughters ages 5-14 (born prior to policy)

	DHS IV (2000)			DHS V (2004-2008)		
	Standard Reweighting	Modified Reweighting		Standard Reweighting	Modified Reweighting	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0
S_2=1	0.009	0.023*	-0.011	0.005	0.016	-0.007
	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)	(0.015)
Age	0.000	0.001	-0.001	0.004***	0.003	0.004***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
Years of Schooling	0.015***	0.017***	0.012**	0.005	0.010*	0.001
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)
# of children in 1997	-0.005	-0.006	-0.004	-0.010**	-0.008	-0.009
	(0.005)	(0.006)	(0.004)	(0.004)	(0.005)	(0.006)
Age at first birth	0.002	0.001	0.002	-0.001	0.000	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Geographic controls	yes	yes	yes	yes	yes	yes
Controls for child's age	yes	yes	yes	yes	yes	yes
Constant	-0.023	-0.015	-0.041	0.980***	0.900***	0.992***
	(0.048)	(0.050)	(0.049)	(0.050)	(0.056)	(0.067)
R-squared	0.649	0.632	0.681	0.126	0.128	0.150
Observations	11008	11008	11005	6905	6905	6903
Observations with positive $\Delta P(x)$	NA	9176	9176	NA	6153	6153

Notes: See notes to Table 2.4. Robust standard errors in parentheses. ***p<0.01. **p<.05, *p<0.1

Table 2.12. Impact on Education of Daughters 15-22

	DHS V (2004-2008)		
	Standard Reweighting	Modified Reweighting	
	(1)	(2)	(3)
Treatment	Contaminated Treatment Group D=1 & D=0	Sterilized by FSP D=1	Sterilized outside FSP D=0
S_2=1	-0.140	-0.061	-0.155
	(0.129)	(0.186)	(0.131)
Age	0.001	-0.043	0.007
	(0.021)	(0.031)	(0.022)
Years of Schooling	0.119**	0.284***	0.047
	(0.054)	(0.078)	(0.055)
# of children in 1997	-0.268***	-0.160**	-0.290***
	(0.058)	(0.067)	(0.068)
Age at first birth	0.019	0.065**	0.011
	(0.023)	(0.033)	(0.026)
Geographic controls	yes	yes	yes
Constant	10.320***	13.246***	12.350***
	(1.452)	(1.634)	(0.947)
R-squared	0.476	0.425	0.526
Observations	4304	4304	4304
Observations with positive $\Delta P(x)$	NA	3490	3490

Notes: See notes to Table 2.6. Robust standard errors in parentheses. ***p<0.01. **p<.05, *p<0.1

REFERENCES

Allison, Paul D. 1982. "Discrete-Time Methods for the Analysis of Event Histories." *Sociological methodology*, 13(1), 61-98.

Bailey, Martha J. 2006. "More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply." *The Quarterly Journal of Economics*, 121(1), 289-320.

Boesten, Jelke. 2007. "Free Choice or Poverty Alleviation? Population Politics in Peru under Alberto Fujimori." *European Review of Latin American and Caribbean Studies*, 82, 3.

Coe, Anna-Britt. 2004. "From Anti-Natalist to Ultra-Conservative: Restricting Reproductive Choice in Peru." *Reproductive Health Matters*, 12(24), 56-69.

Cox, DR. 1972. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2), 187-220.

DiNardo, John; Nicole M. Fortin and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica*, 64(5), 1001-44.

Fujimori, Alberto, 1999. Statement of Peru, por Sr. Presidente de la Republica del Peru, Ing. Alberto Fujimori at the Twenty-first Special Session of the General Assembly 30 June - 2 July 1999, United Nations Headquarters, New York in the debate on the overall review and appraisal in the implementation of the Programme of Action of the International Conference on Population and Development, http://www.un.org/popin/unpopcom/32ndsess/gass/state/peru.pdf, Monday, September 19, 2011.

Miller, Grant. 2010. "Contraception as Development? New Evidence from Family Planning in Colombia*." *The Economic Journal*, 120(545), 709-36.

Moffitt, Robert. 2005. "Remarks on the Analysis of Causal Relationships in Population Research." *Demography*, 42(1), 91-108.

Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, *World Population Prospects: The 2008 Revision*, http://esa.un.org/unpp, Monday, April 25, 2011

Rodriguez, German. 2007. "Survival Models." *Lecture Notes, Princeton University*.

Schultz, T Paul. 2005. "Effects of Fertility Decline on Family Well-Being: Opportunities for Evaluating Population Programs." *New Haven, Yale University*.

Tamayo, Giulia. 1999. *Nada Personal: Reporte De Derechos Humanos Sobre La Aplicación De La Anticoncepción Quirúrgica En El Perú, 1996-1998*. Latin American Committee for the Defense of Women's Rights (CLADEM, Comité de América Latina y el Caribe para la Defensa de los Derechos de la Mujer).

Vasquez del Aguila, Ernesto. 2006. "Invisible Women: Forced Sterilization, Reproductive Rights, and Structural Inequalities in Peru of Fujimori and Toledo." *Estudos e Pesquisas em Psicologia*, 6(1), 109-24.

Chapter 3

Impact of a Youth-Targeted Reproductive Health Initiative on Teen Pregnancy in South Africa

Abstract:
In the early 2000s, the NGO loveLife, in partnership with the South African Department of Health, rolled out the National Adolescent Friendly Clinic Initiative (NAFCI) with the goal of preventing HIV and unwanted pregnancy through education and increased clinical access to reproductive health services. By 2010, 500 clinics were accredited as "youth friendly." Based on interviews with stakeholders and a series of controls, I argue that the roll-out led to a conditionally random increase in reproductive health knowledge and clinical access for adolescents. I use GPS data and historical residence information from secure National Income Dynamics Study data to geolink respondents' location during their early teen years to the accreditation date and location of NAFCI clinics. Preliminary results show that women who lived within 5 km of a NAFCI clinic when they were 12-17 years old are substantially and statistically significantly less likely to experience a birth before the age of 18.

Acknowledgements:
I am grateful to Cally Ardington, Martha Bailey, Nicola Branson, Diane Cooper, Robert Garlick, David Lam, Murray Leibbrandt, Jeff Smith and Pam Smock for helpful comments and guidance.

3.1 Introduction

There is an unresolved debate about whether family planning policies affect fertility. Bongaarts (1994) argues that the increased supply of contraception provided by family planning interventions drive down fertility. Pritchett (1994) counters that while increased contraceptive use is *coincident* with falling fertility, reduced demand for children drives fertility decline, and the causal impact of family planning interventions is small. As in any market where we only observe the amount consumed, disentangling the role of supply and demand drivers is a challenge. In the realm of fertility determining causal pathways is especially challenging given the inherent endogeneity of fertility decisions and the dearth of good instruments or feasible randomizations. In the case of adolescent fertility there are often frictions on both the demand and supply sides. Lack of sex education can break the link between desire to prevent pregnancy and contraceptive use. Social stigma can be a barrier between contraceptive supply and teen access.

In early post-apartheid South Africa, the supply of family planning was technically unconstrained. The Apartheid regime's goal of controlling the non-white population resulted in widespread, free availability of contraception, and the new democratic government elected in 1994 ushered in some of the most progressive reproductive health laws in the world (Cooper et al., 2004). The 1998 South African Demographic and Health Survey reported that 80 to 90 percent of 20 to 30 year olds had ever used a modern contraceptive (DHS 1998). However, among the 30 percent of all South African women under 20 who already had at least one child, 79% reported they did "not want" their last pregnancy.[1] This clear evidence of an unsatisfied desire among teens to delay pregnancy would be defined by family planning advocates as unmet

[1] Mothers under the age of 20 regarding their last pregnancy in the last 3 years: 20% wanted pregnancy then, 66% wanted later,13% did not want the pregnancy.

need—a supply problem. South African public health researchers and advocates believed that high rates of unintended teen pregnancy which coincided with soaring rates of HIV among young women pointed to a knowledge gap and social and institutional barriers to adolescent access to reproductive health services. Thus, evidence on both the supply and demand side pointed to a more optimal outcome, but informational and social frictions drove a wedge between unmet demand to control fertility and untapped supply of family planning. The National Adolescent Friendly Clinic Initiative (NAFCI) aimed to tackle this wedge.

NAFCI was an intensive clinic accreditation, education and community outreach intervention that rolled out to health clinics across South Africa starting in the early 2000s. Based on interviews with stakeholders, a series of controls, and evidence of a trend breaks in service provision, I argue that the timing and geographic variation of the rollout led to an exogenous increase in adolescent access to reproductive health knowledge and services. My analysis is based on geo-linking data from several sources to construct a measure of access to NAFCI clinics for South African adolescents and comparing fertility outcomes for women who lived near accredited clinics in adolescence compared to those who did not. I find that women who lived within 5 km of a NAFCI clinic when they were 12 to 17 years old are substantially and statistically significantly less likely to experience a birth before the age of 18.

This paper contributes to the public health literature on the effectiveness of programs to increase adolescent access to and usage of reproductive health services, particularly in Africa.[2] There is a growing emphasis in public health on youth friendly services, but to date the evidence

[2] Dick et al. (2006) survey the literature in a World Health Organization report.

on efficacy has been sparse.[3] The paper also ties into the extensive economics literature on the causes and consequences of adolescent childbearing in the US (ex. Geronimus and Korenman, 1992, Lang and Weinstein, 2013) and the expanding literature in developing economies (Herrera and Sahn 2013, Marteleto et al., 2008, Ranchhod et al., 2011).

