WORKING PAPER

Treatment Effects Using Inverse Probability Weighting and Contaminated Treatment Data

An Application to the Evaluation of a Government Female Sterilization Campaign in Peru

Tanya Byker and Italo A. Gutierrez

RAND Labor & Population

WR-1118-1
September 2016
This paper series made possible by the NIA funded RAND Center for the Study of Aging (P30AG012815) and the RAND Labor and Population Unit.

RAND CORPORATION

For more information on this publication, visit www.rand.org/pubs/working_papers/WR1118x1.html

Published by the RAND Corporation, Santa Monica, Calif.

TREATMENT EFFECTS USING INVERSE PROBABILITY WEIGHTING AND CONTAMINATED TREATMENT DATA: AN APPLICATION TO THE EVALUATION OF A GOVERNMENT FEMALE STERILIZATION CAMPAIGN IN PERU

Tanya Byker
Middlebury College
tbyker@middlebury.edu

Italo A. Gutierrez
RAND Corporation
italo@rand.org

September 2, 2016

Abstract

We evaluate the impact of a female sterilization campaign implemented by the Peruvian government in 1996 and 1997 that we estimate impacted nearly 70,000 women. We use an inverse probability weighting (IPW) estimator that accounts for contamination in the available data. The contamination arises because while we observe sterilization status, we do not know if a given sterilization occurred as part of the campaign or whether it was chosen without influence from the campaign. The distinction is important because women targeted by the campaign and women who opted for sterilization outside of the campaign likely differ in many aspects, and we suspect the impact of sterilization is different for each group. We show that it is not necessary to fully observe whether a sterilized woman underwent the procedure because of the campaign to estimate unbiased average treatment effect of the government campaign. It is sufficient to estimate--based on auxiliary data--the conditional probability that if a sterilization is observed, it occurred because of the campaign. Using the proposed IPW estimator, we find that women sterilized because of the campaign had on average fewer 0.95 children. We also find substantial and statistically significant improvements in the height for age--a measure of health--of girls whose mothers were sterilized because of the campaign, and small but positive and statistically significant effects on years of schooling for boys.

Keywords: female sterilization; fertility; family planning; contaminated data models; inverse probability weighting; causal effects; observational data.

JEL codes: C21, J13

ACKNOWLEDGEMENTS

We are grateful to Martha Bailey, Robert Garlick, Andrew Goodman-Bacon, David Lam, and Jeffrey Smith for helpful comments and guidance. We also thank Matthew Cefalu for his help with the Monte Carlo analysis. This research was supported by an NICHD training grant to the Population Studies Center at the University of Michigan (T32 HD007339), by a William and Flora Hewlett Foundation/Institute of International Education Dissertation Fellowship in Population, Reproductive Health and Economic Development, and by funding from the Center for Causal Inference at the RAND Corporation.

1. Introduction

In 1996 and 1997 the Peruvian Government implemented a female surgical sterilization campaign as a key element of a national poverty reduction strategy (Coe, 2004). There is little public information about the campaign beyond total numbers of sterilization procedures recorded by the department of health and reports by human rights activists of failure to obtain proper consent, coercion and abuse. There has been no formal evaluation of the impact of the campaign on fertility or families' wellbeing. In this paper, we use publicly available survey data and statistical methods to uncover the profile of women who were targeted by the sterilization campaign and to estimate the impact of the campaign on completed fertility, female labor supply, and child health and education outcomes.

The main challenge in evaluating the sterilization campaign is the lack of public records about women who were targeted by the campaign and how the campaign was implemented. The best publically available data come from the Peruvian Demographic and Health Surveys (DHS). These nationally representative surveys ask women 15 to 49 years old whether they were sterilized and the year of sterilization. As shown in Figure 1, there is a dramatic spike in female sterilizations in 1996 and 1997 and an equally dramatic fall by 1998, after the campaign was allegedly terminated. However, there is no information in the DHS about whether a woman was approached by public health workers at "health festivals" or at home visits as part of the government campaign and sterilized because of it. This is problematic since we would like to distinguish between women who were sterilized *because* of the campaign and women who were not targeted by the campaign and opted for sterilization *without influence* from the campaign. The distinction is important because sterilization is likely to have a different impact for each group. For example, we suspect women who opted for sterilization without influence from the campaign were on average more educated, more informed about modern contraceptive methods, and more likely to live in urban areas. Conversely, women who were sterilized because they were targeted by the campaign were more likely to be poor, less educated, less informed about modern contraceptive methods and more likely to live in rural areas. If there are heterogeneous effects of sterilization across these groups, simply using the population of all women sterilized in 1996 and 1997 to evaluate the effects of the campaign will result in biased estimates.

[Figure 1 here]

In this paper we propose an inverse probability weighting (IPW) strategy that allows us to estimate the average effect of the sterilization campaign even though the details of who was targeted remain undisclosed. In other words, we recover the average effect of sterilization among women who were sterilized because of the government campaign, despite the fact that whether a sterilized woman belongs to this group is not directly observed in the data. Our strategy requires that two probabilities can be estimated: i) the overall probability of sterilization during the campaign years (1996-1997) and ii) the conditional probability that if a woman was sterilized in those years, it happened because of the campaign. We estimate that women who were sterilized because of the campaign had on average 0.95 fewer children (by year 2004). We also find substantial and statistically significant improvements in the height for age (a measure of health) of girls whose mothers were sterilized because of the campaign, and small but positive and statistically significant effects on years of schooling for boys under the age of fifteen. These results contribute to the economic demography literature on the impact of family planning policies on fertility (surveyed in Miller and Singer Babiarz (2016)) and the related literature on the impact of fertility decline on family wellbeing (surveyed in Schultz (2007)). The nature of the family planning policy in Peru was aggressive and shrouded in controversy; however this type of policy is not an isolated event. Recent reports of quota-based sterilization campaigns in Uzbekistan and Czechoslovakia (Holt 2005 and 2012) make analyzing the impact of the Peruvian campaign on both fertility and household wellbeing important and relevant.

In addition to providing the first formal statistical evaluation of the sterilization campaign in Peru, our paper also contributes to the literature on estimation of treatment effects with contaminated data. In contaminated data models, some observations are contaminated, while others are clean (Chen, Hong and Nekipelov, 2011).[1] For example, Hotz, Mullin and Sanders (1997) adapt the method of Horowitz and Manski (1995) for bounding the distribution of a contaminated random variable to estimate the causal effect of teenage childbearing on the women's education and labor market outcomes. The key problem is to estimate the mean counterfactual outcomes for women who gave birth as a teenager. Hotz, Mullin and Sanders (1997) propose using information from teenagers who experienced a miscarriage to estimate that counterfactual. However, teenagers who experienced miscarriage form a contaminated control group. Some of them experience a miscarriage randomly and would have given birth otherwise. Other teenagers who experience miscarriage would have otherwise decided to abort. And some miscarriages may occur non-randomly to teenagers who smoke, use drugs or have other adverse health behavior, which also affect their schooling and labor market outcomes. Only the first subgroup provides a clean comparison. If the researchers could observe the nature of the miscarriage (random or not) and what the pregnancy outcome would have been in the absence of miscarriage (birth or abortion), they could use only the outcomes from teenagers who experience miscarriages randomly and would have given birth otherwise to estimate the counterfactual--

[1] In contrast, in measurement error models all observations can be measured but with errors of different magnitudes Chen, Hong and Nekipelov (2011).

information that is typically unobservable. Hotz, Mullin and Sanders (1997) show that placing a bound (obtained from epidemiological studies) on the proportion of teen miscarriages that occur randomly and would otherwise have resulted in births allows them to form bounds on the mean counterfactual, and therefore on the mean casual effect. Our empirical problem is similar, but we have a contaminated *treatment* group rather than a contaminated control group. Among women who were sterilized from 1996 to 1997, some were sterilized because of the government campaign (our group of interest), while others opted for sterilization without influence from the campaign. Using our methodology, we are able to obtain point identification of the effects of sterilization for the each group separately.

Our approach is also related to work by Botosaru and Gutierrez (2014), who develop a difference-in-difference estimator when treatment status is observed in only one period. Their methodology uses repeated cross-sections and relies on the ability to predict treatment status in one period based on knowledge of treatment status in the other period and auxiliary information available in both cross-sections. Similar to our approach, the Botosaru and Gutierrez (2014) method provides point estimates of the average treatment on the treated. Our empirical problem is different, however, in that we aim to estimate impacts for a subgroup of interest whose identity is not observable. Rather than a pre-post comparison as in Botosaru and Gutierrez, we rely on cross-sectional data with retrospective information and IPW methods. To the best of our knowledge, this paper is the first that combines IPW methods with contaminated data models.

The method we develop in this paper can be applied to other situations where researchers are interested in evaluating the effect of an intervention for a particular subgroup (e.g. individuals that were more compliant with the intervention, individuals who receive a more intense intervention, or individuals who received a better quality intervention) but identity of the subgroup of interest is not observed in the data. If the probability of belonging to the subgroup of interest can be estimated from observable auxiliary information, then it is possible to apply the strategy we present in this paper to estimate the average effect of the intervention for that subgroup.

The rest of the paper is organized as follows. Section 2 provides a brief background on the sterilization campaign in Peru. Section 3 describes the empirical problem of evaluating the effect of the sterilization campaign when only sterilization status is known and there is no direct observation on whether it occurred because of the campaign or outside of the campaign. Section 4 proposes an IPW estimator that addresses this problem. The estimator requires estimating (from auxiliary information) the conditional probability that if a sterilization is observed, it occurred because of the campaign. Section 5 shows how this auxiliary information can be obtained in our data and the required conditions. Section 6 presents the impact evaluation results of the sterilization campaign, including effects on number of children ever born and on other child health and education outcomes. Section 7 further discusses the plausibility of interpreting our finding as casual and the potential threats to the validity of our estimates. Section 8 presents the conclusion of the study. We also include an Appendix with a Monte Carlo simulation study that indicates that our estimator and empirical approach are able to recover unbiased causal impact estimates, given the assumptions of the model.

2. The Government Sterilization Campaign in 1996-1997

In 1983 the Ministry of Health began to offer public family planning services. However, until the early 1990s, the government gave little support to population issues and its public family planning program was poorly organized and relatively ineffective (Coe, 2004). In part, this was the result of the economic crisis and political turmoil, marked by a severe civil conflict, and by limited service delivery capacity as well as opposition to birth control by religious institutions (Gribble, Sharma and Menotti, 2007).

During Alberto Fujimori first presidential term (1990-1995), the government made an initial attempt to reduce total fertility and the number of unwanted pregnancies.[2] During his second term (1995-2000), a combination of factors allowed Fujimori to implement a more aggressive campaign to tackle high fertility rates among poor Peruvian families. A stabilized economic and political situation, an increase in international donor support for population-related activities and Peru's participation in 1994 of the International Conference on Population and Development (ICPD) Program of Action, permitted the Fujimori government to increase financial and political support for reproductive health initiatives. In 1995, the Ministry of Health instituted a policy to provide free family planning products and services to all who wanted them and the Congress authorized female sterilizations to be a standard family planning method. Before this policy change women could obtain only sterilization if they had a health risk, four or more children, or were above a certain age, and required spousal permission (Coe, 2004, Gribble, Sharma and Menotti, 2007).

In 1996, there was allegedly an abrupt change in the government female reproductive health policy that was not publically announced. The policy shift emerged as a response to the slow decline in poverty rates despite market-oriented economic reforms and sustained economic growth. The Fujimori government made contraceptive services the core component of its mass poverty reduction policy (Coe, 2004). The new policy centered on increasing the use of modern

[2] Data from the Demographic and Health Survey (DHS II, 1991-1992) showed that there was an "unmet need" for family planning. About 35 percent of women reported that their latest birth was not wanted; and this percentage increased to 65 percent among women with three or more children

contraceptives, especially female sterilization, among poor and marginalized women. As a main strategy, "health festivals" promoting sterilizations were held in rural areas (Tamayo, 1999) and mobile sterilizations teams were dispatched to perform procedures in rural and isolated areas (Coe, 2004, Mosher, 2002). There is evidence that health workers were given incentives to fill sterilization quotas (Boesten, 2007, Coe, 2004, Leon, 1999, Mosher, 2002), which led them to withhold other contraceptive methods in order to promote sterilizations (Coe, 2004, del Aguila, 2006).

