CCTs and Fertility: Long-Term Impacts Across Two Generations

Fabien Forge*

May 15, 2021

Abstract

This paper investigates the relationship between income, education and fertility by looking at the long term impact of the Mexican conditional cash transfer program. To do so, I define two cohorts of women that were exposed to the program differently. The older cohort, received cash transfers conditional on sending their children to school, while the younger cohort received additionally extra education. Using spatiotemporal variations in the roll-out of the program at the municipality level I find suggestive evidence fertility declined for both cohorts, 13 years after the program started, and that there is no additional effect of increased schooling.

JEL codes: I380, I320, J130

Data Availability Statement: The data used in this article are available online: Integrated Public Use Microdata Series (https://international.ipums.org/international/)

Disclosure Statement: the author has nothing to disclose.

I thank Abel Brodeur and Jason Garred for their guidance in this project. I greatly benefited from discussions with Marie Connolly, Francisco Costa, Fernanda Estevan, Joanne Haddad, Philippe Kabore, Fabian Lange, Sonia Laszlo, Ted Miguel, Julia Mink, Myra Mohnen and Taylor Wright, as well as from the feedbacks of the participants of several seminars.

*Postdoctoral Fellow, Departments of Economics University of Ottawa, 120 University private, Ottawa, ON K1N 6N5, Canada. Email: fforge@uottawa.ca.

1 Introduction

Starting with Malthus' seminal work, the link between fertility and development is one of the oldest issues in the development literature, yet both the direction of the relationship and its microeconomic determinants remain unclear. Far from the Malthusian view of the world, higher fertility rates may be the consequence rather than the cause of underdevelopment (Potter et al., 2002; Galor, 2012). Conditional Cash Transfer (CCTs) programs present a unique opportunity to revisit this relationship. Twenty-three years ago, Mexico implemented such a program, *Progresa*, that combines immediate poverty relief via cash transfers with longer term goals of poverty reduction via mandatory schooling for children whose parents were enrolled.

In this paper, I investigate how income and education can influence fertility using the Mexican program as a treatment. I test whether this program has tended to reduce fertility for treated women and use the long period of time since the creation of the program to define two types of treated groups. The first group, labelled the *older cohort*, corresponds to women who were potentially enrolled in the program when it started but, because of their age, were not eligible to receive extra schooling. The second group, labelled the *younger cohort*, differ from the first by the fact that these women were young enough, when the program started, to first be treated as school-attending children, and then as parents in the same way as the *older cohort*. Thus, the *younger cohort* potentially captures the additional effect of being more educated when making fertility decisions and speaks to the inter-generational evolution of poverty.

I first motivate theoretically why and how a program such as *Progresa* may reduce fertility in the context of a standard quality-quantity trade-off model commonly used in the literature. I embed the cash transfer and the mandatory education components of the program into the overlapping generations model developed in Moav (2004). I show that the model predicts a reduction in fertility, compared with a situation with no program, for both the *younger* and the *older cohorts* provided that the cash transfer is small enough relative to the impact of the "mandatory" investment in education. I then analyse the conditions under which members of the *younger cohort* could further decrease their fertility compared to the *older cohort* as a result of their increased schooling level.

Empirically, I use a difference-in-difference strategy which takes advantage of spatiotemporal variations in the program's expansion. I link this municipality level exposure measure to the Integrated Public Use Micro Data (IPUMS) to observe the fertility levels of women across decennial censuses 13 years after the onset of the program for the post-treatment group and 7 years before for the pre-treatment group. This use of cross-census variation is dictated by the impossibility of using within census variation, especially for the *younger cohort*, since age jointly determines exposure to the treatment and the outcome variable.[1] Instead, I compare how fertility of women within a given age bracket differentially evolved between 1990 and 2010 for treated compared to control cohorts depending on the proportion of households treated changes in the municipality these women belong to. I construct this ratio using information from the Mexican evaluation organization and the Mexican National Institute of Statistics and Geography for the 2,392 municipalities in my data.

An important threat to the identification comes from the initial treatment of the program being particularly targeted at the poorest areas of Mexico. These areas being the poorest, there is a risk that fertility was already converging towards the Mexican mean, even in the absence of a treatment. This would violate the parallel trends assumption needed for the difference-in-difference estimate to be identified. To circumvent this issue I follow the strategy proposed by Parker and Vogl (2018) and use a unique feature in the roll-out of the program. Specifically, the timing of the expansion was shaped by Mexican national elections during which the law forbids the extension of social programs, resulting in three distinct waves of expansion. The first two waves were separated by 5 years and targeted similar households

[1]This is not true of other outcome variables, such as education, which no longer depends on age once the schooling years are over.

according to the "marginality index" used to chose the first recipients, while the third wave to urban and richer areas. I argue that controlling for exposure intensity in 2005 and the marginality index allows me to account for fertility dynamics linked to poverty and thus capture the true effect of the treatment.

My findings suggest that far from creating an incentive to bear more children in order to receive additional transfers, *Progresa* reduced fertility for both the *younger* and the *older cohorts* in a way predicted by my model. A one percentage point increase in the 1999 coverage of households in a municipality is associated with a decrease in fertility by 0.002 to 0.003 fewer children for the *older cohort*. In the data, these are women, aged 16 to 36 when the program started in 1997, who are observed when they were aged 29–49 in the 2010 census, compared with women of the same age in the 1990 census. For the *younger cohort* aged 7 to 15 in 1997 and observed when they are 20 to 28 in 2010, this number ranges from 0.0001 to 0.002. These results are robust to placebo tests in which the same cross-census comparisons are made for women who already reached the end of their fertility window when the program started. I also provide suggestive evidence this reduction in fertility is mainly coming from adjustments along the intensive margins.

I do not find strong evidence of an additional effect of the *young cohort*'s own education on fertility, using both a within census difference-in-difference strategy in which the *older cohort* constitutes the control group and the *young cohort* the treated group and a triple-difference strategy across censuses. This is suggestive evidence that the fertility decline experienced by both cohorts was in response to changes in the incentives relative to investment in education. The extra education received by the *younger cohort* as a result of their enrollment in the program does not seem to have additionally affected their fertility in a way that would be suggested by an inter-generational quality-quantity trade-off model. It is worth noting that since the *younger cohort* is observed, at still a relatively young age, there is a possibility that the full effect of additional education could be observed closer to the end of the fertility window of these women.

This paper builds on several strands of the literature. First, a large microeconomic literature has linked both monetary and non-economic factors such as female empowerment (access to contraceptives, education) (Duflo, 2012), level of income (Becker and Lewis, 1973), health (child mortality and life expectancy) (Doepke 2005; Jayachandran and Lleras-Muney 2009), and demand for human capital (Galor and Weil, 1999) with fertility. Second, many impact evaluation papers[2] have documented direct effects of CCT programs such as improved school indicators for boys and girls (Coady and Parker 2004; Schultz 2004; Behrman et al. 2009; Todd and Wolpin 2006) or indirect effects such as decreased child labor participation (Skoufias et al. 2001; Rubio-Codina 2010; Schultz 2004), improved health for children (Gertler 2004; Fernald et al. 2008; Fernald et al. 2009; Barham 2011) and better natal care (Urquieta et al. 2009, Barber 2009).