Finally, this paper contributes to the literature on the impact of family planning on fertility outcomes by providing new causal estimates of a youth-targeted intervention on fertility timing. Recent studies in both the US and developing economies use innovative identification strategies to bolster Pritchett's claim that increasing access to family planning has negligible impacts on completed fertility (Bailey, 2006, Miller, 2010). But these papers point to other important and more nuanced impacts of access to contraception such as enabling women to optimally time fertility to maximize human capital investments. My results indicate that an intervention specifically directed at increasing adolescent access to reproductive health knowledge and services has the potential to substantially alter fertility timing among teens

3.2 Background and Description of the National Adolescent Friendly Clinic Initiative

3.2.A Fertility Timing and Contraceptive Access in Early Post-Apartheid South Africa

The 1998 South African Demographic and Health Survey (DHS) provides context for the early post-Apartheid contraceptive and fertility patterns that motivated the National Adolescent Friendly Clinic Initiative.[4] Thirty-five percent of 19 year-olds (and 25 percent of 18 year olds) reported ever being pregnant (DHS 1998: Final Report 2002).[5] While this rate of adolescent

---

[3] A new randomized control trial in Ghana was presented at PAA: Aninanya et al. (2014) http://paa2014.princeton.edu/uploads/141674

[4] Apartheid was a system of strictly enforced racial segregation in South Africa. Apartheid officially ended with multi-racial democratic elections in 1994. While laws no longer classify citizens by the color of their skin, the classifications of White, Coloured, Black African, and Indian are still used in everyday conversation and are designations in surveys including the South African Census.

5 Among 19 year-olds 30.2 report being a mother, and among 18 year olds, 19.8 report being a mother.

childbearing is low compared to other sub-Saharan countries[6], teen childbearing in South Africa is more likely to be non-marital, rather than the result of early marriage (Macleod and Tracey, 2010, United Nations Population Fund, 2003). Only 1.2 percent of South African 15-19 year olds were married in 1998. There was also evidence of “widespread” and “endemic” gender violence and coercive sex experienced by teenage girls in South Africa (Wood et al., 1998).[7] And rates of unintended pregnancy among South African teens were high--78 percent of women under the age of 20 reported that their last birth was not wanted or wanted later (DHS 1998).

The Apartheid regime’s plan to control the non-white population led to relatively high contraceptive prevalence in South Africa compared to other sub-Saharan countries (Cooper, et al., 2004). Contraceptives were widely available at no cost at public clinics, hospitals and through mobile service provision.[8] However, the high rate of unintended pregnancy among teens suggests that South African adolescents had a substantial unmet “need” for family planning among adolescents.

Based on birth histories from the 2012 wave of the National Income Dynamics Study (NIDS)[9], Figure 3.1 shows patterns of age-at-first-birth by cohort, and Figure 3.2 shows children-ever-born by age-at-first-birth across cohorts. Figure 3.1 shows that after a decrease in the early-teen birth rate from the 1960 to 1970 birth cohorts, the proportion of 18 and 19 year olds who had given birth remains nearly constant for the 1970, 1980 and 1990 cohorts. At the same time, Figure 3.2 shows that among women born between 1980 and 1990, who were

[6] Adolescent birth rate per 1,000 women aged 15 to 19, 2005-2010 was 54 in South Africa while the average for Sub-Saharan African was 117. The rate in Uganda was 159 and Zambia 151.(UN Population Fund, 2003).
[7] A qualitative study in an African township in peri-urban Cape Town in the mid-1990s found that over 60% of female respondents aged 14-18 reported having sex against their will, and 59% reported have been beaten by their male partners (Wood, et al., 1998). Note, however, that rates of physical abuse by teenage girls in DHS 1998 are significantly lower. According to Human Rights Watch, in 1995 South Africa had the highest recorded per capita rate of rape of for a country not at war.
[8] Long-acting injectables were, and remain, the most common contraceptive method used (DHS 1998, DHIS 2013).
[9] More details about NIDS are provided in Section 3.3A.

adolescents in this early post-apartheid era, a teen birth was much more likely to be followed by a substantial space before the next birth than for earlier cohorts. This implies that completed fertility is converging for women having early versus late first births in the post-apartheid period. The pattern of falling overall fertility in South Africa combined with little change in age at first birth is consistent with teen mothers only starting to use contraception *after* a first birth.

Why were sexually active teens who did not want to get pregnant not using contraceptives when they were widely available for free? Qualitative studies in various South African regions aimed to address this question (Abdool Karim et al., 1992, Ehlers, 2003, Jewkes et al., 2001, Mfono, 1998, Wood and Jewkes, 2006). The findings pointed to social barriers to adolescent access to family planning. First, teens seemed to lack accurate sexual and contraceptive knowledge. For example there were widespread fears stoked by religious leaders and even nurses that hormone-based contraceptive use by adolescents could cause permanent infertility.[10] Next, stigmatization of adolescent sex by health care providers often made clinics inhospitable. Teens reported scolding and even abusive behavior by staff and nurses at public clinics and hospitals when they sought contraceptives, and in some cases even refusal to provide contraceptives.

Concerns among health advocates and the Department of Health about these barriers to adolescent access to reproductive health services were also driven by the increasing prevalence of HIV among youth, particularly teenage girls. Department of Health surveys found that in 1998 and 1999, approximately 20% of pregnant 15-19 year olds were HIV positive (substantially

[10] (Jewkes, et al., 2001) report that teen "mothers often indicate that teenage pregnancy is infinitely preferable to the possibility of infertility caused by contraceptive use … This is widely perceived by women and family planning nurses to be a side-effect of progesterone based injectable contraceptives, particularly Depo-Provera." This notion was also espoused by "preachers at local African churches".

higher in some regions such as KwaZulu Natal) (Allen et al., 2000, Jewkes, et al., 2001). HIV prevalence among 15-24 year olds was estimated to be three times higher among young women than young men--15.% versus 4.8% (Pettifor et al., 2005).

3.2.B The National Adolescent Friendly Clinic Initiative

High rates of unintended teen pregnancy and escalating rates of HIV among young people were the driving force behind the establishment of the NGO loveLife in 1999. The National Adolescent Friendly Clinic Initiative (NAFCI) was a key element of loveLife's strategy that also included high profile media campaigns and sporting events promoting "more open and better informed communication about sex, HIV, sexuality and gender relations." loveLife launched NAFCI in consortium with several other non-governmental organizations and partnership with the South African Department of Health.[11]

NAFCI had a clinical component aimed at reducing physical and social barriers to accessing reproductive health services, and an education component focused on sex education and life skills. The clinical component was based on an "accreditation model" whereby clinics worked towards service standards through a quality improvement process and were rewarded tiered levels of accreditation based on external assessments. The intensive accreditation process, which typically lasted a year, involved training nurses as well as non-medical staff, equipping facilities to offer the services and pharmaceuticals youth need, youth-targeted educational materials, and publicizing the clinics' youth friendliness through signage and community outreach. NAFCI's education component involved building dedicated spaces *at* clinics for youth

[11] Other organizations involved in the consortium were Planned Parenthood, the Reproductive Health Research Unit (RHRU) at the University of Witwatersrand and the Health System Trust.

education and socialization called “chill rooms” and employing local youth to facilitate sex-education programs.[12] Figure 3.5 gives and example of “youth friendly” signage.

The NAFCI was piloted at 10 clinics in 2000. A major scale-up occurred in 2004 and 2005 resulting in 350 active NAFCI sites by the end of 2005. By 2010, almost 500 clinics across the country were accredited as “youth friendly,” or approximately ten percent of all public clinics. Figure 3.3 shows the rollout of accredited clinics by activation year and province. Each NAFCI clinic is assigned at least one full-time loveLife peer educator (groundBREAKER, aged 18 to 25). In 2013, 1200 groundBREAKERS were employed nationwide, assisted by 6000 to 8000 part-time youth volunteers.[13]

There is limited evidence on the effectiveness of interventions aimed at increasing adolescent utilization of health services in developing countries.[14] In 2006, the World Health Organization reviewed evidence from 16 published and unpublished studies and concluded that “the evidence is weak and the findings not conclusive” due in large part to lack of “detailed descriptions of initiatives” and insufficient outcome data (Dick, et al., 2006). The review cites one randomized control study in Nigeria that finds improvements in knowledge and treatment-seeking behavior for sexually transmitted diseases (STD) among high-school students from a school-level treatment involving “community participation, peer education, public lectures, health clubs in the schools, and training of STD treatment providers” (Okonofua et al., 2003). By estimating the impact of NAFCI on teen pregnancy, this paper is the first to my knowledge to examine the impact of youth friendly services on fertility outcomes.

---

[12] The clinical accreditation process is described in detail in Ashton et al. (2009), and Dickson-Tetteh et al. (2001). The ten NAFCI standards are listed in Appendix Table 3.1. The education component of NAFCI is also discussed in Ashton et al. (2009).

[13] http://www.lovelife.org.za/corporate/lovelife-programmes/youth-leadership-development/groundbreakers/

[14] In the US there have been randomized control trials of impact of sex education programs that find comprehensive sex-education reduces teen pregnancy compared to no sex-education, but that abstinence-only education has statistically indistinguishable impacts on teen pregnancy compared to comprehensive sex-education. (Kohler et al., 2008)

3.3 Data and Empirical Strategy to Measure the Impact of NAFCI on Early Teen Pregnancy

This paper's research design is facilitated by geo-linking data on the timing and location of the NAFCI rollout to birth histories in nationally representative survey data. My empirical specification uses proximity to a NAFCI clinic during adolescence as a plausibly exogenous measure of access to reproductive health services and education and estimates the impact on early teen fertility. Linked census and health provision data provide controls and evidence on the impact of the initiative. Data on satellite locations that only offered the education component of the program but were not linked to a clinic help to separate out the impacts of the clinical versus the education components and suggest that interventions linked to clinics are most effective.