By 1998, controversy erupted regarding documented cases of deception, economic incentives, and even coercion to meet those quotas. Due to increasing pressure from civil society and to the refusal of international donors to fund any further activity related to sterilizations, the campaign was dismantled although its details remained covert. In 1999 new guidelines for delivering family planning services was approved by the Ministry of Health. The new guidelines were more in accordance with the ICPD Program of Action, allowing women to exercise reproductive choice (Coe, 2004).[3]

3. The empirical problem

Let $S = \{0,1\}$ denote a discrete event where $S = 1$ if a woman was sterilized in 1996-1997 (the years of the sterilization campaign). This event S is observed in our data. Let $C = \{0,1\}$ denote a discrete event where $C = 1$ if a woman was sterilized *because* of the government campaign. For example, it includes instances when women were persuaded at government-run

[3] However, the post-Fujimori government of Alejandro Toledo also limited reproductive choice by opposing to birth control methods in general on religious grounds (Boesten, 2007, Coe, 2004, del Aguila, 2006).

health fairs or by mobile teams during home visits in remote areas. We can divide the overall probability of sterilization into the mutually exclusive events shown in equation (1):

$$\mathrm{P}(\mathrm{S}=1)=\mathrm{P}(\mathrm{S}=1,\mathrm{C}=1)+\mathrm{P}(\mathrm{S}=1,\mathrm{C}=0) \quad (1)$$

As mentioned in the previous section, some sterilization procedures performed as part of the campaign were conducted without proper consent, involved deception or were coerced. Thus, we could further divide the probability of sterilization because of the campaign as sterilizations that were voluntary and informed ($\mathrm{V}=1$) and sterilizations that were coerced or misguided ($\mathrm{V}=0$). In other words, $\mathrm{P}(\mathrm{S}=1,\mathrm{C}=1)=\mathrm{P}(\mathrm{S}=1,\mathrm{C}=1,\mathrm{V}=1)+\mathrm{P}(\mathrm{S}=1,\mathrm{C}=1,\mathrm{V}=0)$. However, we do not have enough information to infer whether proper consent for sterilizations was obtained or whether sterilizations were misguided or coerced. Thus, in the rest of the paper we abstract from event V and focus on distinguishing whether a sterilization happened as result of the government campaign ($\mathrm{C}=1$) or without influence from the campaign ($\mathrm{C}=0$). Making this distinction is still important because, as discussed earlier, the government campaign was allegedly directed at poor women, with little or no formal education (Coe, 2004), who had higher fertility rates and less knowledge of alternative contraceptive methods according to DHS II (1991-1992). We believe sterilizations likely had a different impact on this group of targeted women than on more educated or more affluent women who had knowledge about and access to a broader menu of contraceptive choices.

To help fix ideas, let X denote observed characteristics of a woman, such as education and place of residency (e.g. rural, urban). We parametrize the probability of being sterilized *because* of the campaign versus *outside* of the campaign as a function $\Gamma(\cdot)$ of linear indexes in X, as shown in equations (2) and (3).

$$\mathrm{P}(\mathrm{S}=1,\mathrm{C}=0|\mathbf{X})=\Gamma(\boldsymbol{\beta}_0\mathbf{X}) \quad (2)$$

$$P(S = 1, C = 1|\mathbf{X}) = \Gamma((\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1)\mathbf{X}) \tag{3}$$

Note that if $\boldsymbol{\beta}_1 \neq 0$, women sterilized because of the campaign will have a different distribution of characteristics than women sterilized outside the campaign. Denote by Y_1 the potential outcome if a woman is sterilized in 1996 or 1997 ($S = 1$) and denote by Y_0 the potential outcome if a woman is not sterilized ($S = 0$) in those years. In this paper, these potential outcomes include completed fertility, employment and measures of health and educational attainment for children born prior to 1996. It is plausible that the effects of sterilization on these outcomes are heterogeneous depending on the characteristics of the woman. We can characterize the potential outcomes as a function $Y(\cdot)$ of linear indexes in sterilization status S and other characteristics X, as shown in equation (4) and (5).

$$Y_0(\mathbf{X}, S) = Y(\boldsymbol{\eta}_1\mathbf{X}) \tag{4}$$

$$Y_1(\mathbf{X}, S) = Y\left(\boldsymbol{\eta}_1\mathbf{X} + \eta_2 S + \boldsymbol{\eta}_3(\mathbf{X} \times \mathbf{S})\right) \tag{5}$$

We can define three average causal effects of sterilization that are of interest:

$$\psi_1 = E[Y_1 - Y_0|S = 1, C = 1] \tag{6}$$

$$\psi_2 = E[Y_1 - Y_0|S = 1, C = 0] \tag{7}$$

$$\psi_3 = E[Y_1 - Y_0|S = 1] \tag{8}$$

The first effect, ψ_1, is the average causal effect of sterilization for women that were sterilized as a result of the government campaign. This is our object of interest. The second effect, ψ_2, is the average causal effect of sterilization for women that were sterilized without influence from the government campaign. And the third effect, ψ_3, is the average causal effect of sterilization among all sterilized women. It is helpful to notice that ψ_3 is a weighted average of ψ_1 and ψ_2, as shown in equation (9):

$$\psi_3 = \psi_1 \times P(C = 1|S = 1) + \psi_2 \times P(C = 0|S = 1) \tag{9}$$

The empirical problem for estimating our object of interest, ψ_1, arises from the simultaneous occurrence of three conditions. First, women who were sterilized because of the campaign were allegedly very different (e.g. poorer and less educated) than women who were sterilized without being approached by the campaign (i.e. $\boldsymbol{\beta}_1 \neq \mathbf{0}$ in equation (3)); second, the effect of sterilization is likely heterogeneous depending on the characteristics of the woman (i.e. $\boldsymbol{\eta}_3 \neq 0$ in equation (5)); and third, we observe S but not C, and thus we can only directly estimate ψ_3 from the data. The first two conditions imply that $\psi_1 \neq \psi_2 \neq \psi_3$. Thus, the readily available estimate, ψ_3, can be regarded as a contaminated (and biased) estimate of ψ_1 since it also includes the average causal effect for individuals that were not sterilized by the campaign (i.e. we have a contaminated treatment group).

4. A modified IPW estimator to address a contaminated treatment group

We propose an IPW estimator augmented with auxiliary data to tackle the problem of a contaminated treatment group. Even if treatment status (C) is uknown we can still obtain unbiased estimates of the treatment effect of interest (ψ_1) if we have auxiliary information that allows us to estimate the conditional probability that if a woman was sterilized, it was because of the government campaign ($P(C = 1|S = 1, \mathbf{X})$). To see how this conditional probability can be used in the estimation of ψ_1, we separately analyze the expected value of the two potential outcomes $E[Y_1|S = 1, C = 1]$ and of $E[Y_0|S = 1, C = 1]$. Start by defining $E[Y_1|C = 1, S = 1]$ as in equation (10). We then multiply and divide the integrand by the conditional probability density function $f(y_1, \mathbf{X}|S = 1)$ and then apply Bayes rule to the numerator and denominator of the new integrand to arrive at equation (11).

$$E[Y_1|S = 1, C = 1] = \int\int y_1 f(y_1, \mathbf{x}|S = 1, C = 1) d\mathbf{x} dy_1 \qquad (10)$$

$$E[Y_1|S=1,C=1]=\frac{P(S=1)}{P(S=1,C=1)}\int\int y_1\frac{P(S=1,C=1|y_1,\mathbf{x})}{P(S=1|y_1,\mathbf{x})}f(y_1,\mathbf{x}|S=1)d\mathbf{x}dy_1 \quad (11)$$

Next we invoke a strong ignorability condition, common in the application of IPW methods:

Condition 1 (Strong Ignorability): After conditioning on X, both the probability of sterilization because of the government campaign and the probability of sterilization without influence from the campaign are independent of the potential outcomes $\{Y_0, Y_1\}$:

$$P(S=1,C=1|y_1,y_0,\mathbf{X})=P(S=1,C=1|\mathbf{X}) \quad (12)$$

$$P(S=1,C=0|y_1,y_0,\mathbf{X})=P(S=1,C=0|\mathbf{X}) \quad (13)$$

$$P(S=1|y_1,y_0,\mathbf{X})=P(S=1|\mathbf{X}) \quad (14)$$

Condition 1 implies that regardless of potential outcomes, women who are observationally equivalent had the same probability of sterilization because of the government campaign (equation (12)), the same probability of being sterilized outside of the campaign (equation (13)), and therefore the same probability of being sterilized in general (equation (14)). Under Condition 1, equation (11) can be re-written as shown in equation (15). Notice that $\frac{P(S=1,C=1|\mathbf{x})}{P(S=1|\mathbf{x})}$ is equal to the conditional probability $P(C=1|S=1,\mathbf{X}=\mathbf{x})$, which suggests the finite sample estimator for $E[Y_1|S=1,C=1]$ provided in equation (16).

$$E[Y_1|S=1,C=1]=\frac{P(S=1)}{P(S=1,C=1)}\int\int y_1\left(\frac{P(S=1,C=1|\mathbf{x})}{P(S=1|\mathbf{x})}\right)f(y_1,x|S=1)d\mathbf{x}dy_1 \quad (15)$$

$$E[Y_1|S=1,C=1]=\sum_{i=1}^{N}y_i\left(\frac{s_i\phi_i\lambda_i}{\sum_{i=1}^{N}s_i\phi_i\lambda_i}\right) \quad (16)$$

The estimator is a weighted average of observed outcomes, Y, for women who are sterilized where is ϕ_i is the sampling weight and $\lambda_i=P(C=1|S=1,\mathbf{X}=\mathbf{x}_i)$ which is the probability that, conditional on being sterilized, the woman is sterilized because of the government campaign.[4,5]

We derive an estimator for the term $\mathrm{E}[\mathrm{Y}_0|\mathrm{S}=1,\mathrm{C}=1]$ in a similar way. We start with its definition in equation (17). Then, we multiply and divide the integrand by the conditional probability density function $\mathrm{f}(\mathrm{y}_0,\mathrm{x}|\mathrm{S}=0)$, apply Bayes rules and invoke Condition 1, to obtain equation (18).

$$\mathrm{E}[\mathrm{Y}_0|\mathrm{S}=1,\mathrm{C}=1]=\int\int \mathrm{y}_0\mathrm{f}(\mathrm{y}_0,\mathbf{x}|\mathrm{S}=1,\mathrm{C}=1)\mathrm{d}\mathbf{x}\mathrm{dy}_0 \tag{17}$$

$$\mathrm{E}[\mathrm{Y}_0|\mathrm{S}=1,\mathrm{C}=1]=\frac{\mathrm{P(S=0)}}{\mathrm{P(S=1,C=1)}}\int\int \mathrm{y}_0\frac{\mathrm{P}(\mathrm{S=1,C=1}|\mathbf{x})}{\mathrm{P}(\mathrm{S}=0|\mathbf{x})}\mathrm{f}(\mathrm{y}_0,\mathrm{x}|\mathrm{S}=0)\mathrm{d}\mathbf{x}\mathrm{dy}_0 \tag{18}$$

After replacing $\mathrm{P}(\mathrm{S}=0)$ with $1-\mathrm{P}(\mathrm{S}=1)$, multiplying and dividing the first term on the right-hand side by $\mathrm{P}(\mathrm{S}=1)$, and performing a similar operation for the integrand, we obtain the result in equation (19). This equation suggests that $\mathrm{E}[\mathrm{Y}_0|\mathrm{S}=1,\mathrm{C}=1]$ can be estimated with the finite sample estimator provided in equation (20), where $\theta_\mathrm{i}=\left(\frac{\mathrm{P}(\mathrm{S}=1|\mathbf{X}=\mathbf{x}_\mathbf{i})}{1-\mathrm{P}(\mathrm{S}=1|\mathbf{X}=\mathbf{x}_\mathbf{i})}\right)$ which resembles the standard IPW weight and λ_i and ϕ_i are defined as before.[6]

$$\mathrm{E}[\mathrm{Y}_0|\mathrm{S}=1,\mathrm{C}=1]=\left(\frac{1-\mathrm{P(S=1)}}{\mathrm{P(S=1)}}\right)\left(\frac{\mathrm{P(S=1)}}{\mathrm{P(S=1,C=1)}}\right)\int\int \mathrm{y}_0\left(\frac{\mathrm{P}(\mathrm{S=1,C=1}|\mathbf{x})}{\mathrm{P}(\mathrm{S}=1|\mathbf{x})}\right)\left(\frac{\mathrm{P}(\mathrm{S}=1|\mathbf{x})}{1-\mathrm{P}(\mathrm{S}=1|\mathbf{x})}\right)\mathrm{f}(\mathrm{y}_0,\mathbf{x}|\mathrm{S}=0)\mathrm{d}\mathbf{x}\mathrm{dy}_0 \tag{19}$$

---

[4] The finite sample estimator of $E[Y_1|S=1,C=1]$ is given by $\frac{P(S=1)}{P(S=1,C=1)}\sum_{i=1}^{N}y_i\left(\frac{s_i\phi_i\lambda_i}{\sum_{i=1}^{N}s_i\phi_i}\right)$. However, the population value $\frac{P(S=1)}{P(S=1,C=1)}$ can be approximated in finite samples with $\frac{\sum_{i=1}^{N}s_i\phi_i}{\sum_{i=1}^{N}s_i\phi_i\lambda_i}$, giving the expression in Equation (16).