Several papers have looked at the impact of conditional cash transfers on fertility. Considering the short run impacts of the CCT programs of Mexico, Nicaragua and Honduras, Stecklov et al. (2007) found no effect on fertility within the 24 months following the treatment for the first two and a positive effect in Honduras. Both Baird et al. (2010) and Schultz (2004) found that teenage pregnancy was reduced a year after the program started in Malawi and Mexico. Taking a longer term perspective Barham et al. (2018) documented a reduction in teenagers' fertility for 42 communities in Nicaragua 10 years after the program started while Laszlo et al. (2019) found no effect on fertility for the Peruvian program.

Finally, this paper contributes to a nascent strand of the literature that gauges whether the identified short-run positive effects of *Progresa* has persisted since the program started in 1997 (Parker and Vogl 2018, Kugler and Rojas 2018). It also responds to the predictions made by Todd and Wolpin (2006) who predicted that treated Mexican women would keep their fertility at the same level or slightly higher.

The rest of the paper is organized as follows. Section 2 reviews the main features of the

[2]For a thorough review of the Mexican program see Parker and Todd (2017)

program, defines the *older* and *younger cohorts* and documents the main determinants of the roll-out. It then frames the link between the program and fertility using a variant of the workhorse quality-quantity trade-off model. Section 3 describes the data used to define the fertility and enrollment ratio measures used in the main specification. Section 4 presents the results for both cohorts and provides suggestive evidence on the potential mechanisms.

2 Context and Theoretical Motivation

This section provides a brief overview of an already widely studied program and links it to the workhorse quality-quantity trade-off model traditionally used when thinking about the link between poverty and fertility. In Section 2.1, I highlight the main features of *Progresa* and describe how, from a long term perspective, two analytically distinct treated groups emerge which I label the *younger* and *older cohorts*. I also describe the determinants of the roll-out of the program that inspire the identification strategy. In Section 2.2, I build on the overlapping generations model proposed by Moav (2004) and include the two main components of the program, income transfers and conditionality, to predict potential changes in fertility for both the *younger* and the *older cohorts*.

2.1 Context

Main features – The Mexican conditional cash transfer program *Progresa* started in 1997 and is, with the program *Bolsa Familia* in Brazil, one of the oldest CCT programs. This anti-poverty intervention tries to combine short-term financial support to the poor with better long-term investment in their children.[3] Specifically, the program is meant to encourage families to use existing schooling capacity for the education of their children. Thus, conditional on their children attending school regularly and receiving some health care, mothers

[3] Parker and Todd (2017) provide an extensive review of the program and the numerous papers written about it. This section will therefore focus on the main components of the program and how they relate to the research question this paper tries to address.

would receive cash transfers from the government for each child enrolled.[4] The value of these transfers were designed to match the opportunity cost of child labor.

At the heart of the program, women play a very important role. First, they are the only ones allowed to receive the transfers. Second, they were supported in their ability to care for their children's health and nutrition. They, themselves, were recipients of a series of health care interventions, including pre-natal and post-natal care. Female students also benefited from preferential treatment in the form of more generous grants. Girls continuing school after grade 7 were encouraged not to drop out by receiving higher grants than their male counterparts.

The program was also designed to prevent parents from abusing it by having additional children receive benefits. In particular, a cap of three was placed on the number of children that could be enrolled in the program. In addition, in the first years after entering the program it was not possible to add more children. But the key monitoring part of the program has remained the conditionality. Initial take up of the program was very close to 100% suggesting that this program was both understood and in demand.

Expansion – Because of its success, *Progresa* has expanded widely since 1997.[5] Indeed, the program started with 219,944 beneficiary households in that year and counted more than 6 million beneficiaries as of 2010. Which localities would be treated first was determined using a marginality index. This index, using information from the 1990 census, corresponds, for individuals above 15, to the normalized first principal component of: the share of illiterates, without secondary school education and earning less than twice the minimum wage, the share with no access to toilets, electricity or water supply, the share with house crowding (according to the number of rooms per person) and dirt floor, as well as the share living in communities with less than 5,000 inhabitants. Table A.2, reports the summary statistics of these components overall for Mexico.

[4] A detailed breakdown of the grants as of 2003 can be found in table A.1.

[5] In 2002, the program was rename *Oportunidades*.

Importantly for the identification strategy used in this paper, the expansion of the program followed three main waves. Figure 1 displays the number of households enrolled yearly in the program at the national level. These waves appear as a result of the Mexican anti-vote buying law that forbids the extension of such program during election years. Thus, the first wave ranges from 1997 to 1999. Its expansion was stopped by the 2000 Presidential election. The second wave began in 2000 and was stopped in 2005 because of the 2006 Presidential election. Finally, the last wave resumed after the 2009 midterm elections.[6]

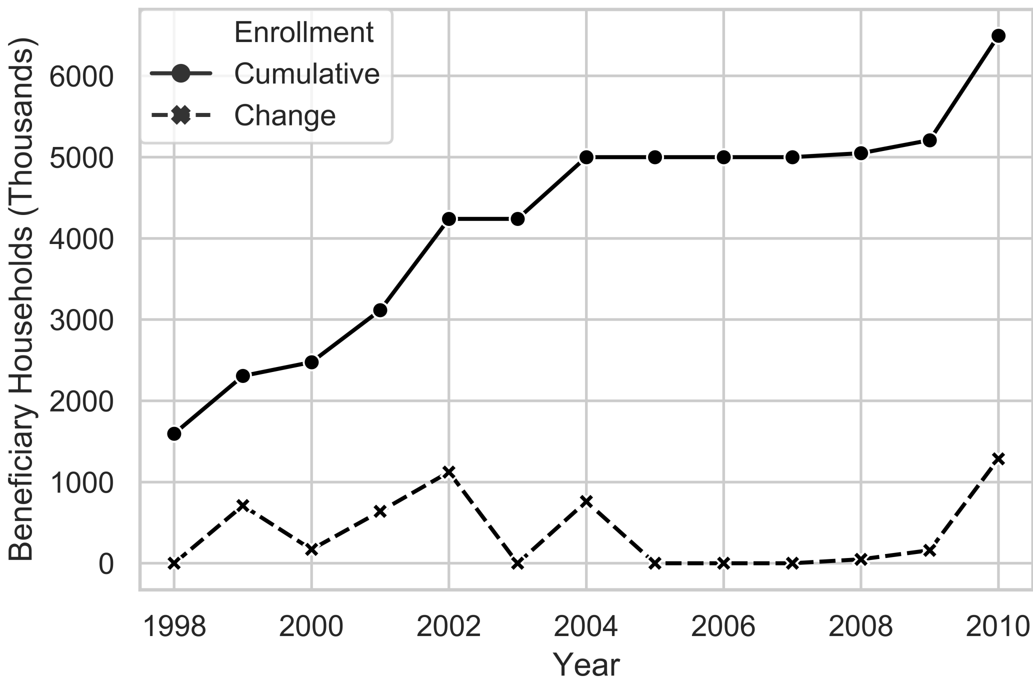

Figure 1: Treatment Expansion Over Time

This figure displays the evolution of the cumulative sum (solid line) and yearly changes (dashed line) in the number of beneficiary households in Mexico. The pattern of this roll-out shows three waves: 1997–1999, 2000–2005 and 2010 onward. The pauses are explained by the national elections of 2000, 2003, 2006 and 2009.

Source: *Progresa/Oportunidades* external examiner.