3.3.A Data to Geo-link NAFCI Rollout to Adolescent Birth Histories

I geo-link several datasets to implement my research design: 1) loveLife Project Monitoring Database, 2) District Health Information System (DHIS) GPS and Service Provision by Facility files, 3) National Income Dynamics Study (NIDS) Secure Data and Secure Administrative Data, and 4) the 2001 South African Census.

The loveLife Project Monitoring Database provides names of each NAFCI clinic and month and year the accreditation process began. Based on interviews with loveLife and clinic staff, I estimate that the effective start date of "youth friendly" services is one year after accreditation began.[15] District Health Information System (DHIS) facility-level files provide GPS coordinates and monthly service provision data for every public health facility in South Africa from 2001 to 2012.[16] Figure 3.4 shows the location and start year of NAFCI clinics from

[15] The database also provides start dates and locations for loveLife "outlets" which provide the education component of the program, but are not linked to a health facility.
[16] The District Health Information System (DHIS) is a health management information system and data warehouse developed by the Health Information Systems Programme (HISP) used by the South African Department of Health (DOH) to collect and monitor routine health data. This data is the basis of the annual South African District Health

2000 to 2010 based on linking the loveLife database with the DHIS. I also use contraceptive distribution (including the two major injectables, pills, iuds, condoms) and reported sexually transmitted diseases (STIs) aggregated to an annual level to track changes in service provision by facility type—NAFCI or non-NAFCI.

The National Income Dynamics Study (NIDS) is a nationally representative longitudinal household survey of over 28,000 individuals in 7,300 households fielded every two years starting in 2008 (SALDRU, 2014). NIDS includes detailed birth histories for all women over the age of 14 at the time of interview. NIDS secure data includes GPS coordinates of residence at time of interview as well as residency history including city/suburb of birth, residence in 1994, 2006, 2009 and 2011.[17] Secure administrative data also include GPS of last school attended, which is important as youth often go to clinics near school rather than home both for convenience and confidentiality.

The key variables in my analysis are age at first birth and distance of residence (or school) to a NAFCI clinic during adolescence. I define a sample of approximately 2000 female NIDS respondents age 19 to 26 in 2010 (wave 2).[18] I create a birth history variable with three categories: first birth by 18, first birth after 18 or no births. I use retrospective residence questions to determine where each woman lived each year when she was aged 12 to 17. If she still resides in that location, I have the GPS of residence during adolescence; otherwise I use the GPS of the centroid of the "main-place" of residence to approximate her location in

---

Barometer that provides indicators of the health system at district level. Monthly facility-level data was obtained with authorization from the DOH and with assistance from HISP.

[17] Secure NIDS data can only be accessed at the DataFirst secure facility at the University of Cape Town upon approval by the NIDS management committee.

[18] This age range was chosen with an eye to examining second-stage employment outcomes—women who are old enough to plausibly be out of secondary school and in the labor force.

adolescence.[19] I calculate the distance in kilometers from residence (school) to the nearest NAFCI clinic for each year a woman was 12 to 17.[20] I create a binary variable for whether she lived (went to school) within five kilometers of a NAFCI clinic.[21] I also create a variable for distance to any public health facility at age 15.

Finally, I link each respondent to her reported sub- (or main-) place of residence at age fifteen. There are 3,109 main-places and 21,243 sub-places designated in the 2001 South African census.[22] I construct a set of variables describing demographic characteristics for each sub- and main-place in the 2001 South African census including rural/urban status, and percent of the population 20 and over with less than 12 years of schooling as a proxy for socio-economic status. These serve as pre-policy control variables—NAFCI was piloted as early as 1999, but the main rollout occurred in the mid-2000s.

3.3.B Empirical Strategy to use NAFCI Rollout as Plausibly Exogenous Increase Access to Reproductive Health Services

My empirical strategy exploits NAFCI's staged rollout across South Africa to identify the impact of changes in adolescent access to contraception and sex education on early teen pregnancy. The key identification assumption is that the timing and location of clinic accreditation is uncorrelated with other determinants of fertility timing. For example, I need to assume that clinic locations were not chosen where teen fertility, either in terms of levels or trends, was different than otherwise similar locations. Similarly, if accreditation was pursued where demand for contraception was highest, this would also threaten identification. I provide

[19] Retrospective location questions were asked at a suburb/town level, but the raw responses are currently coded only to the Main Place level.

[20] Distances are calculated using the user-written command geonear (Picard, 2010). "geonear finds the nearest neighbors using geodetic distances, i.e. the length of the shortest curve between two points along the surface of a mathematical model of the earth."

[21] I chose 5 km based on conversations with clinic staff and loveLife provincial managers.

[22] See Appendix Figure 3.1. provides a description of census geographic sub-categories from province to enumeration area

both empirical and institutional evidence to suggest that geographic distribution of NAFCI clinics is effectively random after controlling for observable characteristics.

First, to provide evidence that NAFCI had a impact on service provision at clinics, I use facility-level DHIS data to measure changes in reproductive health services provided at NAFCI relative to non-NAFCI clinics before and after accreditation using an event-study framework (Jacobson et al., 1993). A descriptive background provides a context for the event study analysis. Appendix Figure 3.2 shows the overall trends in service provision of male condoms, the two major injectables, Depo-Provera and NET-N, and reports of new sexually transmitted infections (STIs) from 2001 to 2011. The figure shows a strong upward trend in condom distribution and fall in new STI cases over the period. Injectable provision was relatively flat with evidence of some switching between the two types.

Figure 3.6 shows relative changes in service provision at NAFCI clinics after accreditation, accounting for the national trends shown in Appendix Figure 3.2 using non-NAFCI clinics as controls and calendar year-fixed effects. After accreditation there is a trend break in the number of condoms distributed (increase) and STIs reported (decrease) at NAFCI clinics compared to non-NAFCI clinics. The relatively flat pre-accreditation trend lends credibility to the assumption that the initiative did not target clinics that had different growth trajectories from the average clinic. Figure 3.6 provides evidence that NAFCI clinics increased condom distribution at a faster rate than non-NAFCI clinics and that the new STI rate fell more quickly at NAFCI clinics.

There is also evidence in the lower two panels of Figure 3.6 for trend breaks in the two main injectable contraceptives, though the provision of NET-EN increases while the provision of

Depo-Provera decreases. This may be explained by NAFCI-trained clinic staff attempting to dispel the perception that Depo-Provera is "not for youth." Depo-Provera is much more likely than NET-EN to cause temporary amenorrhea (absence of a menstrual period)(Draper et al., 2006), which leads some to believe it causes infertility or other health problems (Wood and Jewkes, 2006)

Next I regress whether a respondent lived near a NAFCI clinic during adolescence on 2001 characteristics (urban/rural and percent with less than secondary degree) of the main-place where the respondent lived at age 15 and on 53 district council fixed effects. Many of the coefficients (Appendix Table 3.1) are statistically significant which is not surprising as the initiative was generally focused on areas of high need and low socio-economic status. However, only around 10 percent of all clinics were accredited as youth friendly by 2009 and I control for all of these variables when estimating the impact of NAFCI. As long as accreditation was random *conditional* on these controls, the identification assumption will still hold. Granted, this remains a strong assumption given the evidence provided so far. I continue to collect empirical evidence through more extensive controls and in the meantime I argue the case of conditionally random placement based on institutional knowledge gathered though interviews with stakeholders involved with the rollout process.

Based on interviews with people involved with the rollout of NAFCI, my understanding is that clinics were chosen in a relatively ad hoc way that varied by province and district. [23] According to the first director of loveLife, statistics on teen pregnancy and HIV were not used as selection criteria as those statistics did not exist with any geographic detail at the time.   I was

[23] I conducted extensive interviews with eight current and past employees and consultants of loveLife. I interviewed two provincial managers who were at loveLife during the initial implementation of the program. I visited NAFCI-accredited clinics in Gauteng, Eastern Cape and Western Cape where I met with nurses and local loveLife youth peer educators.

told on multiple occasions that since clinics were usually chosen by provincial or district level departments of health, and that varying personalities and agendas of provincial or district managers led to a “random” mix of clinics across the country (though they did not mean a formal random selection process was used). In some cases “struggling” clinics were targeted, and in others clinics that were perceived to be doing relatively well were rewarded by being chosen for the program. There were many more clinics that either wanted to be involved, or that district managers wanted to be include, than could be accommodated due to the intensity and expense of the program. NAFCI was targeted at high-need communities; however, there are an abundance of high-need communities across South Africa. I feel confident that many clinics that were otherwise similar to chosen clinics were not selected simply due to lack of funds, organization and time.

3.3.C Empirical Specifications

Equation (1) highlights my strategy of measuring the impact of adolescent–friendly clinic access in adolescence ($t-1$) on fertility outcomes measured in adulthood (t).

$$(1) \qquad Birth\ History_{i,t} = \alpha_0 + Access\ to\ NAFCI_{i,t-1} + \phi' X_{i,t-1} + \delta_j + \epsilon_i$$

Specifically, I estimate a linear probability model by ordinary least squares where the outcome is a binary indicator of having a first birth by age 18. I define NAFCI access with a binary indicator for living (or going to school) within 5 kilometers of a NAFCI clinic any time between age 12 and 17 as in Equation (2)

$$(2) \qquad 1[Birth\ by\ 18]_{i,t} = \alpha_0 + 1[NAFCI\ clinic < 5km]_{i,t-1} + \phi' X_{i,t-1} + \delta_j + \epsilon_i$$

where $X_{i,t-1}$ are demographic characteristics of place of residence at age 15 from the 2001 census, and δ_j are district fixed effects.[24]

I also estimate multinomial logit models where the outcome is fertility timing categorized as 1) no births, 2) first birth by 18 or 3) first birth after 18. This specification allows me to separate out impacts on the timing versus extensive margin.