[5] The sampling weight ϕ_i in our application is provided in the data based the on sampling design used by the DHS and it is the inverse of the probability of that each observation was sampled.

[6] The finite sample estimator of $E[Y_0|S=1,C=1]$ is given by $\left(\frac{1-P(S=1)}{P(S=1)}\right)\left(\frac{P(S=1)}{P(S=1,C=1)}\right)\sum_{i=1}^{N}y_i\left(\frac{(1-s_i)\phi_i\theta_i\lambda_i}{\sum_{i=1}^{N}(1-s_i)\phi_i}\right)$. However, the population value $\left(\frac{1-P(S=1)}{P(S=1)}\right)\left(\frac{P(S=1)}{P(S=1,C=1}\right)$ can be approximated in finite samples with $\frac{\sum_{i=1}^{N}(1-s_i)\phi_i}{\sum_{i=1}^{N}(1-s_i)\phi_i\theta_i\lambda_i}$, giving the expression in equation (20).

$$E[Y_0|S = 1, C = 1] = \sum_{i=1}^{N} y_i \left(\frac{(1-s_i)\phi_i\theta_i\lambda_i}{\sum_{i=1}^{N}(1-s_i)\phi_i\theta_i\lambda_i} \right) \tag{20}$$

Thus, the expected value $E[Y_0|S = 1, C = 1]$ is a weighted average of the observed outcome, Y, among women who are not sterilized. The weights are composed of two components. The first component, θ_i, gives higher weights to women who are observationally more similar to women that are sterilized, either due to the government campaign or outside of the campaign. The second factor, λ_i, gives higher weights to those women who, if they are sterilized it is more likely to be due to the government campaign.

Condition 2 summarizes the common support or overlap condition required for our IPW estimator. It requires that $P(S = 1|\mathbf{X} = \mathbf{x}_i) < 1$, so that θ_i is not undefined (i.e. the denominator is not zero). In other words, it requires no perfect predictability of sterilization given observed characteristics.[7]

Condition 2 (overlap): There is no perfect predictability of sterilization given X:

$$P(S = 1|\mathbf{X} = \mathbf{x}_i) < 1 \qquad \text{for all women} \tag{21}$$

By joining the results of equations (16) and (20) for estimating the potential outcomes $E[Y_1|S = 1, C = 1]$ and $E[Y_0|S = 1, C = 1]$, we can estimate our object of interest, ψ_1, as in equation (22). Notice how this estimator differs from the standard IPW estimator, which can be directly applied to estimate ψ_3 (i.e. the average causal effect of sterilization among all sterilized women), as shown in equation (23). The estimation of ψ_1 incorporates an additional term beyond the standard IPW estimator, λ_i, which measures the conditional probability that if a woman is sterilized, it is because of the government campaign.

---

[7] Because we are estimating average causal effects for the treated (i.e. for women that are sterilized) and not average causal effects for the population we do not need to restrict $P(S = 1 \mid \boldsymbol{X} = \boldsymbol{x}_i) > 0$ for all individuals.

$$\widehat{\psi}_1 = \sum_{i=1}^{N} y_i \left(\frac{s_i \phi_i \lambda_i}{\sum_{i=1}^{N} s_i \phi_i \lambda_i}\right) - \sum_{i=1}^{N} y_i \left(\frac{(1-s_i)\phi_i \theta_i \lambda_i}{\sum_{i=1}^{N} (1-s_i)\phi_i \theta_i \lambda_i}\right) \quad (22)$$

$$\widehat{\psi}_3 = \sum_{i=1}^{N} y_i \left(\frac{s_i \phi_i}{\sum_{i=1}^{N} s_i \phi_i}\right) - \sum_{i=1}^{N} y_i \left(\frac{(1-s_i)\phi_i \theta_i}{\sum_{i=1}^{N} (1-s_i)\phi_i \theta_i}\right) \quad (23)$$

Up to this point, we have focused on estimating the impact of the government campaign. We can also estimate ψ_2, i.e. the impact of sterilization for those women who are sterilized without influence the government campaign. The corresponding IPW estimator is given in equation (24). In this case the standard weights are augmented by the probability that if a woman is sterilized it was not due to the government campaign, or $(1 - \lambda_i)$. Note that for this estimator to be defined, we need to strengthen Condition 2 by requiring that $P(C = 1|S = 1, \mathbf{X} = \mathbf{x}_i) < 1$ for all women.

$$\widehat{\psi}_2 = \sum_{i=1}^{N} y_i \left(\frac{s_i \phi_i (1-\lambda_i)}{\sum_{i=1}^{N} s_i \phi_i (1-\lambda_i)}\right) - \sum_{i=1}^{N} y_i \left(\frac{(1-s_i)\phi_i \theta_i (1-\lambda_i)}{\sum_{i=1}^{N} (1-s_i)\phi_i \theta_i (1-\lambda_i)}\right) \quad (24)$$

5. Estimation of key weighting parameters $\boldsymbol{\lambda_i}$ and $\boldsymbol{\theta_i}$ using Peruvian DHS data

We investigate the effect of the sterilization campaign using the fourth and fifth waves of the Peruvian Demographic and Health Surveys (hereafter DHS IV and DHS V). Both DHS IV and DHS V are nationally representative cross sectional surveys conducted after the termination of the campaign allowing us to look at medium and long-term 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 surveys for our purposes is that they collect information on whether women are sterilized and the date when the sterilization occurred. The surveys also contain retrospective information on marital and birth histories, place of residence and basic demographic information like age and educational attainment.

As discussed in the previous section, a key step in our empirical approach is to use auxiliary information to estimate λ_i, the conditional probability that if a woman was sterilized in 1996-1997, it was because of the government campaign. In our case, we use retrospective information about women who were sterilized prior to the start of the campaign to predict the probability that, in the absence of a government campaign, a woman would have been sterilized in 1996 or 1997. To illustrate our strategy, we introduce time subscripts in our notation, such that S_1 denotes sterilization in the years prior to the start of the campaign and S_2 denotes sterilization during the campaign years (1996 and 1997). Thus, $P(S_2 = 1|S_1 = 0, \mathbf{X} = \mathbf{x})$ is the probability that a woman is sterilized during the campaign given that she was not sterilized before. Following equation (1), this probability can be decomposed as the sum of the probability of being sterilized because of the campaign and of the probability of being sterilized outside of the campaign:

$$P(S_2 = 1|S_1 = 0, \mathbf{X} = \mathbf{x}) = P(S_2 = 1, C = 1|S_1 = 0, \mathbf{X} = \mathbf{x}) + P(S_2 = 1, C = 0|S_1 = 0, \mathbf{X} = \mathbf{x}) \quad (25)$$

Condition 3 below allows us to exploit the information we have about the probability of sterilization in years prior to the campaign. A similar assumption is made by Botosaru and Gutierrez (2014) to estimate the probability of treatment for the period where treatment status is missing.

Condition 3 (stationarity in the probability of sterilizations that occur without influence from a government campaign): The probability that a woman with observed characteristics X is sterilized in years 1996 and 1997 without influence from the government campaign is the same as the probability of sterilization in the pre-campaign years for a woman with similar characteristics:

$$P(S_2 = 1, C = 0|S_1 = 0, \mathbf{X} = \mathbf{x}) = P(S_1 = 1|\mathbf{X} = \mathbf{x}) \quad (26)$$

Under Condition 3 we can re-write equation (25) as shown in equation (27) and the probability of sterilization due to the government campaign conditional on sterilization, λ_i, as shown in equation (28):

$$P(S_2 = 1, C = 1 | S_1 = 0, \mathbf{X} = \mathbf{x}) = P(S_2 = 1 | S_1 = 0, \mathbf{X} = \mathbf{x}) - P(S_1 = 1 | \mathbf{X} = \mathbf{x}) \tag{27}$$

$$\lambda_i = P(C = 1 | S_2 = 1, S_1 = 0, \mathbf{X} = \mathbf{x}) = \frac{P(S_2=1, C=1 | S_1=0, \mathbf{X}=\mathbf{x})}{P(S_2=1 | S_1=0, \mathbf{X}=\mathbf{x})} = \frac{P(S_2=1 | S_1=0, \mathbf{X}=\mathbf{x}) - P(S_1=1 | \mathbf{X}=\mathbf{x})}{P(S_2=1 | S_1=0, \mathbf{X}=\mathbf{x})} \tag{28}$$

The result in equation (27) is intuitive. It states that the probability of being sterilized because of the government campaign equals the increase in the probability of sterilization in the years 1996 and 1997 (during the campaign years) in comparison to the previous years. For example, if a woman with certain characteristics (family size, age, education, place of residency, etc.) has a probability of 0.09 of being sterilized prior to start of the campaign, but her probability of sterilization increases to 0.12 during in 1997, then we would estimate her probability of participating in the campaign as 0.03, following equation (27). And if a sterilization is reported, we would estimate that the probability that it occurred because of the government campaign is 0.25 (i.e. 0.03/0.12), following equation (28). In other words, Condition 3 implies that the increases in the incidence of sterilizations in the years 1996 and 1997 shown in Figure 1 are the result of the government campaign. In practice, however, as described in our estimation strategy below, we also allow for a time trend to explain some of the increase in the probability of sterilizations in the years 1996 and 1997.

To implement this empirical strategy we use the date of sterilization and other retrospective variables in the DHS to construct a longitudinal history for each woman describing her fertility and marital time path from 1990 to the end of the government campaign in 1997. Each woman has one observation for each year (indexed by t). We record sterilization in each year with a dichotomous variable S_{it} and once a woman is sterilized she drops from the panel. This rearrangement of the data allows us to estimate the probability of a woman being sterilized in each year, given that she has not been sterilized before. We fit the following model, where the *logit* function is defined as $\mathrm{logit}(z) = \ln(z/(1-z))$:

$$\mathrm{logit}\big(\mathrm{P}(\mathrm{S}_{\mathrm{it}}|\mathrm{S}_{\mathrm{it}-1} = 0, \mathbf{X}_{\mathrm{it}} = \mathbf{x}_{\mathrm{it}})\big) = \alpha + \delta\mathrm{t} + \boldsymbol{\beta_0}\mathbf{x}_{\mathrm{it}} + \boldsymbol{\beta_1}(\mathbf{x}_{\mathrm{it}} \times \mathrm{Campaign}_{\mathrm{t}}) \tag{29}$$

The term $\mathbf{x}_{\mathrm{it}}$ denotes a women's set of observed characteristics, including: geographic location, age group (categorical), age at first birth, education attainment, number of children, whether any of the children is a boy, whether she gave birth in year t, and her mother tongue (Spanish or other). We include a linear time trend ($\delta\mathrm{t}$) in the specification to account for the observed rising trend in sterilizations prior to the start of the government campaign. We posit that in the years 1996 and 1997 the probability of sterilization increased. However, it did not increase for all women, only for women with certain characteristics, who were more likely targeted by the government campaign (e.g. allegedly poor, rural women). We capture this in our specification by interacting the characteristics in $\mathbf{x}_{\mathrm{it}}$ with a dichotomous variable $\mathrm{Campaign}_{\mathrm{t}}$ that equals one if the year t is 1996 or 1997 and zero if the year t is in the prior period 1990-1994. We leave the year 1995 out of the estimation sample because part of this year can be attributed to the pre-campaign period and part to the campaign period.