The first wave ending in 1999 targeted the poorest rural areas. The second ending in 2005, followed the same pattern though it extended to relatively richer areas. Finally, the third wave, in 2010, extended to both richer and more urban areas. Table A.2 presents

[6]Presidential elections occurred in 2000, 2006, and 2012 while midterms occurred in 1997, 2003, and 2009.

the correlation between the marginality index and its components with the changes in the enrollment proportion by municipalities.[7] The results presented in this table suggest that, while richer, the municipalities that saw the largest expansion during the second wave were closer to those of the first wave in terms of poverty profile than the municipalities in the third expansion. Thus, municipalities that benefited most from the first and second waves are presumably more alike.

This pattern can also be seen in Figure 2. This figure plots the count of enrolled households in 1999 against 2005 (left) and the count of enrolled households in 2005 against 2010 (right). The size and color of the points are representative of the total number of households in a municipality. We can see from the left panel that while a fraction of the new enrollments went to municipalities with more than 150,000 households, the second wave mainly targeted small municipalities. As can be seen from the right panel, this is no longer true with the third wave. Indeed, the expansion principally occurred in larger municipalities that had relatively low levels of enrollment in 2005.

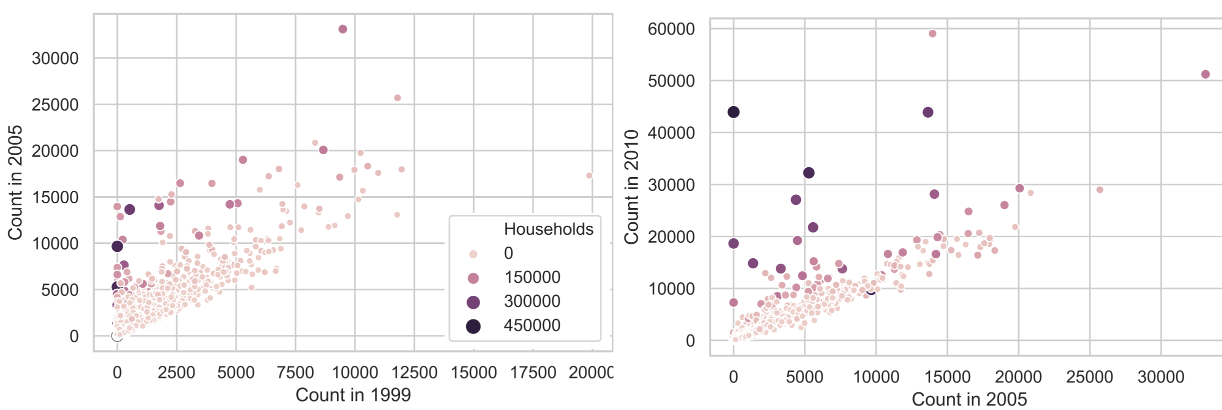

Figure 2: Program Expansion Towards More Urban Areas
This figure plots the count of households treated in 1999 against the number of households treated in 2005 (left) and the count of households treated in 2005 against the number of households treated in 2010 (right). The color and size of the points represent the total number of households in that municipality. This figure shows that the first two waves were mainly targeted towards similarly small municipalities while the third expansion disproportionately favored larger municipalities.
Source: INEGI and *Progresa/Oportunidades* external examiner.

[7] As will be described in more details in Section 3.1, the enrollment proportion is a measure of the cumulative number of households enrolled in the program in a given year over the total number of households in that municipality.

Younger and Older cohorts – Taking a long term perspective on the program suggests the definition of two treated cohorts that are treated differently. Consider the cohort of women aged 16 or higher when the program started in 1997. Provided that they were in a treated area, they had the opportunity to enroll in the program and receive a cash transfer proportional to the number of children they had at the time, conditional on respecting the school and health requirements for their children. This cohort of women, labeled the *older cohort* throughout this paper, was thus treated on two fronts: they received cash transfers and were required to invest in the human capital of their children through school attendance.[8] I define this *older cohort* to be composed of women aged 16 to 36 in 1997. A rationale for the definition of the upper bound is that women aged 36 when the program started will be 49 in 2010 when I observe their fertility. As can be seen in Figure A.1, 49 is the last age where the observed probability for a woman to have a child is non zero. In other words, 36 is the latest age for which fertility (total number of children ever born) can still meaningfully change in 2010.

Now consider the cohort of women aged 11 or lower in 1997. Unlike the *older cohort*, by design of the program, this group of women were more likely to be enrolled in school because of the program when it started.[9] Yet, after these school years, they faced the same program eligibility and conditionality as the *older cohort*. In other words, this group, labelled the *younger cohort*, first received extra schooling and then had children for which they would potentially receive cash transfers from the government under the condition that they go to school, the same way the *older cohort* did. Therefore, to the extent that these three elements have a potential effect on fertility, the *older cohort* can be used to capture the joint effect of the handout and the mandatory investment in their children's education on fertility. The *younger cohort* adds to this list the effect for a woman of being more educated.

I define this *younger cohort* to be composed of women aged 7 to 11 in 1997. In some

[8]I will also consider effects child health and mortality in Table 4. For now I abstract from this dimension to simplify exposition.

[9]This effect has been documented multiple times, see for instance: (Schultz, 2004)

specifications, I extend the range to include women aged 7 to 15. This is because the eligibility for entering the program as a child was 15 or lower. Still, these women aged 12 to 15 were at best partially treated by the program. A rationale for the definition of the lower bound is that women aged 7 when the program started will be 20 in 2010 when I observe their fertility which, as can be seen in Figure A.2, is still very early in the fertility profile of a Mexican woman. In my empirical analysis, I also report results for a *teenagers* category composed of women aged 2 to 6 in 1997 whose fertility I observe when they are 15 to 19 in 2010. A summary of these age thresholds is displayed in Figure 3.

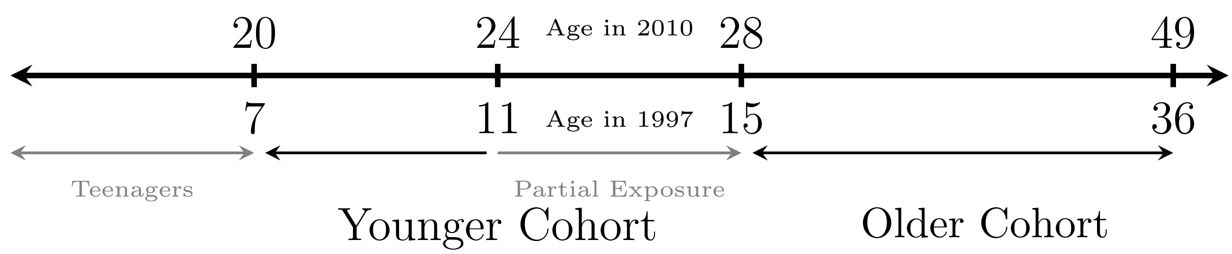

Figure 3: Cohorts Definitions

This figure summarizes the definitions of the *younger* and *older cohorts* depending on their age when the program started in 1997 and when they are observed in the IPUMS data in 2010. The *older cohort* is composed of women aged 16 to 36 in 1997. The *younger cohort*, in its narrow definition, is composed of women age 7 to 11. These women correspond to the first cohort who benefited from the additional schooling component of the program for the maximum number of years possible. A wider definition of the *younger cohort* also includes women aged 12 to 15. These women were partially exposed in the sense that they were eligible to receive additional schooling but not for the full length of their education. The teenagers cohort is composed of women 6 or younger in 1997 whose fertility is observed when they are still teenagers in 2010.