3.4 Estimates of the Impact of NAFCI on Teen Fertility

Table 3.1 compares birth histories for the NIDS sample with and without access to NAFCI-accredited clinics. Of the roughly 2000 women in the sample, 250 lived near a NAFCI clinic in adolescence; and despite the fact that NAFCI clinics were placed in poor communities the prevalence of early teen births is lower among those who lived near a clinic. This result is shown to be marginally significant in the first column of Table 3.2, which gives estimates of equation (2) without controls or fixed effects. As seen in the second column, the estimated effect of living near a clinic is stronger when I control for geographic characteristics. This makes sense if clinics that became accredited were on average in poorer, higher teen-pregnancy locations. The estimated 7.7 percentage point drop in the likelihood of having a first birth by age 18 is significant at a 5 percent level. The third column of Table 3.2 shows that attending a school within 5 km of a NAFCI clinic also reduced the likelihood of a birth by 18 by a statistically significant 10.8 percentage points. Given the mean rates of teen childbearing in non-NAFCI areas shown in Table 3.1, a 7.7 percentage points represents a 44 percent drop in the rate of early teen pregnancy.

The multinomial logit coefficients shown in Table 3.3 show a statistically significant reduction in the odds of having a first births by age 18 compared to no births, but a small and

[24] There are 53 districts in South Africa. I choose district level fixed effects as districts can be thought of as local labor markets as described in (Dinkelman, 2011).

insignificant reduction on the odds of having a birth *after* 18 compared to no births.[25] These results imply that NAFCI's effect was to delay rather than reduce fertility.

Table 3.4 compares the results of living near a NAFCI clinic to living near other types of loveLife facilities that only provided an education intervention but were not linked to a clinic—Outlets, Franchises and Y-Centers. These results may allow me to parse out the separate impact of education and clinical access and suggest that the education component alone had a smaller and less significant impact. Living near an education-only facility, does not have a significant impact on likelihood of having a birth by age 18. These results suggest that the education intervention alone is not as effective, but cannot resolve if linking the two components is more important or if the clinical component alone is driving the results in Table 3.1.

3.5 Conclusion

My results for the impact of NAFCI in South Africa indicate that a youth-targeted reproductive health initiative has the potential to substantially and significantly reduce the likelihood of early teen pregnancy. More research is required to disentangle the education and clinical access components of the NAFCI intervention, but it is clear that increasing teen access to clinics is key to the success of the program. While not ruling out the importance of the demand side, my results imply that reducing barriers to contraceptive supply has a causal impact on fertility timing among adolescents in the South African context.

The policy concern surrounding teen childbearing stems largely from a belief that early age at first birth leads to negative consequences for later life outcomes. The findings of the current paper and its data and empirical strategy lay the groundwork to study these longer-term

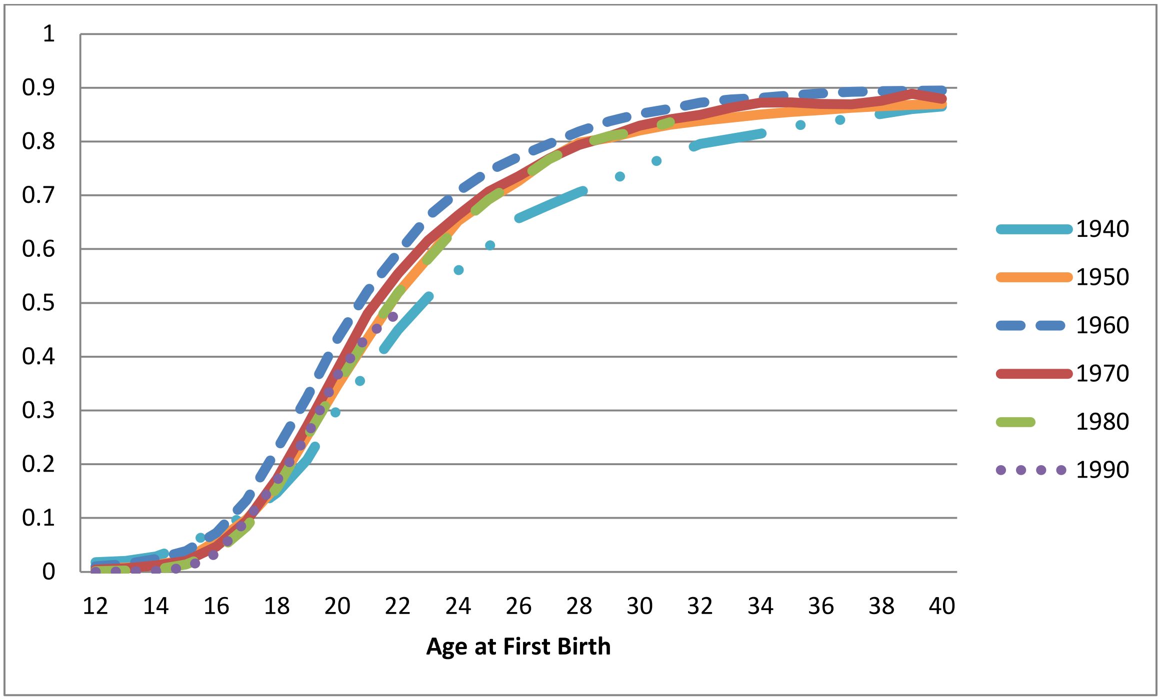
[25] A negative coefficient implies that living near a clinic reduces the odds of the given outcome relative to the reference category (no births) compared to someone not living near a clinic. A positive coefficient would imply increased odds.

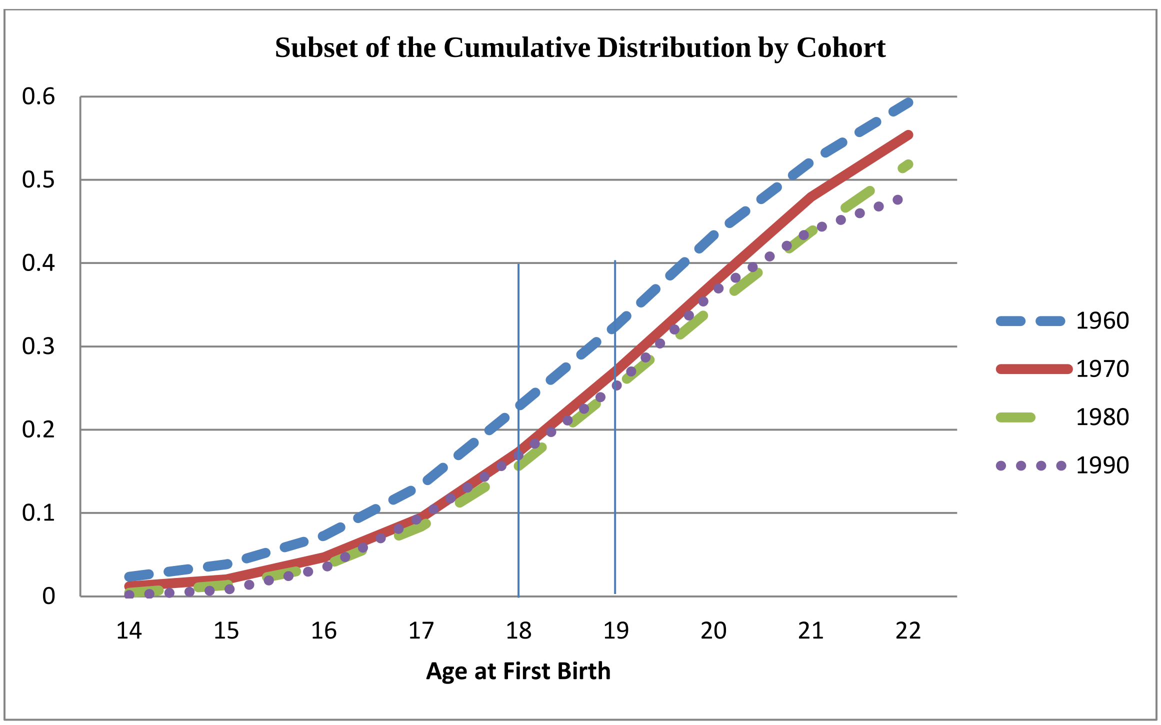
outcomes. Since the 1990s, there has been mounting evidence from the US that adverse socio-economic outcomes for teen mothers are not the causal impact of early fertility, but rather echoes of the correlation between circumstances in adolescence and later life outcomes.[26] However, there is little evidence to disentangle the causes and consequences of teen childbearing in low and middle-income countries. Recent work by Lang and Weinstein (2013) point out that the consensus debunking the negative consequences of teen motherhood is based on the analysis of teen births that occurred after Roe v. Wade when contraceptive access was widespread in the US. Analyzing an earlier period when family planning was not easily accessible for minors, they estimate negative causal impacts of teen motherhood.[27] Limited or uneven access to contraception characterizes many developing countries today. Herrera and Sahn (2013) studying Madagascar and Ranchhod, et al. (2011) studying an urban area in South Africa conclude that a negative impact of early childbearing remains after controlling for family background and economics circumstances in adolescence. These findings suggest an important role in developing economies for policies like NAFCI that improve young people's ability to plan fertility.

[26] In 1992 Geronimus and Korenman showed that, controlling for family background by comparing sister pairs with one teen and one non-teen mother, much of the cross-sectional difference in outcomes for teen versus non-teen mothers disappears and in terms of child outcomes even reverses.