Table 1 shows the logit coefficients for DHS IV and DHS V. We present the main coefficients associated with each characteristic (β_0) and the coefficients associated with the interactions (β_1), which measure the change in the variable coefficient in the years 1996 and 1997. In our model specification, we only keep the interactions that are statistically significant in DHS IV and DHS V logit models, as shown in Table 1. We exclude interaction terms with statistically insignificant coefficients as the point estimates of some of those coefficients were negative. Including the negative interactions led to lower predicted probabilities of sterilization during the campaign period for women with certain combinations of observed characteristics. This creates a problem in the application of our empirical strategy since Condition 3 rules out the probability that the risk of sterilization decreases for any women during the years 1996 and 1997. However, all of the statistically significant coefficients associated with the interactions terms are positive indicating that there was an unambiguous increase in the probability of sterilization during the years 1996 and 1997 for women with certain characteristics, as it would be required by Condition 3.[8] In particular, women under 30 years old, less educated women and women in the rural areas of the Andes and Amazon experienced an increased likelihood of sterilization in the years 1996 and 1997 consistent with allegations about who was targeted by the campaign. The probability of sterilization in 1996 and 1997 also increased for women who did not give birth in those years, indicating that a larger share of sterilizations during the campaign period were not performed at the time of a delivery. Sterilization at the time of delivery is common practice with planned sterilization. Sterilizations that do not coincide with a birth event are consistent with women being encouraged to undergo the procedure during sporadic health fairs or during visits from mobile sterilization teams.

[Table 1 here]

Using the estimation results of equation (29), we can predict for women who were not yet sterilized as of 1996 their probability of sterilization in 1996 and in 1997, as shown in equation (30) where the *expit* function is defined as $\text{expit}(\text{z}) = e^z/(1+e^z)$:

$$\text{P}(\text{S}_{\text{it}} = 1 | \text{S}_{\text{it}-1} = 0, \mathbf{X}_{\text{it}} = \mathbf{x}_{\text{it}}) = \text{expit}[\alpha + \delta\text{t} + (\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2)\mathbf{x}_{\text{it}}] \tag{30}$$

[8] In alternative specifications of the logit model we added a main effect for the indicator variable $Campaign_{it}$, but the estimated coefficient was negative, small and not statistically significant.

And under Condition 3, the probability of sterilization for women not influenced by the government campaign equals to the predicted probability we obtain after "turning off" the interaction term. In other words, it is equal to:

$$P(S_{it} = 1, C_i = 0 | S_{it-1} = 0, \mathbf{X}_{it} = \mathbf{x}_{it}) = \text{expit}[\alpha + \delta t + \boldsymbol{\beta_1} \mathbf{x}_{it}] \qquad (31)$$

With these probabilities we construct the weights λ_i and θ_i that are employed in our IPW estimator:[9]

$$\lambda_i = \frac{\text{expit}[\alpha+\delta t+(\boldsymbol{\beta_1}+\boldsymbol{\beta_2})\mathbf{X}_{it}]-\text{expit}[\alpha+\delta t+\boldsymbol{\beta_1}\mathbf{X}_{it}]}{\text{expit}[\alpha+\delta t+(\boldsymbol{\beta_1}+\boldsymbol{\beta_2})\mathbf{X}_{it}]} \qquad (32)$$

$$\theta_i = \frac{\text{expit}[\alpha+\delta t+\boldsymbol{\beta_1}\mathbf{x}_{it}]}{1-\text{expit}[\alpha+\delta t+\boldsymbol{\beta_1}\mathbf{x}_{it}]} \qquad (33)$$

6. Empirical findings

Table 2 and Table 3, based on DHS IV and DHS V respectively, show the sample means of demographic variables before and after applying the weights we developed in Sections 3 to 5. Column 1 shows the unweighted sample means for all women with at least one child who were not sterilized as of the end of 1995 (no women without children report sterilization in the data); column 2 shows the unweighted sample means for women who were sterilized in 1996-1997; and column 3 shows the unweighted sample means for women (with at least one child) who were not sterilized in those years. We observe that sterilized women are older, have more children, are slightly less educated and are less likely to live in rural mountain areas than non-sterilized women. However, once we apply the proposed weights we find that there are important differences in the average characteristics of women we estimate are more likely to have been

[9] Since the campaign spanned two years (1996 and 1997), each woman in the control group (i.e. not sterilized in that period) has in practice two estimates of the predicted probabilities in equations (32) and (33). We use the average of these two predicted probabilities in our analysis. For women sterilized in 1996 we use the predicted probabilities in that year (since those women are dropped from the data afterwards); whereas for women sterilized in 1997 we use the average of the predicted probabilities in 1996 and1997.

sterilized because of the government campaign (column 4 in Table 2 and Table 3) compared to women who are more likely to have been sterilized without any influence from the campaign (column 5 in Table 2 and Table 3), in line with the results of the logit models in Table 1. We estimate that women sterilized because of the government campaign were relatively younger, less educated and considerable more likely to live in rural areas, especially in the mountain (Andes) and jungle regions of Peru. To obtain these weighted means, we selected the sample of women sterilized in 1996 and 1997 and weight them according to their probability of being sterilized because of the government campaign λ_i (column 4) and their probability of opting for sterilization without influence from the government campaign $1 - \lambda_i$ (column 5).

[Table 2 and Table 3 here]

Before presenting the estimated impact of the campaign, we first calculate the number of women that were affected by it. We estimate the number of total female sterilizations in a given year with the estimator $\sum_{i=1}^{N} s_i \phi_i$ and the total number of sterilizations due to the campaign with the estimator $\sum_{i=1}^{N} s_i \phi_i \lambda_i$. Using information from DHS IV, which has a shorter recall period, we estimate that 159,709 women were sterilized in Peru in the years 1996-1997, and of them about 66,981 women were sterilized because of the campaign. In other words, 42% of the sterilizations that occurred during the years 1996-1997 were part of the government campaign; or alternatively, the campaign increased the number of sterilizations by 72%.

Table 4 shows the estimated impact of the campaign on fertility and other outcomes. Columns 1 and 2 give results for outcomes from DHS IV which was collected in 2000, i.e. four years after the start of the campaign, and allow us to evaluate medium-term outcomes. Columns 3 and 4 show estimates for outcomes from DHS V which was collected from 2004 to 2008, i.e. eight to twelve years after the start of the campaign, and allow us to evaluate long-term outcomes. Column 1 and 3 provide estimates of the impact of the campaign, using the estimator $\widehat{\psi}_1$ from equation (22). Column 2 and 4 provide estimates of the effects of sterilization for women who were sterilized outside the campaign using the estimator $\widehat{\psi}_2$ from equation (24). The standard errors were obtained using bootstrap methods. For example, to obtain the standard errors in column 1, we sampled with replacement women from DHS IV. Then for the sampled women we reconstructed their longitudinal history and estimated the logit model in equation (29), and use the logit model to construct the weights λ_i and θ_i and estimate $\widehat{\psi}_1$. In total we ran 500 bootstrap replications. We followed a similar procedure to estimate the standard errors of $\widehat{\psi}_1$ in DHS V (column 3) and to estimate the standard errors of $\widehat{\psi}_2$ in DHS IV (column 2) and DHS V (column 4)

As seen in Table 4, we find substantially different impacts of sterilization on fertility for women who we estimate were sterilized because of the campaign (columns 1 and 3) compared to women we estimate were sterilized outside of the campaign (columns 2 and 4). By 2000, women likely to have been sterilized because of the campaign had on average 0.42 fewer children than similar women who were not sterilized in 1996 and 1997. This difference increases to 0.95 fewer children by 2004 at which time the average affected woman was approaching 40 years of age and we thus take this to be an estimated drop in completed fertility of approximately one child (both effects are significant at the 1% level). In comparison, the effect of sterilization on the fertility rates for women sterilized in 1996 and 1997 outside the campaign are much smaller, on average 0.12 fewer children by 2000 (and not statistically significant) and 0.64 fewer children by 2004 (statistically significant at the 1% level).

[Table 4 here]

We also find that by 2000, the probability of working for pay among women likely to have been sterilized because of the campaign increased by 2.3 percentage points, although the estimate is not statistically significant. Similarly small and insignificant estimates of the impact on paid work are obtained for DHS V in 2004-2008.

Next we examine outcomes for children who were born before 1996. 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 may have gone on to have younger siblings. This kind of comparison would allow examine whether a quality/quantity trade-off (Becker and Lewis (1973)) mechanism is evident. First we estimate the impact of sterilization on height for age in standard deviations from the reference median, which is a long-term measure of health. The DHS only records this biometric information among children under age four. Thus, Table 4 shows only height for age results for DHS IV, since children born before 1996 in DHS V are over four years old by the time of the survey. Furthermore, since children's characteristics were not included in the logit models, we use an IPW-regression adjusted estimator controlling for children's age.[10] Impacts on height for age are positive and significant at the 5% level for children whose mothers we estimate were sterilized because of the campaign. The effects are also positive and of similar magnitude but only significant at the 10% level for children whose mothers are likely to have been sterilized outside of the campaign.

Interestingly, a closer examination reveals that the improvements in health are concentrated among girls. Sterilization of their mothers has a large, positive and strongly significant effect on height for age among girls under the age of four. The estimates are similar for daughters of

[10] IPW-regression adjusted estimators are obtained through a weighted ordinary least square (OLS) regression of the outcome of interest on mother's sterilization in 1996-1997 (S_i) and additional child-level covariates, and weighting observations in the sterilized group ($S_i = 1$) by λ_i and in the non-sterilized group ($S_i = 0$) by $\theta_i \lambda_i$.

women likely to have been sterilized because of the campaign and women likely to have been sterilized without influence from the campaign.

We also investigate the effect of sterilization on education for children born prior to 1996. The DHS only records information on schooling for male children under the age of 15, but records information on schooling for all females in the household. Thus, we do separate analysis for boys and girls under the age of 15 and for girls over the age of 15. For the first group, we find small but statistically significant positive impacts of the sterilization on schooling. Using DHS V we find 0.13 additional years of schooling for daughters and 0.17 additional years for sons of women likely to have been sterilized because of the government campaign, with the effect for boys reaching statistical significance at the 5% level. The effect for daughters and sons of women likely to have been sterilized outside of the campaign are similar in magnitude and in statistically significance. For daughters 15 and older, find no statistically significant effects of sterilization on years of schooling.

7. Discussion

To evaluate the validity of our estimates we need to assess several factors. First, we need to assess whether our proposed estimator and empirical approach to calculate the weights λ_i and θ_i deliver unbiased estimates of $\widehat{\psi}_1$, given that the conditions described in Section 4 and Section 5 (the strong ignorability condition, the overlap condition, and the stationarity in the probability of sterilization without influence from a government campaign) are met. Second, we need to assess whether the assumed conditions are plausible in our application. And finally, we need to assess the potential limitations of the data. In this section we discuss each of these factors.

To assess the first issue, whether our estimator and empirical approach deliver unbiased estimates, we implement a Monte Carlo study with simulated data. The details of the study are available in the Appendix. We find that our estimator and empirical approach perform very well with an average bias close to zero. In the few scenarios when the average bias is more than 1%, we also observe a higher variability in the estimates so that the (standardized) effect size of the average bias remains small and non-statistically significant. Thus, our simulations indicate that our proposed estimator and our empirical approach are able to recover unbiased causal impact of sterilization for the group of interest (in this case women sterilized because of the government campaign) even if group membership is not directly observed, as long and group membership can be predicted using auxiliary information (in this case past behavior of observationally similar women).

Next, we assess each of the conditions needed for causal interpretation of our estimates. We start with the strong ignorability condition (Condition 1). Although there is no way to directly test this condition, we test whether we obtain balance on characteristics between sterilized women who are likely to have been sterilized because of the government campaign and unsterilized women (as of 1997) who are observational similar to them. We do this analysis for both variables that were included in our logit models (equation 29) and for variables that were not included in the logit models. If these variables appear balanced among the two groups of women, then there is less concern of unequal distribution of unobserved variables that might correlate with sterilization and its effects on outcomes.