2.2 Theoretical Motivation

This section details how the three main treatment components (cash transfer, conditionality and schooling for the *younger cohort)* are expected to influence fertility from a theoretical standpoint. The effect of *Progresa* on fertility is not *ex-ante* self-evident. First, although there is a large literature documenting the negative correlation between income and fertility, the program may instead create an incentive for parents to bear more children since the program is an increasing function of the number of children. Second, it is not obvious whether the two different treatments received by the *older* and the *younger cohorts* (defined in the previous section) should result in different fertility changes. If the *younger cohort* reduces fertility further compared to the *older cohort*, this would be in line with *Progresa*'s ambition to alleviate poverty in the short run via cash transfers and to eradicate poverty in the long run thanks to its mandatory schooling component.

In order to frame the research question and capture both the static effect of cash transfer and mandatory schooling and the inter-generational dynamic coming from women's own education, I use an overlapping-generations model developed in Moav (2004). I first present the mechanisms by which fertility may be reduced in this *quality-quantity* trade-off model in the absence of treatment. I then show successively i) how fertility may be impacted by the cash transfer alone, ii) the effect of the introduction of conditionality which may reverse the first prediction and iii) the conditions under which the *younger cohort* would further decrease its fertility compared with the *older cohort*.

2.2.1 Benchmark model - no *Progresa*

Consider the following overlapping generations model inspired by Moav (2004) in which a representative parent lives two periods. In the first period, she acquires human capital. In the second period, she consumes, decides the number of children to have and decides whether and how much to (uniformly) invest in her children's education. She maximizes the log-linear utility function[10] over her own consumption (c_t), her number of children (n_t) and her children's future income (h_{t+1}).

$$U_t(c_t, n_t, h_{t+1}) = (1-\beta)\log(c_t) + \beta[\log(n_t) + \theta\log(h_{t+1})] \tag{1}$$

Here $\beta \in (0,1)$ is the weight determining the parent's preference for consumption over the number and income of her children, while $\theta \in (0,1)$ drives the preferences between quantity

[10]Many of the results derived in this model do not demand log-linear utility. Yet, as noted by Jones et al. (2008), the utility function does need to be separable in consumption, quantity and quality of children for the negative relationship between income and fertility to hold across generations.

and quality captured by the child's income, h_{t+1}, in the next period. Wages are normalized to 1 in this economy so that income only depends on human capital h.

Let human capital production be a function of the parent's investment in her children,[11] so that :

$$h_{t+1} = h(e_{t+1}) := \delta_0 + \delta_1 e_{t+1} \tag{2}$$

Here $\delta_0 > 0$ is the human capital endowment which is independent of investment in quality and $\delta_1 > 0$ is the return to investment on children e_{t+1}.

The representative parent is endowed with a unit of time and a constant fraction τ is needed to rise each child n_t regardless of their quality.[12] The cost of quantity is therefore given by τh_t due to forgone income and increases with the parent's human capital h_t while the cost of quality is given by e_{t+1} and does not depend on h_t. The budget constraint is given by:

$$c_t + n_t(\tau h_t + e_{t+1}) \leq h_t \tag{3}$$

Combining equations (1),(2) and (3) gives the optimization problem of the representative parent. First, note that the optimal consumption of a parent is a fixed fraction of the parent's income:

$$c_t^* = (1 - \beta)h(e_t) \tag{4}$$

[11]This is a common functional form first given by Becker and Tomes (1976). Although Moav (2004) does not define the human capital functional form, the author requires for the elasticity $\frac{e_{t+1}h'(e_{t+1})}{h(e_{t+1})}$ to be increasing in e_{t+1}. $h_{t+1} := \delta_0 + \delta_1 e_{t+1}$ represents a special case also used by Vogl (2015). This functional form allows both for fertility to decline in parental skills in an interior solution, and for a corner solution with no investment in quality if parental skills are too low. Vogl (2015) also includes a subsistence level of consumption which implies that the demand for children is first increasing in h_t for low levels of h_t before decreasing in this argument. This stylized fact does not seem to hold in the Mexican context. Indeed, childlessness or more broadly a small number of children for low levels of income do not seem to be important. This will be apparent in the empirical part below. In particular, I find that there is no extensive margin response to *Progresa* which would confirm this stylized fact.

[12]Another assumption is that $\beta > \tau$. This is because a corner solution of no investment in quality e_{t+1} implies a constant and positive number of children given by β/τ.

This depends on the investment made by her parent in her education e_t in the previous period. In other worlds, consumption is increasing in fixed proportion, across generations, provided that education is increasing as well.

The non-linear interaction between n_t and e_{t+1} in the budget constraint generates the usual quality-quantity trade-off first introduced by Becker (1960). In this model, the shadow price of the number of children is increasing with the representative parent's level of capital h_t and the investment in the quality of her children e_{t+1}. The marginal rate of substitution between quality and quantity for an interior solution is thus given by:

$$\frac{h(e_{t+1})}{\theta n_t} = \frac{\tau h(e_t) + e_{t+1}}{n_t/h'(e_{t+1})} \tag{5}$$

We can see that the cost of quantity $[\,h(e_t)+e_{t+1}]$ is increasing in e_{t+1} but also in the parent's own education e_t. The dynamic implications are that the price of quantity increases across generations if education is increased over time.

Finally, the optimal number of children as a function of investment decision includes a possible corner solution with no investment and is given by:[13]

$$n_t(e_{t+1}) = \begin{cases} \beta/\tau & \text{, for } e_{t+1} = 0 \\ \beta h(e_t)/(\tau h(e_t) + e_{t+1}) & \text{, for } e_{t+1} > 0 \end{cases} \tag{6}$$

2.2.2 Fertility decisions with *Progresa*

I now introduce, in turn, two of the main components of *Progresa*: the cash transfer and mandatory schooling. Results derived here are meant to represent *Progresa*'s treatment effect in a context where the circumstances of the baseline model from Section 2.2.1 represent the counterfactual. As noted in Parker and Todd (2017) take up of treatment were very high so

[13] The closed form solution for the optimal investment in education is given by: $e^*_{t+1} = \frac{\theta}{1-\theta}(\tau h_t - \frac{\delta_0}{\delta_1})$, which, for an interior solution, yields the following optimal number of children: $n^*_t = \frac{(1-\theta)\beta h_t}{\tau h_t - \delta_0/\delta_1}$.

I do not model decisions to enroll in the program, nor do I model when to exit it.

Unconditional cash transfer – First consider the effect of the introduction of a cash transfer without mandatory schooling. Let γ represent the average per child cash transfer.[14] For simplicity, the model abstracts from the fact that transfers were capped to a maximum of three children. The new budget constraint is now given by:

$$c_t + n_t(\tau h_t + e_{t+1}) \leq h_t + \gamma n_t \tag{7}$$

Equations (1),(2) and (7) now form the optimization problem. Note first that optimal consumption is unchanged compared to the benchmark model (4) which means that the relaxation of the budget constraint offered by the cash transfer only influences spending on quantity and/or quality.[15] The optimal fertility decision from (6) is now given by:

$$n_t(e_{t+1}) = \begin{cases} \beta h(e_t)/(\tau h(e_t) - \gamma) & \text{, for } e_{t+1} = 0 \\ \beta h(e_t)/(\tau h(e_t) + e_{t+1} - \gamma) & \text{, for } e_{t+1} > 0 \end{cases} \tag{8}$$

In the absence of investment in the quality of children ($e_{t+1} = 0$), an unconditional cash transfer will increase the number of children and the magnitude of this effect depends on the distance between the parent's income $h(e_t)$ and the transfer amount γ.[16] It is easy to see that the positive impact of the cash transfer on fertility is dampened when the parent does not opt for a corner solution and invests in quality ($e_{t+1} > 0$).