[27] They hypothesize that once contraception and abortion are widely available, anyone for whom the cost would of early childbearing is high will avoid or terminate the pregnancy.

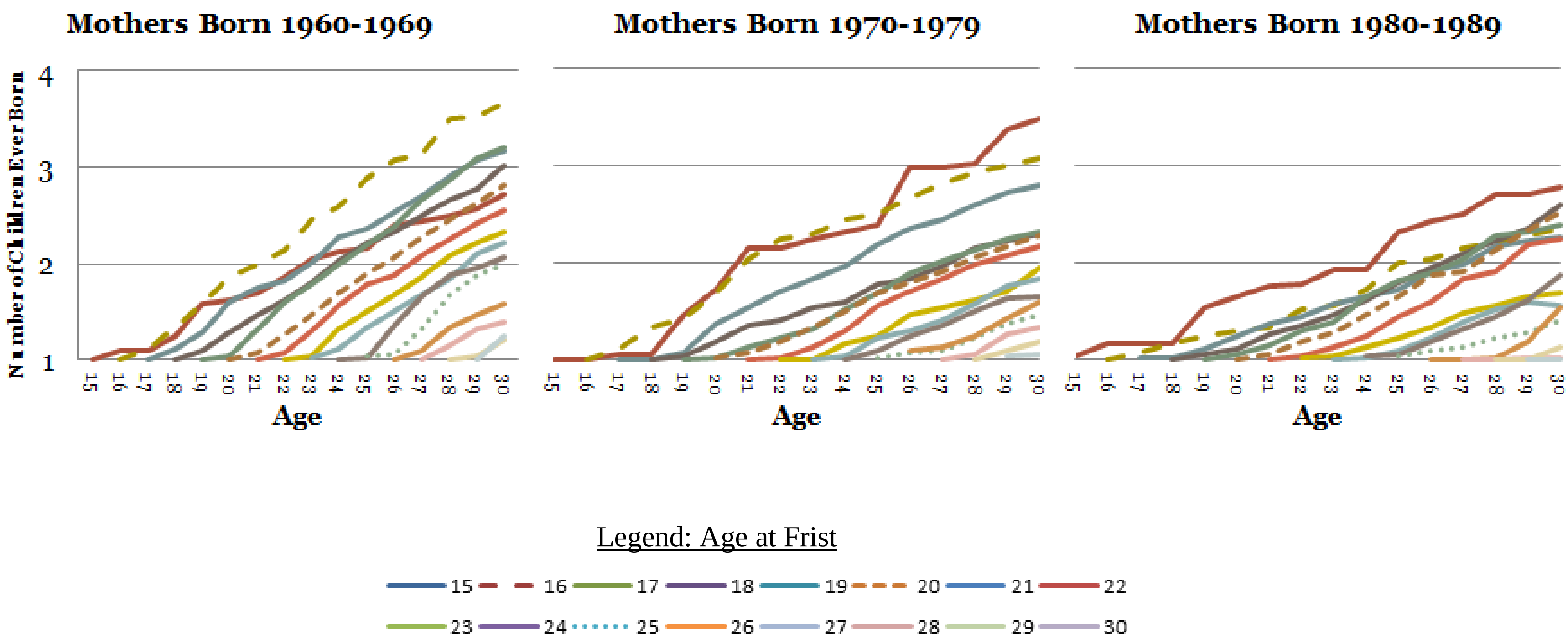
Figure 3.1. Cumulative Distribution of Age at First Birth by Cohort





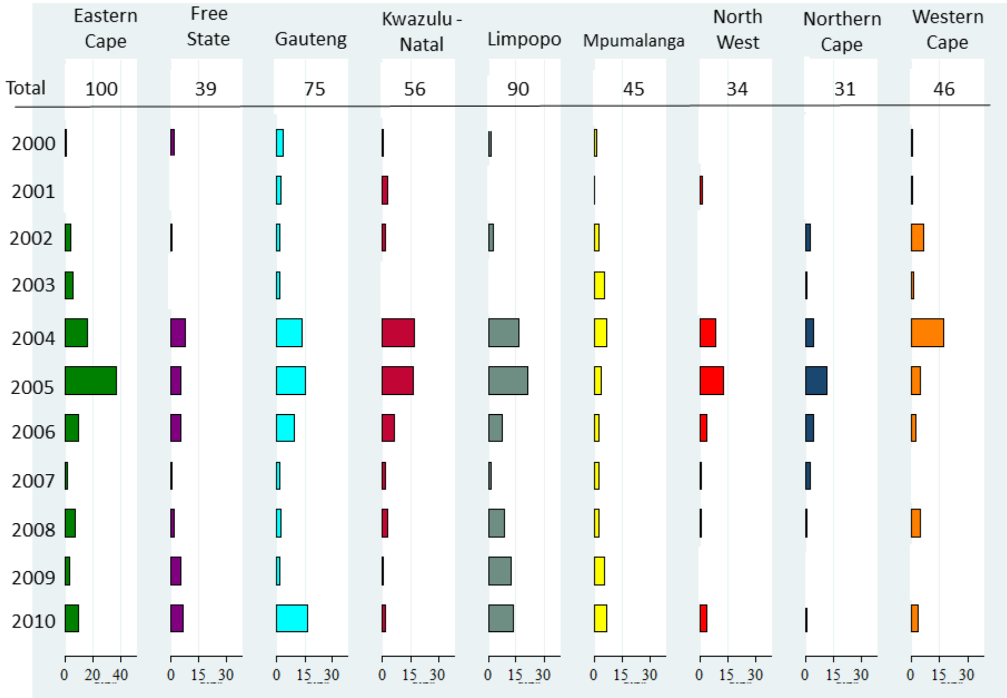
Notes: Author's calculations based on birth histories in the South African National Income Dynamics Study (NIDS) Wave 3, 2012..

Figure 3.2. Children Ever Born by Age at First Birth, across Cohorts



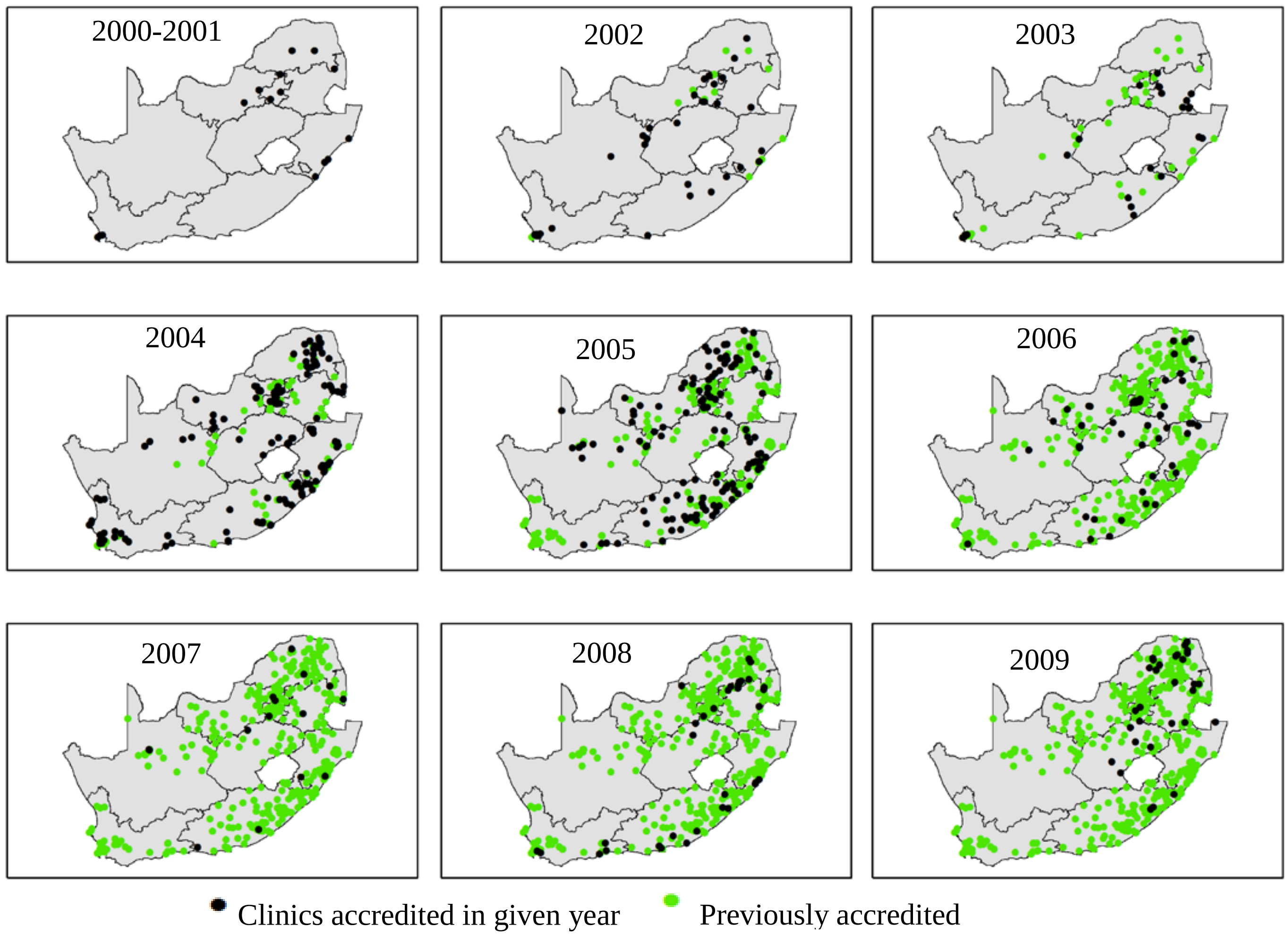
Notes: Author’s calculations based on birth histories in the South African National Income Dynamics Study (NIDS) Wave 3, 2012.