Table 5 presents the results of the balancing tests. In column 1 we weight the sample of women in DHS IV who were sterilized in 1996 and 1997 according to their probability of having been sterilized because of the government campaign, i.e. λ_i. In column 2 we weight the sample of women in DHS IV who were not sterilized in 1996 and 1997 according to how observationally similar they are to women who we estimate were sterilized in those years because of the government campaign, i.e. $\theta_i\lambda_i$. Columns 4 and 5 present the same calculations for DHS V. We obtain relatively balanced characteristics between the two (weighted) groups of women. We are particularly encouraged by the fact that the balancing in covariates is not limited to the variables included in the logit models. In particular, the last two rows of Table 5 are the weighted average responses to DHS survey questions about the wantedness of the woman's last pregnancies. We only consider pregnancies before 1997 and thus we can only do this analysis using DHS IV since the information is available for pregnancies in the last five years. The percentage of women who did not want their last pregnancy is 55% for the unweighted sample of sterilized women and 31% for the unweighted sample of unsterilized women (see last row of columns 2 and 3 in Table 2). We find that the weights λ_i and $\theta_i\lambda_i$ help to close much of the difference in the unweighted means. The weighted percentage of sterilized women that did not want their last pregnancy is 57%, whereas the weighted percentage of non-sterilized women is 50% (Table 5). The remaining difference between is not statistically significant at the 5% level. We also find that the weights reduced the difference in the percentage of women that wanted their last pregnancy but at a later time. The unweighted means are 14% and 22% for sterilized and non-sterilized women, respectively (Table 2). In comparison, the weighted means are 15% for both groups of women (Table 5). Since these variables are not included in the logit models, this is suggestive evidence that, while we are matching on observed characteristics, our treatment and control group may also match on unobserved that are likely to correlate with sterilizations and potential outcomes. This is reassuring for the causal interpretation of our treatment effects

[Table 5 here]

Next we assess the overlap condition. Figure 2 and Figure 3 present the distribution of the probabilities $P(S_{it} = 1|S_{it-1} = 0, X_{it} = x_{it})$ and $P(C_i = 1|S_{it} = 1, S_{it-1} = 0, X = x_{it})$ for women sterilized in 1996 and 1997 and for women not sterilized in those years. These probabilities are the key components in the weights λ_i and θ_i. It can be observed that there is a reasonable overlap in both probabilities for both groups of women in DHS IV and DHS V. Also, the condition $P(S_{it} = 1|S_{it-1} = 0, X_{it} = x_{it}) < 1$ is satisfied in both datasets.

[Figure 2 and Figure 3 here]

Next, we assess the plausibility of Condition 3, stationarity in the probability of sterilization without influence from a government campaign. Previous research has indicated that after its authorization as a standard family planning method in 1995, sterilization was mainly promoted and implemented by the public health sector, which resulted in a dramatic increase in the family planning market share for the Ministry of Health (Gribble, Sharma and Menotti, 2007). In other words, sterilization was not given preference in the private sector and thus we can attribute most of the uptake after its legalization to the sterilization campaign in the public sector. Moreover, our logit models indicate that after 1996 the probability of sterilization changed statistically (an increase) only for marginalized women (low educated and living in rural areas of the mountain and jungle regions). Both of these observations imply that in 1996 and in 1997, in comparison to the previous years, the incidence of sterilizations *not* encouraged by the campaign remained stable. In other words, the probability of sterilizations among women who were unlikely to be targeted by the campaign did not change. Thus, evidence is supportive of the plausibility of Condition 3.

Finally, we assess the limitations of the data for our purposes. First, there may be a recall bias arising from women erroneously reporting their year of sterilization, which would affect the sample composition in our estimation as well as the estimated probabilities and weights. Second, some of the more marginalized women might not be aware of whether they have been sterilized. As noted in Section 2, there are reports that some sterilizations were performed without women's knowledge, especially at the time of delivery. Third, some sterilizations were done without adequately trained personnel or equipment (Coe, 2004), which might have resulted in complications and even death. The DHS data do not include such cases. These data issues could introduce bias in our estimated results. Although it cannot be proven, it is likely that these limitations would result in smaller estimates of the impact of the campaign. For example, misreporting the year of sterilization would result in misclassifying women sterilized because of the campaign as being sterilized prior to the campaign or unsterilized as of 1997. As a consequence, we would erroneously assign them smaller weights in the IPW estimator or we may assign them entirely to the control group. Similarly, misclassifying sterilized women as unsterilized due to lack of knowledge would result in smaller estimated impacts of sterilization because they are assigned to the control group.

8. Conclusion

We evaluate the impact of a large-scale government sterilization campaign in Peru in 1996 and 1997, which focused mostly on female surgical sterilization. We use a modified inverse probability weighting (IPW) estimator to tackle the challenge of contaminated treatment data—while we know who was sterilized during the campaign period, we do not know who was sterilized because of the campaign and who opted for sterilization without influence from the campaign. The proposed method can be applied to other situations where researchers are interested in evaluating the effect of an intervention for a particular subgroup--for instance individuals that were more compliant with the intervention--but the membership to the subgroup of interest is not observed in the data. The key requirement of the estimator is the ability to construct (from auxiliary data) a conditional probability of belonging to the unobserved subgroup of interest if the contaminated treatment is observed. This conditional probability is then used as an additional weight in the IPW estimator. In this study, we use information available (from retrospective data) on women's probability of sterilization in the years prior to 1996 to estimate the conditional probability that if a woman was sterilized in 1996 or 1997, it was because of the government campaign. Using this conditional probability, we apply the modified IPW estimator to evaluate the impact of the sterilization campaign on fertility and household wellbeing.

We find that the group of women more likely to have been sterilized because of the campaign was different on average from the group of women who were more likely to have been sterilized outside of the campaign. In particular, women more likely to have been sterilized because of the campaign were less educated and more likely to live in rural areas, especially in the mountain and jungle regions of Peru. We also find that women more likely to have been sterilized because of the campaign, had on average 0.42 fewer children by 2000, and 0.95 fewer children by 2004, in comparison to similar women who were not sterilized in 1996 and 1997. We find that sterilization has a smaller impact on fertility for women who were more likely to have been sterilized outside of the campaign in 1996 and 1997. They had 0.12 fewer children by 2000 and 0.64 fewer children by 2004 than comparable women not sterilized in those years, which is consistent with those women having knowledge about and access to a broader range of contraceptive options

We find that the campaign did not lead to an increase in the probability working for pay. However, we find that it led to substantial and statistically significant improvements in the height for age (a measure of health) of girls whose mothers we estimate were sterilized because of the campaign. We also find small but positive and statistically significant effects of the campaign on years of schooling for boys under the age of fifteen. In both cases, the size of the effects are larger than those found for children of women likely to have been sterilized outside the campaign, although the differences are smaller than in the case of fertility.

How should we view these findings of decreased fertility and evidence of improved health and education outcomes for children as a result of the Peruvian sterilization campaign? Fujimori claimed to both domestic and international audiences that his family planning campaign was a tool to break the "vicious circle [of] poverty—unwanted child-poverty."[11] These claims hint at mechanisms long debated in the economic demography literature--the causal link between family planning interventions and fertility decline and the related issue of if and how fertility decline impacts family wellbeing.

Grant Miller and Kimberly Singer Babiarz (2016) provide a review of the evidence on the causal impacts of family planning programs. They highlight the debate between the "supply" and "demand" sides of the literature: Scholars such as Bongaarts (1994) argue that by lowering the cost and increasing the supply of contraception, family planning programs drive down fertility. Pritchett (1994) argues, on the contrary, that while increased contraceptive use may be coincident with falling fertility; it is the demand for smaller families that causes the decline. Resolving this debate empirically is notoriously difficult. Miller and Singer Babiarz conclude that the existing evidence based on microdata shows that family planning programs in middle and low income countries do play a causal role in reducing the number of children ever born

[11] Statement of President Alberto Fujimori of Peru at a Special Session of the General Assembly 30 June - 2 July 1999, United Nations Headquarters, New York: http://www.un.org/popin/unpopcom/32ndsess/gastatements.htm

between 5 and 35 percent, but that this only explains between four and 20 percent of observed fertility decline. The estimated impacts of the compulsory, quota-based policies in China bound the upper ends of those ranges. The authors also note that the argument for family planning policies directly causing fertility decline are stronger "for poor households facing tight credit or liquidity constraints." This reasoning is in line with our finding that the impact of sterilization was greater for the less-educated rural women more likely to be targeted by the campaign than for the more affluent women who likely faced a more complete menu of contraceptive options and opted for sterilization with no influence from the campaign.

Using a contraceptive campaign as a poverty reduction strategy relies on the notion that having fewer children will improve the economic prospects of the household. As specifically relates to our findings of improved health and education outcomes for existing children of sterilized women, Becker and Lewis (1973) hypothesized that the fewer children parents have the more they invest in each child. Given that the theory envisions a household making a *joint* decision about family size and investments in human capital, it is not surprising that this so-called quality-quantity tradeoff has proven difficult to test empirically. Schultz (2007) conducts an extensive review of the existing examples in the literature that use plausibly exogenous changes in fertility (in most cases based on twinning or gender composition of previous birth) to test for a causal link between fertility decline and investments in child health and education. The evidence is mixed with some studies finding a negative causal relationship between number of children and investments in human capital; and others finding no significant relationship. Schultz concludes that "[t]he trade-off between the quantity of children a woman bears and the quality of those children is viewed by many as a stylized fact, but this behavioral regularity may be more common in high-income urban societies than elsewhere;" and indicates the need for further evidence from low and middle income economies. Evidence that is more in line with our findings comes from Liu (2014), who uses the relaxation of China's one-child policy to show that having more siblings has a negative causal impact on children's height which he hypothesizes is due to lower nutritional intake per child given constrained household budgets. When the level of family wealth is incorporated, the impact is only on girls' height, which the author argues is due to strong son preference in China. Researchers have found similar evidence of son preference in the Andean region of Peru (the region most affected by the sterilization campaign) in the allocation of health care (Larme, 1997) and that changes in household welfare affect girls more than boys (Ilahi, 2001).[12]

Finally, regarding the effect of increased access to contraception methods and maternal labor supply, recent research in both the United States (Bailey, 2006) and Colombia (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. In contrast to these results, our findings in Peru suggest that the mere reduction in fertility rates allowed by sterilizations, without changes in the ability to optimally plan the timing of births, does not lead to increases in the probability of working for pay.

[12] However, the evidence is mixed. For example, Graham (1997) finds no gender differences in food allocation in rural Peruvian households.

References

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*, 289-320.

Becker, Gary S and H Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy*, 81(2), S279-S88.

Boesten, Jelke. 2007. "Free Choice or Poverty Alleviation? Population Politics in Peru under Alberto Fujimori." *Revista Europea de Estudios Latinoamericanos y del Caribe/European Review of Latin American and Caribbean Studies*, 3-20.

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

Botosaru, Irene and Federico H Gutierrez. 2014. "Difference-in-Differences When Treatment Status Is Observed in Only One Period." *Available at SSRN 2377483*.

Chen, Xiaohong; Han Hong and Denis Nekipelov. 2011. "Nonlinear Models of Measurement Errors." *Journal of Economic Literature*, 49(4), 901-37.

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

del Aguila, Ernesto Vasquez. 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.

Graham, Margaret A. 1997. "Food Allocation in Rural Peruvian Households: Concepts and Behavior Regarding Children." *Social Science & Medicine*, 44(11), 1697-709.

Gribble, James N; Suneeta Sharma and Elaine P Menotti. 2007. "Family Planning Policies and Their Impacts on the Poor: Peru's Experience." *International family planning perspectives*, 33(4), 176-81.

Horowitz, Joel L and Charles F Manski. 1995. "Identification and Robustness with Contaminated and Corrupted Data." *Econometrica: Journal of the Econometric Society*, 281-302.

Hotz, V Joseph; Charles H Mullin and Seth G Sanders. 1997. "Bounding Causal Effects Using Data from a Contaminated Natural Experiment: Analysing the Effects of Teenage Childbearing." *The Review of Economic Studies*, 64(4), 575-603.

Ilahi, Nadeem. 2001. "Children's Work and Schooling: Does Gender Matter? Evidence from the Peru Lsms." *Evidence from the Peru LSMS (December 2001). World Bank Policy Research Working Paper*, (2745).

Larme, Anne C. 1997. "Health Care Allocation and Selective Neglect in Rural Peru." *Social Science & Medicine*, 44(11), 1711-23.