Conditional cash transfer – Now consider the effect of the cash transfer when coupled with mandatory schooling. Define the new investment in quality $\tilde{e}_{t+1}$. I can assume that $\tilde{e}_{t+1} >$

[14]Transfers amount are increasing with age and larger for female students in practice but I abstract from these dimensions here.

[15]This is a property of the model used. Because γ multiplies n_t it follows that $h_t - c_t = n_t(\tau h_t + e_{t+1} - \gamma)$. If γ does not depend on n_t, the optimal consumption is given by $c_t^* = (1-\beta)(h_t + \gamma)$.

[16]The special case $\gamma = 0$ yields the original corner solution in (6)

0 because investment in education is mandatory under the program and cannot be zero. $\tilde{e}_{t+1} > e_{t+1}$ implies that the parent does not reduce her own investment compared to a situation without the program. Thus, I can rewrite (8) dropping the corner solution as:

$$n_t(\tilde{e}_{t+1}) = \frac{\beta h(e_t)}{\tau h(e_t) + \tilde{e}_{t+1} - \gamma} \tag{9}$$

The relationship between $\tilde{e}_{t+1}$ and e_{t+1} (investment in the absence of the program) depends in part on the response of the parent; e.g. whether mandatory schooling and the parent's own investment are complements instead of substitutes. Equation (9) suggests that, at the margin, fertility will be reduced for all treated cohorts if the total induced investment in quality, $\tilde{e}_{t+1} - e_{t+1}$, exceeds the benefit from the transfer γ such that:

$$\tilde{e}_{t+1} - \gamma > e_{t+1}$$

Younger vs older cohort – Finally, there is a possibility that, out of the two treated groups, the *younger cohort* will experience an even larger reduction in fertility.[17] To see this let the superscripts $\{Y, O\}$ represent respectively the *younger* and *older cohort* and recall from equation (5) that the cost of quantity is given by: $[\,h(e_t) + e_{t+1}]$. Therefore, this cost of quantity will increase for the younger cohort if the following inequality is respected:

$$h(\tilde{e}_t^Y) + \tilde{e}_{t+1}^Y > h(e_t^O) + \tilde{e}_{t+1}^O$$

Which, assuming that the education investment under mandatory schooling is the same for both cohorts ($\tilde{e}_{t+1}^Y = \tilde{e}_{t+1}^O$), yields the following inequality:

$$h(\tilde{e}_t^Y) > h(e_t^O) \iff \delta_1 \tilde{e}_t^Y > \delta_1 e_t^O$$

[17]Remember that the *younger cohort* first receives extra schooling as a result of exposure to the program and then make fertility decisions otherwise facing the same conditions as their parents.

Given the investment in her education, the model predicts that an individual in the *younger cohort* will further decrease her fertility and that the increase in the cost of quantity is driven by $\delta_1 > 0$. Thus, overall, the program is expected to reduce the fertility of both cohorts as a result of the increase in education investment, provided that this boost to investment is large enough compared to the amount of the transfer. Furthermore, the *younger cohort* is expected to further reduce fertility as a consequence of a higher education level which increases income and the cost of quantity.

3 Empirical Strategy and Data

Identifying the potential effect *Progresa* has had on the *younger* and *older cohorts* presents a double challenge: i) the difference of treatment intensities between treatment and control need to be large enough for an effect to actually be detected, ii) both treatment exposure and fertility levels are jointly determined by age. This section motivates the difference-in-difference strategy used in this paper, which relies on cross-census variation to identify pre and post treatment cohorts, and the characteristics of the program's roll-out to determine areas treated and account for pretrends. Section 3.1 describes the construction of the intensive margin treatment variable based on the different waves characterizing the expansion of the program. It also describes the method used to account for pre-trends. Section 3.2 uses the timing of *Progresa*'s implementation to define the two treated cohorts and a control cohort in the IPUMS data.

3.1 Treatment Exposure

I identify the effect of *Progresa* on fertility using a difference-in-difference strategy inspired by Parker and Vogl (2018). This strategy relies on two measures of treatment intensity, interacted with a post treatment indicator, varying at the municipality level. The first measure represents the share of households enrolled in the program in 1999 and the second represents the analogous share for the year 2005. This yields a specification in which the coefficient on the interaction term for the 1999 enrollment measure captures an effect conditional on enrollment levels in 2005. In other words, conditional on having a high level of enrollment in 2005, the impact of a high level of enrollment in 1999 is given by the coefficient on the first interaction term.[18] This strategy retains variation in the timing and intensity of the treatment by municipality while also accounting for heterogeneity between municipalities. Specifically, the second interaction term aims at capturing the time varying effect of being in a poor municipality, allowing for the first interaction term to measure the effect of earlier high intensity enrollment.[19] The identification relies on the assumption that earlier treatment is not correlated with relevant unobservables.[20]

These two exposure measures are defined by the following ratios:

$$enroll_m^y = \frac{\text{Treated Households}_m^y}{\text{Total Number of Households}_m^y} \text{ , for } y \in 1999, 2005 \tag{10}$$

Enrollment in municipality m in year y is defined as the total of the cumulative sum of treated households at the end of that year divided by the number of households in the same location in that year.[21] I obtain geostatistical information from the Mexican external evaluation organization[22] which provides yearly information on the cumulative number of beneficiary

[18]Recall from Section 2.1 that municipalities relatively more intensively treated during the second wave of the program's expansion were closer, as defined by marginality indicators, to municipalities treated during the first wave.

[19]See equation (11) in Section 3.2 for more details.

[20]Note that looking at the long term effect of *Progresa* one could instead use the initial randomization of the program. The effect captured would thus be the effect of being exposed to the program for up to 18 months ahead of the control group. While this may be enough to capture differences in market outcomes such as education or income it may not be enough for fertility. To see this, consider a women who is 20 when the program starts in 1997. When her fertility is observed in 2010, she will have been exposed to the program for 13 years while a woman of the same age in the control group will have been exposed for 11.5–12 years. Therefore, by 2010, both treatment and control will have made fertility decisions facing very close conditions thus making the detection of an effect unlikely.

[21]Although this ratio is not bounded above by 1 (because over time, cumulatively, there can be more beneficiary households than the total number of households) allowing the denominator to vary over time helps with the interpretability of the results.

[22]The website is no longer in maintained but the author can share the data used upon request. In addition, the numbers used here are close to those found in Parker and Vogl (2018).

households by municipality. For the total number of households in a municipality, I use the household count provided by the Mexican national institute of statistics INEGI.[23]

Table 1 presents the summary statistics for the municipalities included in my analysis. The average municipality has an enrollment ratio of 0.26 in 1999 and 0.48 in 2005 reflecting the large increase in beneficiaries between the first and the second waves. Notice that while the bottom 75% of municipalities have exposure measures below .70 in 2005, this variable is not bounded above by 1. This is because the number of treated households corresponds to both current and formerly enrolled households.[24]

	1999 exposure ratio	2005 exposure ratio
mean	0.26	0.48
std	0.24	0.28
min	0.00	0.00
25%	0.02	0.23
50%	0.21	0.49
75%	0.46	0.70
max	1.05	1.26

Table 1: Exposure Measures

This table displays summary statistics of the 1999 and 2005 enrollment share measures defined in equation (10) across Mexican municipalities. Between 1999 and 2005, the average Mexican municipality ratio went from 0.26 to 0.48. Measures above 1 are indicative of a cumulative number of enrolled households larger than the total number of households in that municipality.
Source: *Progresa*/*Oportunidades* external examiner and INEGI.