Figure 3.3. Rollout of National Adolescent Friendly Clinic Initiative by Year of Accreditation and Province



Notes: Source: loveLife project monitoring database

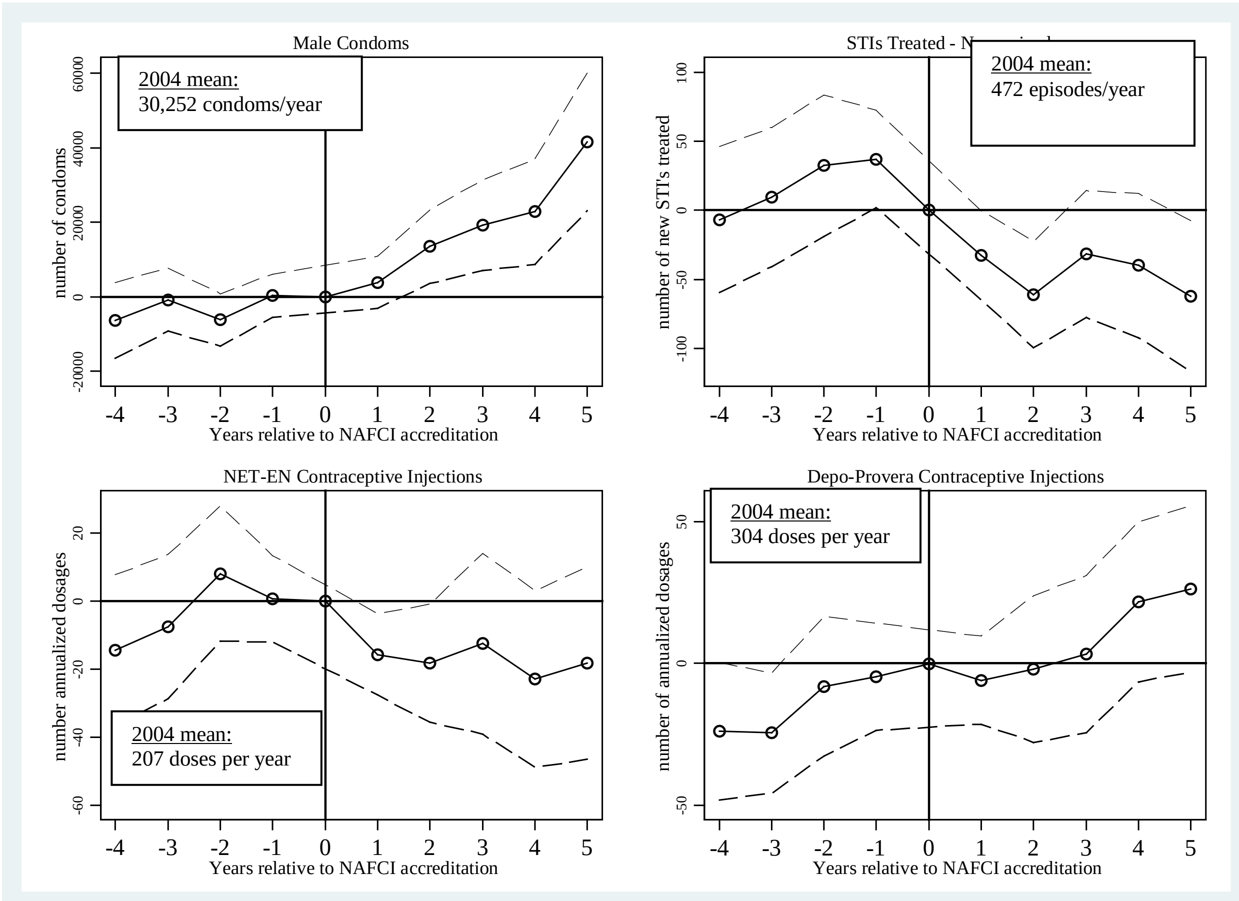
Figure 3.4. Geography and Timing of National Adolescent Friendly Clinic Initiative Rollout



Source: loveLife project monitoring database and District Health Information System (DHIS)

Figure 3.5. Youth Friendly Clinic Signage

Figure 3.6. Changes in Reproductive Health Service Provision Relative to Year of NAFCI Accreditation



Notes: Source: Service provision data from the South African District Health Information System. Information on timing of clinic accreditation form loveLife Project Monitoring Databases. Each plot shows the change in amount of a given service/contraceptive provided at the clinic relative to the level in year zero (one year after the accreditation process started). These estimates control for national trends in service provision by including controls for trends among non-NAFCI accredited clinics and calendar year fixed effects. Results, not shown, including province fixed-effects and year×province fixed-effects show very similar patterns. Depo-provera (depot medroxyprogesterone acetate) and NET-EN (norethisterone oenanthate) are long-acting injectable contraceptives. Depo-provera is given every three months and NET-EN every two months. In the figures above the number of injections are annualized by dividing the number to Depo injections by six and the number of NET-EN injections by four. The 2004 mean levels listed in each section of the graph are calculated among all public clinics. See Appendix Figure 3.2 for average trends from 2001 to 2011.

Table 3.1. Women age 19 to 26, by birth category and NAFCI access

	Ever lived within 5km of a YFS clinic when age 12-17?			
	Sample Size		Percentage in each birth category (using NIDS sampling weights)	
	no	yes	no	yes
No Births	717	124	41.41	52.18
Birth by 18	302	46	13.77	9.85
Birth after 18	753	80	44.82	37.97
Total	1772	250		

Notes: Source: Respondents from the South African National Income Dynamics Study Wave 2 geo-linked to NAFCI clinics using data from District Health Information System and loveLife Project Monitoring Databases.

Table 3.2. Estimated Impact of Access to NAFCI Clinics on Likelihood of Birth by Age 18

	Type of Access		
Dependent variable: birth by age 18	Lived Near NAFCI Clinic		School Near NAFCI Clinic
Had Access to Youth Friendly Clinics any time when aged 12-17	-0.039*	-0.077**	-0.108**
	(0.023)	(0.035)	(0.045)
Population group (African omitted)			
Coloured		0.047	0.174*
		(0.047)	(0.090)
White and other		-0.063*	-0.090*
		(0.034)	(0.051)
Age		0.118	0.203
		(0.092)	(0.142)
Age^2		-0.003	-0.005
		(0.002)	(0.003)
Kilometers to any public clinic		-0.002	
		(0.002)	
Main Place controls		yes	yes
District Council fixed effects		yes	yes
Constant	0.138***	-1.201	-2.057
	(0.011)	(1.044)	(1.589)
Observations	2,025	2,009	931
R-squared	0.002	0.063	0.099

Robust standard errors in parentheses

** p<0.05, * p<0.1

Notes: Source: NIDS wave 2. Sample includes all women age 19 to 26 including women who have not given birth. Being "near" means that the clinic was within 5km of residence or school respectively. Main place is a South Africa Census geographic designation one level below a Municipality. Main Place controls are i) percent of the population 20 and older with less than 12 years of schooling, and ii) rural/urban status. Main place controls are based on the 2001 census and are linked to the respondent's location in 2001 -- prior to roll out of NAFCI. Sample size smaller for school-proximity analysis becuase not all respondents provided information on school attended.

Table 3.3. Multinomial Logit Estimates of Impact of NAFCI on Fertility Timing

Multinomial Logit Estimates of Impact of NAFCI on Fertility Timin

	Odds of Birth Before or After 18 Relative to No Birth	
	Birth by 18	Birth after 18
Lived within 5km of NAFCI clinic	-0.8389**	-0.1145
any time when aged 12-17	(0.3864)	(0.2956)
Age	1.6525	2.3052
	(0.9832)	(0.8389)
Age^2	-0.0327	-0.0423
	(0.0220)	(0.0187)
Population group (African omitted)		
Coloured	0.0165	-0.1940
	(0.4863)	(0.4253)
White and other	-2.7717	-1.1523
	(0.9530)	(0.6160)
Main Place controls	yes	
District Council fixed effects	yes	
Constant	-21.9013	
	(10.9628)	