Leon, Federico R. 1999. "Peru: Providers' Compliance with Quality of Care Norms," Lima: Population Council,

Liu, Haoming. 2014. "The Quality–Quantity Trade-Off: Evidence from the Relaxation of China's One-Child Policy." *Journal of Population Economics*, 27(2), 565-602.

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

Miller, Grant and Kimberly Singer Babiarz. 2016. "Family Planning Program Effects: A Review of Evidence from Microdata." *Center for Global Development Working Paper*, (422).

Mosher, Steven W. 2002. "President's Page: Peru Sterilization Campaign," *Population Research Institute.*

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

Schultz, T Paul. 2007. "Population Policies, Fertility, Women's Human Capital, and Child Quality." *Handbook of development economics*, 4, 3249-303.

Tamayo, Giulia. 1999. *Nada Personal: Reporte De Derechos Humanos Sobre La Aplicación De La Anticoncepción Quirúrgica En El Perú, 1996-1998*. CLADEM, Comité de América Latina y el Caribe para la Defensa de los Derechos de la Mujer.

Figure 1. Number of Reported Sterilizations by Year
Peruvian Demographic and Health Surveys

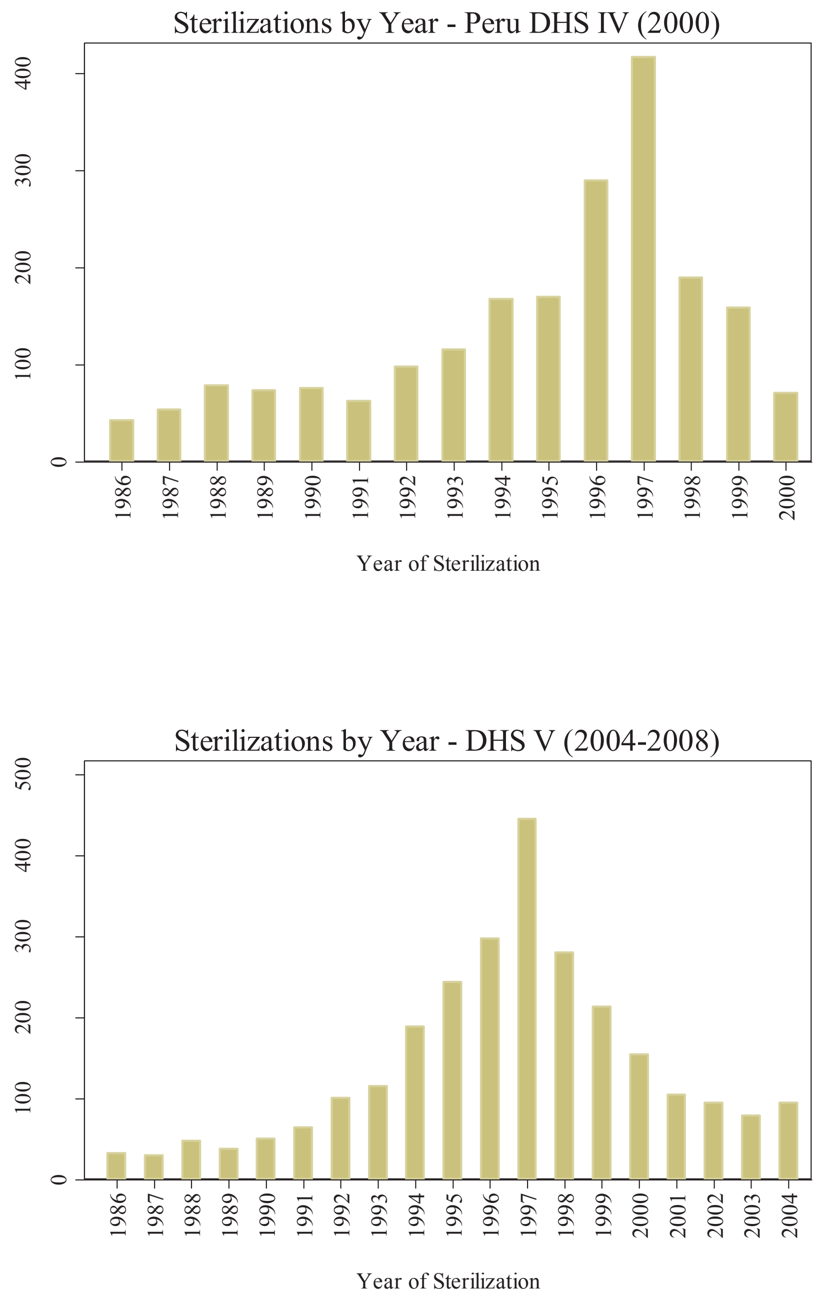

Figure 2. Common Support for Weighting Variables by Sterilization Status in 1996-1997 (DHS IV)

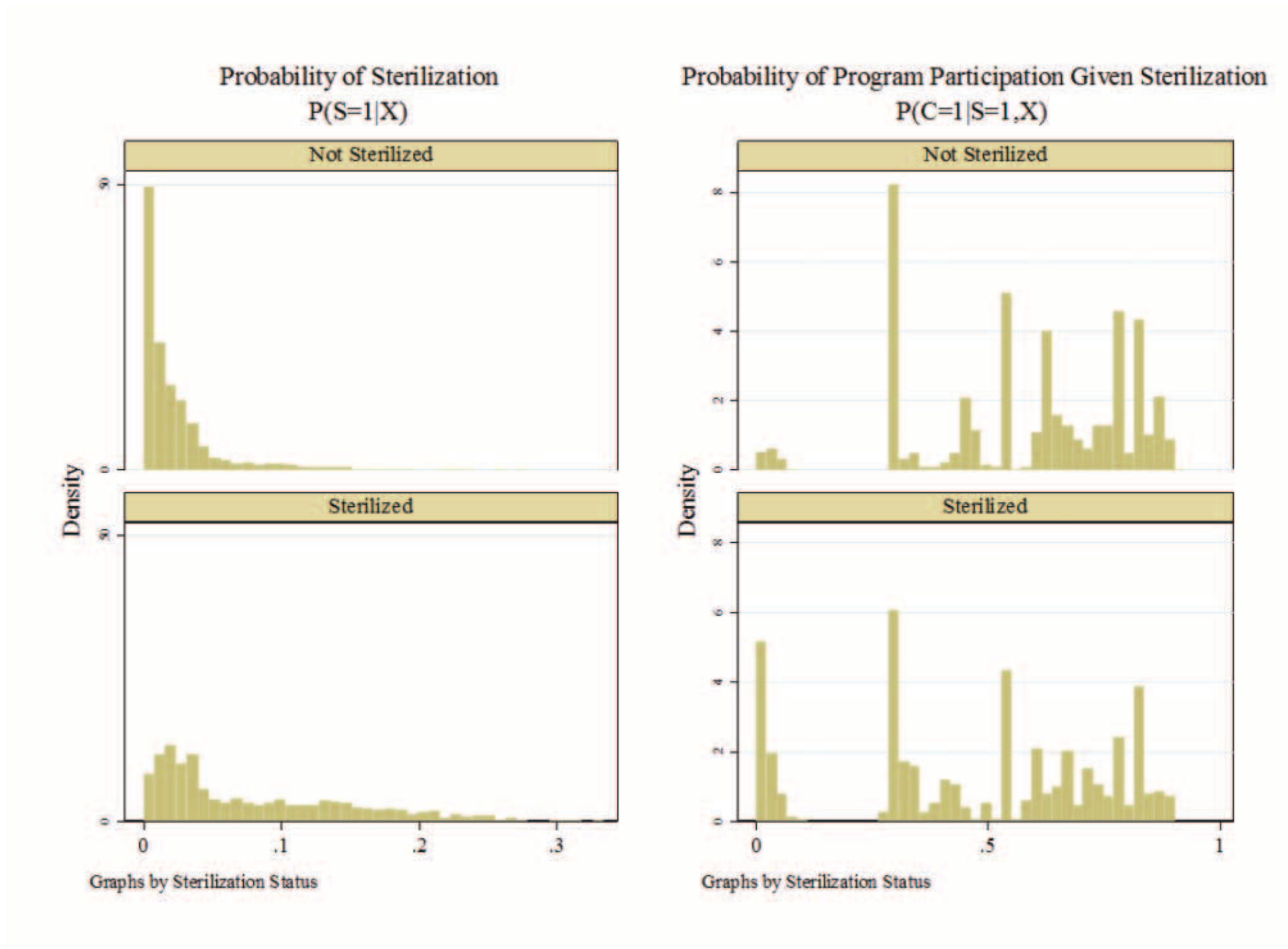

Figure 3. Common Support for Weighting Variables by Sterilization Status in 1996-1997 (DHS V)

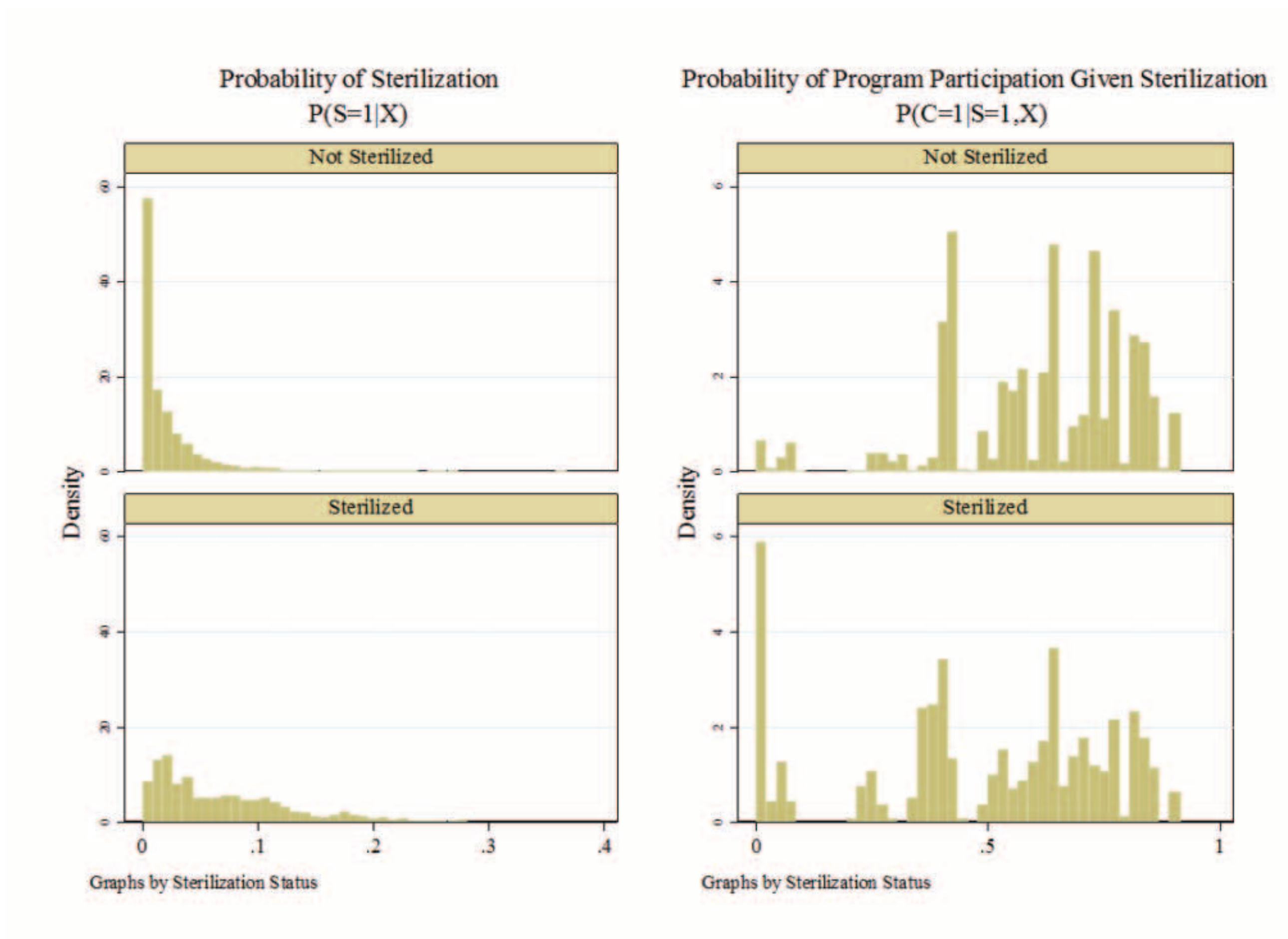

Table 1. Determinants of Sterilization Prior to and During the Fujimori Sterilization Campaign (Logit Coefficients)