3.2 Cross-Census Variation

I obtain information on fertility and other socio-economic features such as schooling and income, using the Integrated Public Use Micro Sample (IPUMS) census surveys for Mexico

[23] In their paper, Parker and Vogl (2018) use the the household count provided by the IPUMS census survey. An issue with this method is that households belonging to small municipalities are more likely to be under-counted as noted in the IPUMS documentation.

[24] This is the case for 3 municipalities out of 2,392. The definition of a household in *Progresa* may also be looser than the statistical definition used by INEGI. Therefore it is possible for *Progresa* to report more households treated than the total number of statistical households even early on. To the extent this would represent measurement error I replicated the main regression having top-coded these ratios. Results remain quantitatively similar.

for the years 2010, 2000 and 1990.[25] The main dependent variable used throughout this study is the number of children ever born. This corresponds to the total number of children a woman ever gave birth to. Because it includes both surviving and dead children the effects found using this variable should be interpreted as the change in the total number of births.[26]

The use of this dependent variable implies that some of the fertility decision can potentially be made prior to the beginning of the treatment and are therefore unrelated to it. Depending on when the treatment occurs with respect to one's fertility timing, a woman can either stop having additional children or plan to have fewer children compared to what she would have had in the absence of the program. It is only by the end of a woman's fertility window that one can observe the total chosen number of children. In this study, complete fertility will be observed only for some women belonging to the *older cohort*. It is unlikely to be observed for the *younger cohort* since the oldest are 28 in 2010. Therefore, an important assumption of this work is that reductions in fertility can be at least partially observed before a woman's fertility window closes. In other words, the decision to reduce fertility due to the treatment can be observed only if, on average within a cohort, there is reductions of the number of children throughout the fertility window.

In order to measure the total effect of the treatment, I follow the methodology introduced by Aaronson et al. (2014) and use cross-census variation between the 2010 and the 1990 census surveys instead of within 2010 census variation across age groups. For the *older cohort*, I include women, from the 2010 and 1990 IPUMS, aged 29 to 49 at the time of the census. For the *younger cohort*, I include women, from the 2010 and 1990 IPUMS, aged 20 to 24 (or 28) at the time of the census. Thus, in each case, women appearing in the 2010 census belongs to the post-treatment category while women appearing in the 1990 census

[25] I use the 10% sample from 1990, the 10.6% sample from 2000 and the 10% sample from 2010.

[26] The IPUMS censuses also include those two variables although the information on the number of dead children is absent from the 1990 census.

belongs to the pre-treatment category.[27]

I merge this individual level information with information on the treatment exposure intensity, detailed in Section 3.1. This measure, which enters the interaction term of interest, varies at the municipality level. When necessary these municipalities are merged into a single, time-invariant, municipality.

Thus, the main identification strategy takes advantage of the repeated cross Section offered by the different decennial censuses and applies the difference-in-difference strategy inspired by Parker and Vogl (2018). Combined, these two components yield the following empirical specification:

$$y_{imt} = \beta(enroll_m^{1999} \times post_t) + \gamma(enroll_m^{2005} \times post_t) + \Psi(X_m^{1990} \times post_t) + \delta_m + \eta_t + \epsilon_{imt} \tag{11}$$

This model relates the total number of children y_{imt} for woman i, in municipality m, aged t, to a municipality fixed effect δ_m, a year of census fixed effect η_t and the interaction of the municipal enrollment share in 1999 and an indicator equals to 1 if the observation belongs to the 2010 census, as well as the interaction between the post indicator and 2005 enrollment. I also include an interaction term between the post indicator and the marginality index described in Section 2.1. This strategy retains the timing and intensity variations in the treatment captured by β while accounting for heterogeneity between municipalities in the data set. Thus, γ and Ψ, jointly capture the time varying effect of being in a poor municipality. This allows β to measure the effect of earlier high intensity enrollment without this cofounder.[28]

[27] I will also report the results from an exercise in which I use the 2000 census as pre-treatment. The advantage of using the 2000 census is that it makes post and pre-treatments more comparable and the requirement on pretends less likely to be violated. Nevertheless, this does not constitute the main specification because the program started in 1997 which means that the 2000 census does not truly represent a pre-treatment group but a short term treatment instead.

[28] In appendix A.1, I present elements defending both the relevance of the research question and the validity of the cross-census specification.

4 Results

This section presents the results of estimating the empirical model described in Section 3. Section 4.1 displays the main results of the analysis and shows the extent to which fertility declined in areas treated by *Progresa*. Section 4.2, tries to isolate the effect of schooling for the *younger cohort* by making comparisons to the *older cohort*. Finally, Section 4.3 attempts to shed light on the mechanisms by which fertility was reduced.

4.1 Main Results

Table 2 reports the results from estimating equation (11), for the different cohorts, using women of the same age in the 1990 census as comparison group. I also report the average fertility for each cohort-census combination – the difference between which is captured by the year dummy. Throughout, I report the estimate for the 2005 enrollment share interacted with the post indicator. Recall that, jointly with the marginality index, this interaction term should capture the time varying effect of belonging to a poor locality.

Column (1) presents the results for a teenager cohort composed of women aged 2 to 6 in 1997 when the program started and whose fertility I observe when they were 15 to 19 in the 2010 census. The coefficient on the interaction term between the 1999 enrollment share and the post indicator is negative but small and insignificant. This is unsurprising given the average fertility for Mexico in this age category, displayed at the top of the table for each cohort, which remained stable at low levels between 1990 and 2010. Columns (2)–(4) present the results for different definitions of the *younger cohort*. The negative effect of the program on fertility is not significant in column (2) for the fully treated cohort (aged 7 to 11 in 1997). When considering the partially treated, either in isolation in column (3) or in combination with the fully treated cohort in column (4) this negative effect becomes significant. Thus, for the extended definition of the *younger cohort* presented in column (4) a 1 percentage point increase in the enrollment share in 1999 is associated with a fertility decline of 0.002 fewer

	Teenagers	*Younger Cohort*			*Older Cohort*	*All Cohorts*	*Placebo Cohorts*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age in 1997	2-6	7-11	11-15	7-15	16-36	7-36	40-47
	-	-	-	-	-	-	-
Age in 2010	15-19	20-24	20-28	15-28	29-49	20-49	53-60
Fertility							
2010 Mean of Cohort	0.17	0.87	1.65	1.19	3.14	2.44	5.13
2010 SD of Cohort	(0.45)	(1.06)	(1.42)	(1.28)	(2.22)	(2.15)	(3.22)
1990 Mean of Cohort	0.17	1.03	2.09	1.46	4.25	3.08	6.45
1990 SD of Cohort	(0.53)	(1.28)	(1.73)	(1.57)	(2.95)	(2.82)	(4.02)
Enrollment 1999	-0.0124	-0.0904	-0.283***	-0.198***	-0.377***	-0.308***	-0.291
×*Post*	(0.0215)	(0.059)	(0.0914)	(0.0695)	(0.128)	(0.0911)	(0.210)
Enrollment 2005	-0.112***	-0.439***	-0.436***	-0.331***	-0.453***	-0.455***	0.937***
×*Post*	(0.0242)	(0.0668)	(0.0787)	(0.0542)	(0.141)	(0.102)	(0.234)
Marginality	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	961,253	834,974	1,422,252	2,383,505	2,294,703	3,716,955	484,997
R^2	0.017	0.055	0.063	0.032	0.140	0.071	0.123