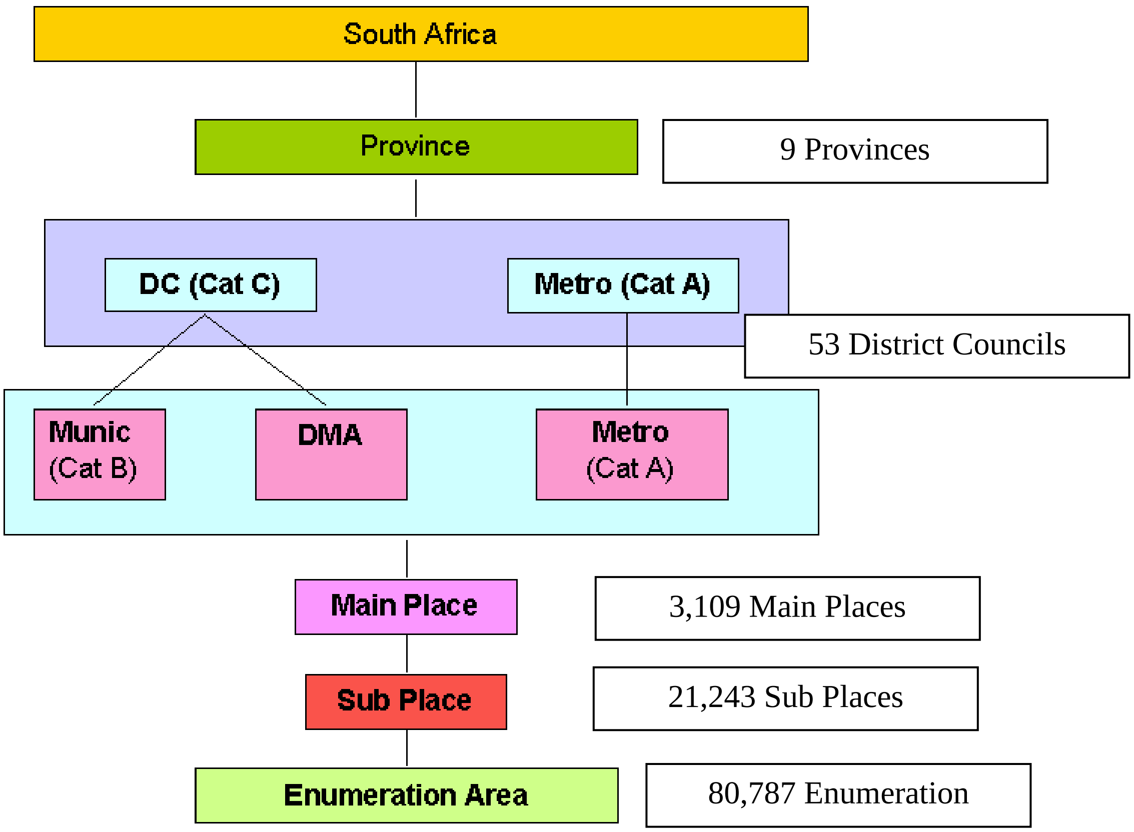
Notes: Source: NIDS wave 2. Sample includes all women age 19 to 26 including women who have not given birth. Main place is a South Africa Census geographic designation one level below a Municipality. Main Place controls are i) percent of the population 20 and older with less than 12 years of schooling, and ii) rural/urban status. Main place controls are based on the 2001 census and are linked to the respondent's location in 2001 -- prior to roll out of NAFCI.

Table 3.4. Comparing Impact of Access to NAFCI Clinics to Impact of Other loveLife facilities on Likelihood of Birth by Age 18

	Type of Access			
Dependent variable: birth by age 18	NAFCI Clinic	loveLife Outlet	loveLife Franchise	loveLife Y-center
Lived near facility any time when aged 12-17:				
NAFCI Clinic	-0.077**			
	(0.035)			
loveLife Outlet		0.008		
		(0.088)		
loveLife Franchise			-0.041	
			(0.044)	
loveLife Y-center				-0.010
				(0.044)
Population group (African omitted)				
Coloured	0.047	0.089	0.057	0.054
	(0.047)	(0.076)	(0.048)	(0.047)
White and other	-0.063*	-0.013	-0.061*	-0.062*
	(0.034)	(0.041)	(0.033)	(0.034)
Age	0.118	-0.013	0.106	0.115
	(0.092)	(0.324)	(0.092)	(0.092)
Age^2	-0.003	0.001	-0.002	-0.003
	(0.002)	(0.008)	(0.002)	(0.002)
km_to_PC	-0.002	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Main Place controls	yes	yes	yes	yes
District Council fixed effects	yes	yes	yes	yes
Constant	-1.201	0.160	-1.089	-1.189
	(1.044)	(3.324)	(1.060)	(1.055)
Observations	2,009	1,303	2,009	2,009
R-squared	0.063	0.067	0.059	0.059

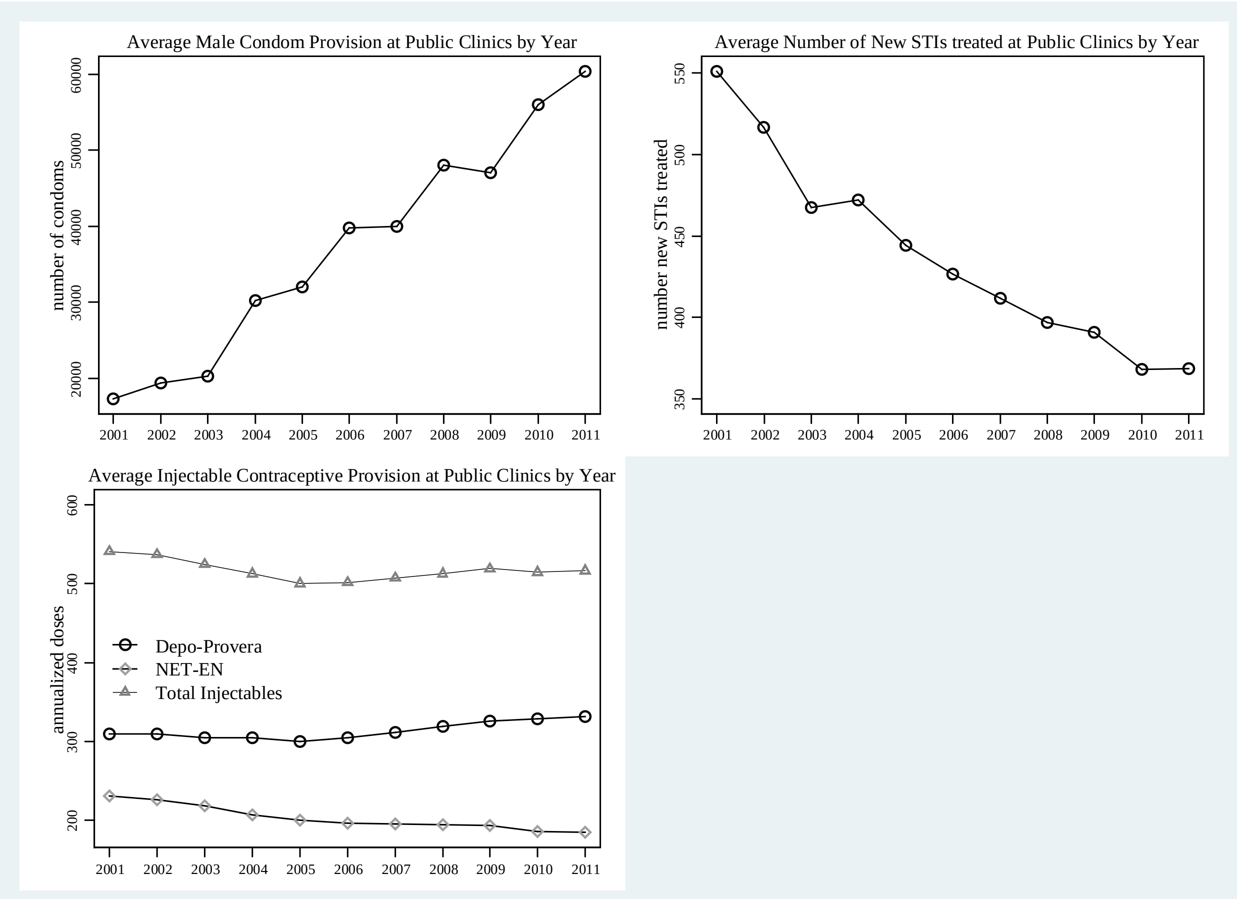
Notes: Source: NIDS wave 2. Sample includes all women age 19 to 26 including women who have not given birth. Being "near" means that the facility was within 5km of residence. Main place is a South Africa Census geographic designation one level below a Municipality. Main Place controls are i) percent of the population 20 and older with less than 12 years of schooling, and ii) rural/urban status. Main place controls are based on the 2001 census and are linked to the respondent's location in 2001--prior to roll out of NAFCI. Robust standard errors in parentheses. ***p<0.01. **p<.05, *p<0.1

Appendix Figure 3.1. South African Census 2001 geographical area hierarchy structure



Source: (Statistics South Africa 2001) http://www.statssa.gov.za/census01/html/Geography_Metadata.htm

Appendix Figure 3.2. Trends in Reproductive Health Service Provision among All Public Clincs by Year



Notes: Source: South African District Health Information System facility-level data.

Appendix Table 3.1. National Adolescent Friendly Clinic Initiative Standards

1. Management systems are in place to support the effective provision of adolescent-friendly services.
2. The clinic has policies and processes that support the rights of adolescents.
3. Clinic services appropriate to the needs of adolescents are available and accessible.
4. The clinic has a physical environment conducive to the provision of adolescent friendly health services
5. The clinic has the drugs, supplies and equipment necessary to provide the essential service package for adolescent-friendly health care.
6. Information, education and counseling consistent with the Essential Service Package are provided
7. Systems are in place to train staff to provide effective adolescent-friendly services.
8. Adolescents receive an accurate psychosocial and physical assessment.
9. Adolescents receive individualized care on based on standard service delivery guidelines.
10. The clinic provides continuity of care for adolescents.