Variables	DHS IV		DHS V	
	Main effects	Interactions effects	Main effects	Interactions effects
Number of kids				
1	-4.903***		-5.826***	
	(0.000)		(0.000)	
2	-1.797***		-2.065***	
	(0.000)		(0.001)	
3	-0.649		-0.463	
	(0.170)		(0.426)	
Age at first birth	0.038***		0.065***	
	(0.006)		(0.000)	
Rural Coast	-0.032		0.281*	
	(0.811)		(0.064)	
Urban Mountain	-0.343***		-0.404***	
	(0.002)		(0.001)	
Rural Mountain	-1.568***	0.717***	-1.163***	0.444*
	(0.000)	(0.002)	(0.000)	(0.053)
Urban Jungle	0.003		0.028	
	(0.977)		(0.810)	
Rural Jungle	-1.417***	0.923***	-1.253***	0.624**
	(0.000)	(0.001)	(0.000)	(0.046)
Primary Education	-0.716***	0.430***	-0.875***	0.516***
	(0.000)	(0.008)	(0.000)	(0.003)
Secondary Education	-0.199		-0.406***	
	(0.128)		(0.004)	
Age				
under 26	-1.159***	0.604**	-0.935***	0.762***
	(0.000)	(0.035)	(0.000)	(0.006)
26 to 29	-0.611***	0.599***	-0.246	0.309
	(0.000)	(0.003)	(0.116)	(0.114)
over 36	-0.357***		-0.442***	
	(0.001)		(0.003)	
First language not Spanish	-0.662***		-0.469***	
	(0.000)		(0.002)	
No male children	-0.167		-0.449***	
	(0.269)		(0.004)	
No birth that year	-2.400***	0.363**	-2.396***	0.540***
	(0.000)	(0.020)	(0.000)	(0.002)
Year (linear trend)	0.149***		0.148***	
	(0.000)		(0.000)	
Constant	-2.066***		-2.906***	
	(0.000)		(0.000)	
Observations	*84,717*		*93,734*	

Notes: Main effects refer to coefficient β_1 from equation 21. Interaction effects refer to coefficient β_2. Indicator variables for number of kids greater than 3 (up to 9) were included in the regressions but are not reported for the sake of space. Coefficient on the constant term is also omitted for space. P-values are included in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 2. Characteristics of Women Sterilized in Not Sterilized in 1996-1997, DHS IV

	Means among all women not sterilized and with at least one child by 1995, by sterilization status in 1996-1997 (Unweighted)			Means among all women sterilized in 1996-1997 (Weighted by the conditional probability of belonging to each group)	
	all	S = 1	S =0	S=1, C = 1	S=1, C = 0
	(1)	(2)	(3)	(4)	(5)
Pre-Campaign Characteristics					
Age in 1996	31.5	32.4	31.4	31.1	33.3
# Kids in 1996	2.9	3.9	2.8	4.0	3.8
Years of education	7.9	7.1	7.9	5.4	8.4
Age at first birth	20.7	20.1	20.7	19.1	20.8
Geography					
rural	35.9%	35.2%	35.9%	53.3%	21.8%
coast	51.3%	58.3%	51.0%	42.2%	70.2%
mountain	35.7%	26.5%	36.2%	36.9%	18.7%
jungle	13.0%	15.3%	12.9%	20.9%	11.1%
urban coast	45.4%	48.1%	45.3%	33.3%	59.1%
rural coast	5.9%	10.1%	5.7%	8.8%	11.1%
urban mountain	12.5%	9.4%	12.7%	7.2%	11.0%
rural mountain	23.2%	17.1%	23.5%	29.7%	7.7%
urban jungle	6.2%	7.3%	6.1%	6.2%	8.2%
rural jungle	6.8%	7.9%	6.7%	14.8%	2.9%
Among pre-1997 pregnancies:					
Wanted last Pregnancy Later	21.6%	14.4%	22.3%	15.4%	13.8%
Did not Want last Pregnancy	33.1%	55.4%	31.0%	56.7%	54.7%

Notes: Columns 1 to 3 give unweighted mean characteristics for the sample of women with at least one child who were not sterilized as of the end of 1995. Column 1 shows the means for the full sample. Column 2 shows means for women who were sterilized in 1996-1997 (S=1). Column 3 shows means for women not sterilized in that period (S=0). Column 4 shows means for women sterilized during the campaign years weighted by the probability that they were sterilized because of the campaign (λ_i). Column 5 shows means for women sterilized during the campaign years weighted by the probability that they were sterilized for reasons other than the campaign ($1 - \lambda_i$).

Table 3. Characteristics of Women Sterilized in Not Sterilized in 1996-1997, DHS V

	Means among all women not sterilized and with at least one child by 1995, by sterilization status in 1996-1997 (Unweighted)			Means among all women sterilized in 1996-1997 (Weighted by the conditional probability of belonging to each group)	
	all	S = 1	S =0	S=1, C = 1	S=1, C = 0
	(1)	(2)	(3)	(4)	(5)
Pre-Campaign Characteristics					
Age in 1996	29.1	31.1	29.0	29.9	32.1
# Kids in 1996	2.5	3.7	2.5	3.9	3.6
Years of education	8.0	6.9	8.0	5.0	8.5
Age at first birth	20.5	20.3	20.5	19.3	21.2
Geography					
rural	34.1%	36.4%	34.0%	53.3%	22.5%
coast	50.1%	53.5%	49.9%	39.7%	64.9%
mountain	37.3%	31.9%	37.5%	41.8%	23.8%
jungle	12.7%	14.6%	12.6%	18.5%	11.4%
urban coast	44.9%	43.6%	45.0%	29.0%	55.7%
rural coast	5.1%	9.9%	4.9%	10.8%	9.2%
urban mountain	14.4%	11.5%	14.5%	9.8%	12.8%
rural mountain	22.9%	20.4%	23.0%	31.9%	10.9%
urban jungle	6.6%	8.5%	6.5%	7.9%	8.9%
rural jungle	6.1%	6.1%	6.1%	10.6%	2.4%

Notes: See notes to Table 3.

Table 4. Effects of Sterilizations during the Campaign years (1996-1997)

	DHS IV (2000)			DHS V (2004-2008)		
	(1)	(2)		(3)	(4)	
	S=1, C=1 (1)	S=1, C=0 (2)	# of obs.	S=1, C=1 (3)	S=1, C=0 (4)	# of obs.
<u>Fertility Outcome:</u>						
Number of Children	-0.423***	-0.124	14,430	-0.945***	-0.635***	16,545
	(0.085)	(0.086)		(0.109)	(0.101)	
<u>Women's Outcomes</u>						
Worked for pay?	0.023	-0.000	14,430	0.002	-0.011	16,675
	(0.029)	(0.031)		(0.033)	(0.035)	
Domestic Violence				0.042	-0.014	13,383
				(0.034)	(0.022)	
<u>Children's Outcomes</u>						
Height for age	0.251**	0.201*				
	(0.120)	(0.116)				
Girls	0.512***	0.479***	2,898			
	(0.184)	(0.153)				
Boys	-0.001	-0.123	1,457			
	(0.155)	(0.177)				
Years of Schooling if age<15 (Boys and Girls)	0.037	0.030		0.148**	0.122**	26,649
	(0.053)	(0.069)		(0.093)	(0.066)	
Girls (age< 15)	0.052	0.103	10,984	0.132	0.120	13,061
	(0.076)	(0.089)		(0.132)	(0.086)	
Boys (age< 15)	0.020	-0.029	11,529	0.172**	0.126*	13,588
	(0.069)	(0.085)		(0.109)	(0.081)	
Years of Schooling if age>15 (Only girls)				0.070	0.033	4,302
				(0.348)	(0.260)	

Notes: Each entry in columns 1 and 3 gives the average treatment effect of sterilizations that occurred because of the campaign. Each entry in columns 2 and 4 gives the average treatment effect of sterilizations that occurred for reasons other than the campaign. In the case of child outcomes we use an IPW-regression adjusted estimator which controls for child's age as discussed in footnote 8. Standard errors calculated based on 500 bootstrap replications appear in parentheses. *** denotes p-value<0.01; ** denotes p-value <.05; * denotes p-value <0.1 .

Table 5. Balancing Tests

	Sample means of sterilized (S=1) and non-sterilized women (S=0), weighted by the likelihood of being sterilized because of the campaign.					
	DHS IV			DHS V		
	S = 1 (1)	S = 0 (2)	p-value of difference (3)	S = 1 (4)	S = 0 (5)	p-value of difference (6)
Pre-Campaign Characteristics						
Age in 1996	31.1	32.4	0.00	29.9	30.4	0.05
# Kids in 1996	4.0	4.1	0.62	3.9	3.9	0.97
Years of education	5.4	5.8	0.04	5.0	5.4	0.04
Age at first birth	19.1	19.6	0.01	19.3	19.5	0.28
Geography						
rural	53.3%	50.3%	0.27	53.3%	47.9%	0.06
coast	42.2%	42.0%	0.96	39.7%	43.1%	0.26
mountain	36.9%	37.0%	0.96	41.8%	39.7%	0.46
jungle	20.9%	21.0%	0.99	18.5%	17.2%	0.48
urban coast	33.3%	34.1%	0.78	29.0%	34.3%	0.07
rural coast	8.8%	7.9%	0.55	10.8%	8.8%	0.30
urban mountain	7.2%	8.9%	0.13	9.8%	10.3%	0.75
rural mountain	29.7%	28.1%	0.53	31.9%	29.3%	0.33
urban jungle	6.2%	6.7%	0.58	7.9%	7.5%	0.72
rural jungle	14.8%	14.3%	0.78	10.6%	9.7%	0.54
Among pre-1997 pregnancies:						
Wanted last Pregnancy Later	15.4%	15.2%	0.92			
Did not Want last Pregnancy	56.7%	50.2%	0.08			

Notes: Columns 1 and 4 show means for women sterilized during the campaign years weighted by the probability that they were sterilized because of the campaign (i.e. using λ_i). Columns 2 and 5 shows means for women not sterilized during the campaign years weighted to resemble women sterilized because of the campaign (i.e. using $\theta_i \lambda_i$).

APPENDIX A: Validation using Monte Carlo simulations

We use Monte Carlo simulations to study whether our proposed estimator and empirical approach to calculate the weights λ_i and θ_i deliver unbiased estimates, given that all the conditions described in the paper are met. In other words, in our simulations we hold true the strong ignorability condition, the overlap condition, and the stationarity in the probability of sterilization without influence from a government campaign.

To mimic our estimation strategy we simulate data for two periods. In each period women have a probability of being sterilized, but in period 2 there is an increase in the probability of sterilization due to a government campaign. Importantly, the probability of sterilization in period 2 increases only for a subgroup of individuals, as evidence from our empirical estimates suggests it was the case with the sterilization campaign in Peru. We then model the effect of sterilization on outcomes, in this total fertility rate, which is also heterogeneous across subgroups. Finally we follow the estimation strategy discussed in the paper to calculate the weights λ_i and θ_i and implement our IPW estimator. We perform this simulation for various scenarios. In each scenario there are 50,000 observations and 1,000 simulations.