All regressions are estimated using ordinary least squares. Standard errors clustered at the municipality level are in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Effects on Fertility
This table displays estimated effects of being exposed to *Progresa* on fertility for different cohorts. The dependent variable is the total number of children ever born and varies at the individual-census level. The effect of the treatment is captured by the coefficient on the interaction of a 2010 census dummy with the 'Enrollment 1999' share, which varies at the municipality level and corresponds to the cumulative sum of households enrolled in the program divided by the total number of households in that municipality. The variable 'Enrollment 2005' is the equivalent for the year 2005. The variable 'Marginality' is the index that determined the location of the initial treatment. Combined, these two variables aim to capture the time varying effect of belonging to a poor municipality. 'Municipality' represents fixed effects for the 2,392 municipalities. The 'Year' fixed effect corresponds to a dummy equal to 1 if an observation belongs to the 2010 census and 0 if it belongs to the 1990 census.

children. For the average Mexican municipality this corresponds to a decline in fertility by 0.052 children per woman.[29] This number seems sensible compared with the average 0.33 decrease in children in Mexico for this age group over the same period.

Column (5) reports the change in fertility experienced by the *older cohort*, whose fertility I observe when they were aged 29 to 49 in the 2010 census. The coefficient on the 1999 enrollment share suggests a reduction in fertility by 0.004 fewer children for each 1 percentage point increase in the 1999 enrollment ratio. This corresponds to a decline of 0.099 children per woman, for the average Mexican municipality, compared with a decline in fertility by

[29] As can be seein in Table 1, the mean of the enrollment measure in 1999 is 0.26.

1.21 children that this cohort experienced on average between 1990 and 2010. Column (6) pools the *younger* and the *older cohorts* together and confirms the individual results: the program did not act as an incentive to have more children and instead negatively affected women's fertility, a result compatible with the quality-quantity trade-off model developed in Section 2.2.

Finally, column (7) provides the results of a placebo regression including a cohort of older women. These are women aged 40 to 47 when the program started and whose total fertility I observe when they were 53 to 60 in 2010 the census. Because of their age at the time, these women are expected to have been at the end of their fertility window when the program started. Reassuringly, the effect for this cohort of woman is no longer significant.[30] In Table A.3, I reproduce the same regressions using the 2000 census as pre-treatment.[31] The main conclusions remain unchanged.

4.2 Additional Effect of Schooling

Given that the program is estimated to have impacted women's fertility negatively, including for the *younger cohort*, I now ask whether we can detect an extra effect of schooling on fertility. Recall from Section 2.2, that the *younger cohort* is expected to further decrease its fertility in comparison with the *older cohort* as a result of their increased education level which translates into a greater cost of quantity. The first cohort who received a full dose of extra education because of the program were women aged 7 to 11 in 1997. The youngest cohort that did not receive the education treatment in 1997 were aged 16 to 20 in 1997.

Because these cohorts are close in age, I first display results of estimating a model using within census comparison under the assumption that the key issue with such a model – the fact that age jointly determines treatment exposure and fertility – is mitigated by the age

[30]Note that from a biological standpoint these women could still modify marginally their fertility. The issue with including older women as placebo would then be age expectancy related. To the extent that the poorest women are both more likely to bear more children and to die younger this would constitute a wrong placebo group.

[31]This, of course, is not a true pre-treatment since *Progresa* started in 1997.

proximity of the two cohorts.[32] I also present results from a triple-difference strategy, similar to the main model presented in equation (11, in which the comparison between the younger and older cohorts constitutes the third difference. Both methods attempt to isolate the effect of having received the education treatment.

Table 3 presents the results from both strategies. All regressions include controls for the interaction measure of the 2005 enrollment share and marginality index with a post dummy. These are once more included in order to control for the time varying effect of belonging to a poor municipality. Panel A displays the results using the narrow definition of the *younger cohort* limited to the fully treated category of women aged 7 to 11 in 1997. They are compared to women from the *older cohort* aged 16 to 20 in 1997. These are the youngest women from the *older cohort* not to have been treated by the schooling component of the program and thus the closest comparison group available. Panel B displays results extending the definition of the *younger cohort* to partially treated women and includes women aged 7 to 15 in 1997. The comparison group is also extended to match the same number of age cohorts and includes women from the *older cohort* aged 16 to 24 in 1997.

Column (1) presents the results for the within census strategy. Using the narrow definition of the *younger cohort* composed of women aged 20 to 24 in 2010 yields a small negative and insignificant estimated additional effect of schooling on fertility, which appears in Panel A. The extended definition of the *younger cohort*, composed of women aged 20 to 28 in 2010, appearing in Panel B, suggests a significant negative effect. This effect is not robust to the alternative specifications using the triple-difference strategy presented in column (2). Overall these results suggest no detectable additional effect on the fertility of the younger cohort. This could be explained by a return on investment of schooling[33] that is too low. Alternatively, I note that when observed in 2010 between 20 and 28, most of the women belonging to *younger cohort* are at a relatively early stage of their fertility window. Thus,

[32] See Appendix A.1 for a full discussion of this issue.

[33] δ_1 in Section 2.2.

	(1) Enrollment 1999 × Educated (Within 2010)	(2) Enrollment 1999 $\times Post \times Educated$ (2010 vs 1990)
Panel A: [7-11] vs [16-20] in 1997		
Interaction 1999	-0.0730	0.0820
	(0.0625)	(0.0617)
Interaction 2005	-0.229***	0.203***
	(0.0579)	(0.0443)
Double Interactions	No	Yes
Age	Yes	No
Marginality	Yes	Yes
Municipality	Yes	Yes
Observations	915,661	1,521,826
R^2	0.278	0.277
Panel B: [7-15] vs [16-24] in 1997		
Interaction 1999	-0.151***	-0.0147
	(0.0577)	(0.0530)
Interaction 2005	-0.229***	0.195***
	(0.0577)	(0.0394)
Double Interactions	No	Yes
Age	Yes	No
Marginality	Yes	Yes
Municipality	Yes	Yes
Observations	1,576,745	2,576,501
R^2	0.292	0.259

All regressions are estimated using ordinary least squares. Standard errors clustered at the municipality level are in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Extra Effect of Schooling

This table displays estimated effects of being exposed to the schooling component of *Progresa* on fertility. The dependent variable is the total number of children ever born. Panel A presents the results where the educated cohort was 7 to 11 while the control group was 16 to 20 in 1997. Panel B presents the results where the educated cohort was 7 to 15 while the control group was 16 to 24 in 1997. Results in column (1) use treatment and control groups from the 2010 census only. Results in column (2) use the 2010 and 1990 censuses. The effect of interest is captured by the variable 'Interaction 1999' which corresponds to the interaction of the 1999 enrollment share with an educated dummy variable equal to 1 if an individual belongs to the educated cohort in column (1). In column (2), 'Interaction 1999' captures the triple interaction of the 1999 Enrollment, with the educated dummy and a 'Post' dummy equal to 1 if an individual is observed in the 2010 census. The variable 'Interaction 2005' follows the same definition using the 2005 Enrollment ratio. The variable 'Marginality' is the index that determined the location of the initial treatment. Combined, these two variables aim to capture the time varying effect of belonging to a poor municipality. 'Municipality' represents fixed effects for the 2,392 municipalities. The 'Year' fixed effect corresponds to a dummy equal to 1 if an observation belongs to the 2010 census. The variable 'Double Interactions' corresponds to the set of double-interactions of the control variables and fixed effects for the triple-difference estimation. .

these results do not rule out an effect of schooling on fertility manifesting itself in later years.