Source: (Ashton, et al., 2009)

Appendix Table 3.1. Correlates of NAFCI Clinic Placement

Dependent Variable: NAFCI Clinic near Residence at age 15-17				
	coefficient	Std. Err.	t	P>|t|
2001 Main Place Controls (Rural omitted)				
Urban	0.2153	(0.0282)	7.64	0.000
Semi/urban	0.1226	(0.0252)	4.86	0.000
Percentage less then 12 yrs	0.0049	(0.0007)	7.04	0.000
District Council fixed effects				
2	0.2380	(0.1165)	2.04	0.041
3	-0.1955	(0.1794)	-1.09	0.276
4	0.2337	(0.1396)	1.67	0.094
5	-0.2074	(0.2635)	-0.79	0.431
6	-0.1822	(0.1905)	-0.96	0.339
7	-0.1171	(0.2760)	-0.42	0.671
8	-0.1023	(0.1678)	-0.61	0.542
9	0.5446	(0.1321)	4.12	0.000
10	0.0994	(0.1343)	0.74	0.46
12	0.0313	(0.1150)	0.27	0.786
13	-0.0655	(0.1185)	-0.55	0.581
14	0.1814	(0.1308)	1.39	0.165
15	-0.1248	(0.1049)	-1.19	0.234
16	-0.1817	(0.1716)	-1.06	0.29
17	-0.0870	(0.1125)	-0.77	0.439
18	-0.1724	(0.1153)	-1.49	0.135
19	-0.1354	(0.1169)	-1.16	0.247
20	-0.1246	(0.1260)	-0.99	0.323
21	-0.0980	(0.1144)	-0.86	0.392
22	0.0423	(0.1109)	0.38	0.703
23	-0.1066	(0.1160)	-0.92	0.358
24	-0.1081	(0.1269)	-0.85	0.394
25	0.2526	(0.1276)	1.98	0.048
26	-0.0928	(0.1162)	-0.8	0.425
27	-0.0998	(0.1207)	-0.83	0.408
28	-0.1039	(0.1187)	-0.88	0.381
29	-0.0699	(0.1252)	-0.56	0.577
30	0.0982	(0.1175)	0.84	0.403
31	-0.0010	(0.1077)	-0.01	0.993
32	0.0523	(0.1092)	0.48	0.632
33	-0.0862	(0.1092)	-0.79	0.43
34	0.0734	(0.1161)	0.63	0.527
35	-0.0433	(0.1063)	-0.41	0.684
36	0.0010	(0.1198)	0.01	0.993
37	0.0382	(0.1120)	0.34	0.733
38	-0.0646	(0.1138)	-0.57	0.571
39	0.0262	(0.1218)	0.22	0.830
40	-0.1023	(0.1455)	-0.7	0.482
42	0.1941	(0.1063)	1.83	0.068
43	-0.1426	(0.1331)	-1.07	0.284
44	-0.1209	(0.1306)	-0.93	0.355
76	0.1374	(0.1059)	1.3	0.195
81	0.1783	(0.1843)	0.97	0.333
82	0.1815	(0.1588)	1.14	0.253
83	-0.0508	(0.1072)	-0.47	0.635
84	0.1010	(0.1138)	0.89	0.375
88	0.3582	(0.1227)	2.92	0.004
171	0.5153	(0.1033)	4.99	0.000
275	0.2626	(0.1086)	2.42	0.016
572	-0.0515	(0.1014)	-0.51	0.611
773	0.3873	(0.1045)	3.71	0.000
774	0.1038	(0.1031)	1.01	0.314
constant	-0.3255146	0.115812	-2.81	0.005
Observations	2030			
R-squared	0.3156			

REFERENCES

Abdool Karim, Q.; E. Preston-Whyte and S. S. Abdool Karim. 1992. "Teenagers Seeking Condoms at Family Planning Services. Part I. A User's Perspective." *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde*, 82(5), 356-59.

Allen, David M; Nolhemba P Simelela and Lindiwe Makubalo. 2000. "Epidemiology of Hiv/Aids in South Africa." *Southern African Journal of HIV Medicine*, 1(1).

Ashton, Joanne; Kim Dickson and Melanie Pleaner. 2009. "Evolution of the National Adolescent-Friendly Clinic Initiative in South Africa," World Health Organization,

Bailey, Martha J. 2006. "More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply." *The Quarterly Journal of Economics*, 121(1), 289-320.

Bongaarts, John. 1994. "The Impact of Population Policies: Comment." *Population and Development Review*, 20(3), 616-20.

Cooper, Diane; Chelsea Morroni; Phyllis Orner; Jennifer Moodley; Jane Harries; Lee Cullingworth and Margaret Hoffman. 2004. "Ten Years of Democracy in South Africa: Documenting Transformation in Reproductive Health Policy and Status." *Reproductive Health Matters*, 12(24), 70-85.

Department of Health Republic of South Africa and Macro International Inc. . 2002. "South African Demographic and Health Survey 1998: Final Report,"

Dick, B.; J. Ferguson; V. Chandra-Mouli; L. Brabin; S. Chatterjee and D. A. Ross. 2006. "Review of the Evidence for Interventions to Increase Young People's Use of Health Services in Developing Countries," *Technical Report Series-World Health Organization. 938:151.*

Dickson-Tetteh, Kim; Audrey Pettifor and Winnie Moleko. 2001. "Working with Public Sector Clinics to Provide Adolescent-Friendly Services in South Africa." *Reproductive Health Matters*, 9(17), 160-69.

Dinkelman, Taryn. 2011. "The Effects of Rural Electrification on Employment: New Evidence from South Africa." *American Economic Review*, 101(7), 3078-108.

Draper, Beverly H; Chelsea Morroni; M Hoffman; J Smit; M Beksinska; J Hapgood and Lize Van der Merwe. 2006. "Depot Medroxyprogesterone Versus Norethisterone Oenanthate for Long-Acting Progestogenic Contraception." *Cochrane Database Syst Rev*, 3.

Ehlers, V. J. 2003. "Adolescent Mothers' Utilization of Contraceptive Services in South Africa." *International Nursing Review*, 50(4), 229-41.

Geronimus, Arline T. and Sanders Korenman. 1992. "The Socioeconomic Consequences of Teen Childbearing Reconsidered." *The Quarterly Journal of Economics*, 107(4), 1187-214.

Herrera, Catalina, and David E. Sahn. 2013 "The Impact of Early Childbearing on Schooling and Cognitive Skills Among Young Women in Madagascar." *Cornell University Food and Nutrition Policy Program Working Paper* 247.

Jacobson, Louis S.; Robert J. LaLonde and Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *The American Economic Review*, 83(4), 685-709.

Jewkes, Rachel; Caesar Vundule; Fidelia Maforah and Esmé Jordaan. 2001. "Relationship Dynamics and Teenage Pregnancy in South Africa." *Social Science & Medicine*, 52(5), 733-44.

Kohler, Pamela K.; Lisa E. Manhart and William E. Lafferty. 2008. "Abstinence-Only and Comprehensive Sex Education and the Initiation of Sexual Activity and Teen Pregnancy." *Journal of Adolescent Health*, 42(4), 344-51.

Lang, Kevin and Russell Weinstein. 2013. "The Consequences of Teenage Childbearing before Roe V Wade." *National Bureau of Economic Research Working Paper Series*, No. 19627.

Macleod, Catriona Ida and Tiffany Tracey. 2010. "A Decade Later : Follow-up Review of South African Research on the Consequences of and Contributory Factors in Teen-Aged Pregnancy." *South African Journal of Psychology*, 40(1), 18-31.

Marteleto, Letícia; David Lam and Vimal Ranchhod. 2008. "Sexual Behavior, Pregnancy, and Schooling among Young People in Urban South Africa." *Studies in Family Planning*, 39(4), 351-68.

Mfono, Zanele. 1998. "Adolescent Contraceptive Needs in Urban South Africa: A Case Study." *International Family Planning Perspectives*, 24(4), 180-83.

Miller, Grant. 2010. "Contraception as Development? New Evidence from Family Planning in Colombia*." *The Economic Journal*, 120(545), 709-36.

Okonofua, Friday E.; Paul Coplan; Susan Collins; Frank Oronsaye; Dapo Ogunsakin; James T. Ogonor; Joan A. Kaufman and Kris Heggenhougen. 2003. "Impact of an Intervention to Improve Treatment-Seeking Behavior and Prevent Sexually Transmitted Diseases among Nigerian Youths." *International Journal of Infectious Diseases*, 7(1), 61-73.

Pettifor, Audrey E; Helen V Rees; Immo Kleinschmidt; Annie E Steffenson; Catherine MacPhail; Lindiwe Hlongwa-Madikizela; Kerry Vermaak and Nancy S Padian. 2005. "Young People's Sexual Health in South Africa: Hiv Prevalence and Sexual Behaviors from a Nationally Representative Household Survey." *AIDS*, 19(14), 1525-34.

Picard, Robert. 2010. "Geonear: Stata Module to Find Nearest Neighbors Using Geodetic Distances,"

Pritchett, Lant H. 1994. "Desired Fertility and the Impact of Population Policies." *Population and Development Review*, 20(1), 1-55.

Ranchhod, V; D Lam; M Leibbrandt and L Marteleto. 2011. "Estimating the Effect of Adolescent Fertility on Educational Attainment in Cape Town Using a Propensity Score Weighted Regression." *A Southern Africa Labour and Development Research Unit Working Paper Number 59. Cape Town*, SALDRU, University of Cape Town.

Statistics South Africa. 2001. "Metadata: Geography Hierarchy and Attributes," *Census 2001*.

Southern Africa Labour and Development Research Unit. 2014 National Income Dynamics Study (2012, Wave 3 [dataset]. Version 1.2, 2010-2011, Wave 2 [dataset]. Version 2.2., 2008, Wave 1 [dataset]. Version 5.2) Cape Town: Southern Africa Labour and Development Research Unit [producer], 2014, 2009. Cape Town: DataFirst [distributor], 2014.

United Nations Population Fund. 2003. "Motherhood in Childhood: Facing the Challenge of Adolescent Pregnancy. State of World Population 2013.," New York: United Nations:

Wood, Kate and Rachel Jewkes. 2006. "Blood Blockages and Scolding Nurses: Barriers to Adolescent Contraceptive Use in South Africa." *Reproductive Health Matters*, 14(27), 109-18.

Wood, Katharine; Fidelia Maforah and Rachel Jewkes. 1998. "“He Forced Me to Love Him”: Putting Violence on Adolescent Sexual Health Agendas." *Social Science & Medicine*, 47(2), 233-42.