A.1 Simulations set up

In each scenario we first generate 5 bivariate covariates $\mathbf{X} = [x_1, x_2, x_3, x_4, x_5]$ with success probability equal to $\mathbf{p} = \{0.5, 0.2, 0.3, 0.2, 0.1\}$. Then, we simulate the probability of sterilization in period 1 (S_1) and in period 2 (S_2), using the following models, where the *expit* function is defined as $\text{expit}(z) = e^z/(1 + e^z)$:

$$P(S_1 = 1|\mathbf{X}) = \text{expit}[\alpha_0 + \mathbf{X}_i{}'\boldsymbol{\alpha}] \quad (A.1)$$

$$P(S_2 = 1|S_1 = 0, \mathbf{X}) = \text{expit}[\alpha_0 + X_i{}'\boldsymbol{\alpha} + \gamma x_{1i}] \quad (A.2)$$

Where $\alpha_0 = -2.5$ and $\boldsymbol{\alpha} = [-2, -1, 1, 1, 0.5]$. The parameter γ increases the probability of sterilization in period 2 for individuals with $x_1 = 1$. The parameter γ takes the following values $\{0.5, 1, 1.5, 2, 2.5\}$ depending of the simulation scenario. Using Condition 3, we can further decompose the probability of sterilization in period 2 as being sterilized because of the campaign ($S_2 = 1, C = 1$) or because of other reasons ($S_2 = 1, C = 0$):

$$P(S_2 = 1, C = 1 | S_1 = 0, \mathbf{X}_i) = \text{expit}[\alpha_0 + \mathbf{X}_i{}'\boldsymbol{\alpha} + \gamma_1 x_{1i}] - \text{expit}[\alpha_0 + \mathbf{X}_i{}'\boldsymbol{\alpha}] \quad \text{(A.3)}$$

$$P(S_2 = 1, C = 0 | S_1 = 0, \mathbf{X}_i) = \text{expit}[\alpha_0 + \mathbf{X}_i{}'\boldsymbol{\alpha}] \quad \text{(A.4)}$$

Using the probabilities in equations (A.1), (A.3) and (A.4) we predict each woman's sterilization outcomes in period 1, and conditional on not being sterilized in period 1, sterilization in period 2 either because of the campaign or due to other reasons. Finally, we posit that total fertility rate follows the following data generating process:

$$y_i \sim \text{Poisson}(\exp[\beta_0 + \mathbf{X}_i{}'\boldsymbol{\beta} + \pi_0 S_i + \pi_1 (S_i \times x_{1i})]) \quad \text{(A.5)}$$

Where $\beta_0 = 1$, $\boldsymbol{\beta} = [-0.5, 0.1, 0.3, -0.25, 1]$, $\pi_0 = -0.05$, and π_1, which creates heterogeneous treatment effects, can take the following values $\{-0.5, -0.2\}$ depending on the simulation scenario. Notice that sterilization S is modeled to reduce total fertility rate more for the same subgroups of women for whom their probability of sterilization increased because of the campaign (i.e. those with $x_1 = 1$). This is important because if the effects of sterilization were not heterogeneous we could evaluate the effects of the campaign by looking at the average effects of sterilization. The problem for the econometrician of not observing who is sterilized because of the campaign and who is sterilized because of other reasons is that both groups are on average different in their distribution of x_1 and effect of sterilization varies with x_1.

A.II Implementing the estimator

To implement our estimator we first need to estimate the conditional probability that if a sterilization occurred in period 2 it was because of the campaign. Following the approach described in the paper, we use the information from period 1, before the campaign started, to help pin down this conditional probability. Specifically, we fit the following model where the *logit* function is defined as $\mathrm{logit}(z) = \ln(z/(1-z))$:

$$\mathrm{logit}\big(\mathrm{P}(\mathrm{S_{it}}|\mathrm{S_{it-1}} = 0, \mathbf{X}_\mathrm{i} = \mathbf{x}_\mathrm{i})\big) = \alpha_0 + \mathrm{X_i}'\boldsymbol{\alpha} + \gamma(\mathrm{x_{1i}} \times \mathrm{Period2}) \qquad \text{(A.6)}$$

Using this logit model, we predict $\mathrm{P}(\mathrm{S_2} = 1|\mathrm{S_1} = 0, \mathbf{X}_\mathrm{i})$ and the conditional probability $\mathrm{P}(\mathrm{C} = 1|\mathrm{S_2} = 1, \mathrm{S_1} = 0, \mathbf{X}_\mathrm{i})$ using the same approach followed in the paper. In other word, these probabilities are predicted as shown in equation (A.7) and (A.8):

$$\mathrm{P}(\mathrm{S_2} = 1|\mathrm{S_1} = 0, \mathbf{X}_\mathrm{it} = \mathbf{x}_\mathrm{it}) = \mathrm{expit}[\hat{\alpha}_0 + \mathbf{X}_\mathrm{i}'\hat{\boldsymbol{\alpha}} + \hat{\gamma}\mathrm{x_{1i}}] \qquad \text{(A.7)}$$

$$\mathrm{P}(\mathrm{C} = 1|\mathrm{S_2} = 1, \mathrm{S_1} = 0, \mathbf{X}_\mathrm{i}) = \frac{\mathrm{expit}[\hat{\alpha}_0 + \mathbf{X}_\mathrm{i}'\hat{\boldsymbol{\alpha}} + \hat{\gamma}\mathrm{x_{1i}}] - \mathrm{expit}[\hat{\alpha}_0 + \mathbf{X}_\mathrm{i}'\hat{\boldsymbol{\alpha}}]}{\mathrm{expit}[\hat{\alpha}_0 + \mathbf{X}_\mathrm{i}'\hat{\boldsymbol{\alpha}} + \hat{\gamma}\mathrm{x_{1i}}]} \qquad \text{(A.8)}$$

These two probabilities are then used to construct the weights $\lambda_i = \mathrm{P}(\mathrm{C} = 1|\mathrm{S_2} = 1, \mathrm{S_1} = 0, \mathbf{X}_\mathrm{i})$ and $\theta_i = \frac{\mathrm{P}(\mathrm{S_2}=1|\mathrm{S_1}=0, \mathbf{X}_\mathrm{it}=\mathbf{x}_\mathrm{it})}{1-\mathrm{P}(\mathrm{S_2}=1|\mathrm{S_1}=0, \mathbf{X}_\mathrm{it}=\mathbf{x}_\mathrm{it})}$, which we use in the IPW estimator. Because we can simulate the counterfactual outcomes in the case of sterilization ($\mathrm{Y_1}$) and no sterilization ($\mathrm{Y_0}$) and because we observe in the generated data whether the woman was sterilized because of the campaign or not, we can also compute the true average effect of sterilization for the group of women that were sterilized due to the campaign. In other words, we can calculate the true value of $\psi_1 = \mathrm{E}[\mathrm{Y_1} - \mathrm{Y_0}|\mathrm{S} = 1, \mathrm{C} = 1]$, following the notation of Section 3. We use this true value to compute the average bias of our proposed finite sample estimator of ψ_1, given in equation (A.9). For comparison purposes, we also estimate the overall average effect of sterilization, or $\psi_3 = \mathrm{E}[\mathrm{Y_1} - \mathrm{Y_0}|\mathrm{S} = 1]$ following the notation of Section 3. This estimate is implemented using the finite sample estimator given in equation (A.10), which is a standard IPW estimator. We then compute what would be the average bias if we use the estimator or ψ_3 to estimate ψ_1.

$$\widehat{\psi}_1 = \sum_{i=1}^{N} y_i \left(\frac{s_{2i}\lambda_i}{\sum_{i=1}^{N} s_{2i}\lambda_i}\right) - \sum_{i=1}^{N} y_i \left(\frac{(1-s_{2i})\theta_i\lambda_i}{\sum_{i=1}^{N}(1-s_{2i})\theta_i\lambda_i}\right) \tag{A.9}$$

$$\widehat{\psi}_3 = \sum_{i=1}^{N} y_i \left(\frac{s_{2i}}{\sum_{i=1}^{N} s_{2i}}\right) - \sum_{i=1}^{N} y_i \left(\frac{(1-s_{2i})\theta_i}{\sum_{i=1}^{N}(1-s_{2i})\theta_i}\right) \tag{A.10}$$

A.III Simulation results

The simulation results are presented in Figure A.1 and Table A.1. The probability of sterilization in the second period because of the campaign and the overall probability of sterilization increases across the scenarios as γ increases. Note also that the true value of ψ_1 depends on parameter π_1, and it is either around 1.35 or 2.55 less children.

[Figure A.1 and Table A.1 here]

For each scenario, Figure A.1 and Table A.1 show the average bias in the 1,000 replications of using $\widehat{\psi}_1$ (Modified IPW) and of using $\widehat{\psi}_3$ (Standard IPW) to estimate ψ_1. We draw some important observations Figure A.1 and Table A.1. First, as expected, the Standard IPW delivers biased estimates of the ψ_1. In all scenarios the average biases are large in comparison to the standard deviations so we can strongly reject the null of zero average bias. Second, the average bias of the Standard IPW gets larger when the group of women sterilized because of the campaign is a smaller group in comparison to the total number of sterilized women. In other words, the average bias gets larger when $P(C = 1|S_2 = 1, S_1 = 0)$ is small. This is because the average sterilized women becomes less representative of the women sterilized because of the campaign if $P(C = 1|S_2 = 1, S_1 = 0)$ is small. And third, our simulations show that the average bias of our Modified IPW approach is small and close to zero. In the few cases when the average bias is more than 1% (scenario 2 and scenario 4) we also observe a higher variability in the estimates. Note, however, that in all scenarios the average bias is small in comparison to the standard deviations, so we cannot reject the null hypothesis of an average bias equal to zero. Thus, our simulations indicate that our proposed estimator and our empirical approach are able to estimate unbiased impact estimates of the sterilization campaign, even if we do not directly observe in the data whether a woman was sterilized because of the campaign or because of other reasons. As explained in the paper, the necessary key element for identification of this effect is the ability to predict, using auxiliary data, the conditional probability that if a woman was sterilized it was because of the campaign. In our study and in this simulation we use sterilizations in the year prior to the campaign to pin down that conditional probability.

Figure A.1. Average Bias in Simulated Scenarios

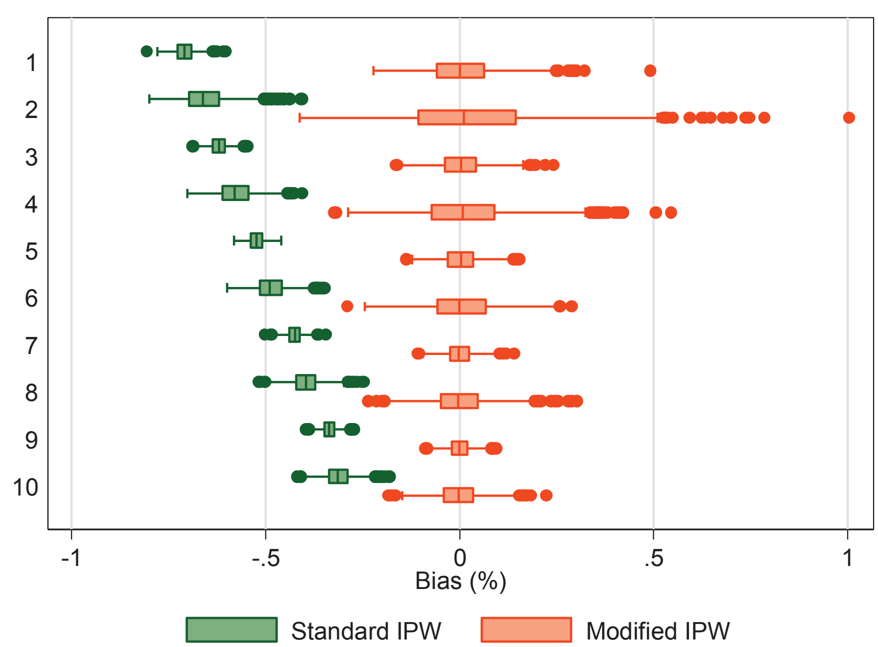

Table A.1. Monte Carlo Simulations

Scenario	γ	π_1	P(S_2=1)	P(C=1|S_2 =1)	True ψ_1	Bias of Estimator (%) Standard IPW ($\hat{\psi}_3$) Mean	Std. Dev	Modified IPW ($\hat{\psi}_1$) Mean	Std. Dev
1	0.50	-0.50	0.07	0.09	-2.59	-70.85	2.70	0.64	9.21
2	0.50	-0.20	0.07	0.09	-1.35	-65.44	6.16	3.34	19.31
3	1.00	-0.50	0.08	0.20	-2.58	-62.05	2.33	0.55	6.09
4	1.00	-0.20	0.08	0.20	-1.35	-57.57	4.96	1.30	12.54
5	1.50	-0.50	0.10	0.33	-2.57	-52.28	2.14	0.30	4.59
6	1.50	-0.20	0.10	0.33	-1.35	-48.67	4.35	0.59	9.25
7	2.00	-0.50	0.12	0.45	-2.56	-42.39	2.13	0.10	3.79
8	2.00	-0.20	0.12	0.45	-1.34	-39.57	3.98	0.08	7.42
9	2.50	-0.50	0.15	0.57	-2.54	-33.60	1.92	-0.03	3.00
10	2.50	-0.20	0.15	0.57	-1.33	-31.29	3.71	-0.15	5.91