4.3 Mechanisms

In this section, using the same cross-census specification as in equation (11), I test whether several of the assumptions and my interpretation of the results made so far are in line with what can be found in the data. Columns (1)–(4) of Table 4 present results for the narrow definition of the *younger cohort* (women aged 7 to 11 in 1997). Columns (5)–(8) present the results for the same dependent variables using the *older cohort*. All results displayed use the cross-census variation between 2010 and 1990 except for columns (4) and (8). Because the count of dead children is not available for in 1990, I instead use the 2000 census as pre-treatment.

	younger cohort				*older cohort*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Schooling	Extensive	Intensive	Children Dead	Schooling	Extensive	Intensive	Children Dead
Enrollment 1999	0.480*	-0.0002	-0.245**	-0.0194	0.0808	-0.0103	-0.336***	-0.0394
	(0.209)	(0.0172)	(0.0633)	(0.0108)	(0.191)	(0.00760)	(0.127)	(0.0271)
Enrollment 2005	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Marginality	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	865,015	878,085	439,467	489,470	2,317,792	2,337,470	2,076,353	2,402,181
R^2	0.233	0.034	0.067	0.017	0.282	0.012	0.175	0.059

All regressions are estimated using ordinary least squares.
Standard errors clustered at the municipality level are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Potential Mechanisms

This table displays estimated effects of being exposed to *Progresa* on range of outcome variables for difference cohorts. These dependent variables are 'Schooling' (number of years of education), 'Extensive' (dummy variable equal if a woman has any child), 'Intensive' (number of children conditional on having at least one child), 'Children dead' (total number of children dead). The effect of the treatment is captured by the 'Enrollment 1999' ratio which varies at the municipality level and corresponds to the cumulative sum of households enrolled in the program divided by the total number of households in that municipality. This variable is interacted with the 'Year' dummy variable which equals 1 if an observation belongs to the 2010 census. The variable 'Enrollment 2005' is the equivalent for the year 2005. The variable 'Marginality' corresponds to the index that determined the location of the initial treatment. Combined, these two variables aim to capture the time varying effect of belonging to a poor municipality. The variable 'Municipality' is a dummy variable for the time invariant 2,392 municipalities.

Column (1) presents the effect of the program on the number of years of education for the *younger cohort*. As documented in the literature and in Appendix A.1, *Progresa* positively impacted the education of the young cohort. As expected, I find no such effect of the program on the education of the *older cohort* as can be seen in column (6). These results support two key assumptions of this paper. First, this suggests that the *younger* and *older cohorts* were indeed treated differently by the program thus confirming that they represent two distinct cohorts. Second, columns (1) and (5) are suggestive evidence that the reduced form results obtained in this paper are consistent with the theory in Section 2.2. Specifically, that fertility may have been reduced in response to changes in investment in the education of children, which reduced the optimal number of children.

Columns (2), (3), (6) and (7) decompose the effect on fertility along the intensive and extensive margins. Column (2) reports a precisely estimated zero on the extensive margin for the *younger cohort*. The effect on the extensive margin of the older cohort in column (6) is also estimated to be zero. These results, in combination with the significant intensive margin results in columns (3) and (7), suggest that the effect on fertility was entirely driven by the intensive margin. This may be surprising for the *younger cohort* as one could have expected that spending more time at school might translate into a delay in the arrival of the first child. But these results are consistent with Figure A.3 which shows that by age 20 close to 50% of Mexican woman have had at least one child.

Finally columns (4) and (8) show the effect of the program on child mortality. The inclusion of these regressions is motivated by the fact that, if a woman does not maximize the total number of children ever born but the total number of children alive instead, then total fertility could also decline as a result of improvements in the health of children enrolled.[34] The estimates in columns (4) and (8) suggest a small negative and insignificant effect of the program on the number of deceased children. However, because the results use variation between the 2010 and 2000 censes, these point estimates may be upward biased, since the program had existed for 3 years by 2000.

[34]As discussed in Section 2.1, *Progresa*'s payments were also conditional on receiving some health care.

5 Conclusion

In this paper I have studied the impact that one of the world's oldest conditional cash transfer programs has had on fertility in the long run, in Mexico. I defined two cohorts of women that were treated differently by Mexico's *Progresa* program based on their age, and motivated theoretically why such a program may reduce each group's fertility via its impact on education investment, income and cost of quantity. The results of this paper suggest that, far from creating an incentive for women to have more children in order to receive more transfers, *Progresa* reduced the total number of children to a small but significant extent. Perhaps surprisingly, the increase in human capital encouraged by the program does not seem to have had any detectable differential impact on the fertility of the *younger cohort*. The results also offer suggestive evidence that the effect on fertility corresponds less to a change in the timing of births (first births don't seem to be postponed), but rather to women tending to have fewer children as they progress towards the end of their fertility window. If true, then I may observe the fertility of the *younger cohort* too early in their fertility cycle to capture the true additional effect of schooling.

References

Aaronson, D., Lange, F., and Mazumder, B. (2014). Fertility transitions along the extensive and intensive margins. *The American Economic Review*, 104(11):3701–3724.

Baird, S., Chirwa, E., McIntosh, C., and Özler, B. (2010). The short-term impacts of a schooling conditional cash transfer program on the sexual behavior of young women. *Health economics*, 19(S1):55–68.

Barber, S. L. (2009). Mexico's conditional cash transfer programme increases cesarean section rates among the rural poor. *European journal of public health*, 20(4):383–388.

Barham, T. (2011). A healthier start: the effect of conditional cash transfers on neonatal and infant mortality in rural mexico. *Journal of Development Economics*, 94(1):74–85.

Barham, T., Macours, K., and Maluccio, J. A. (2018). Experimental evidence of exposure to a conditional cash transfer during early teenage years: Young women's fertility and labor market outcomes.

Becker, G. S. (1960). An economic analysis of fertility. In *Demographic and economic change in developed countries*, pages 209–240. Columbia University Press.

Becker, G. S. and Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of political Economy*, 81(2, Part 2):S279–S288.

Becker, G. S. and Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of political Economy*, 84(4, Part 2):S143–S162.

Behrman, J. R., Parker, S. W., and Todd, P. E. (2009). Schooling impacts of conditional cash transfers on young children: Evidence from mexico. *Economic development and cultural change*, 57(3):439–477.

Coady, D. P. and Parker, S. W. (2004). Cost-effectiveness analysis of demand-and supply-side education interventions: the case of progresa in mexico. *Review of Development Economics*, 8(3):440–451.

Doepke, M. (2005). Child mortality and fertility decline: Does the barro-becker model fit the facts? *Journal of population Economics*, 18(2):337–366.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, 50(4):1051–1079.

Fernald, L. C., Gertler, P. J., and Neufeld, L. M. (2008). Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of mexico's oportunidades. *The Lancet*, 371(9615):828–837.

Fernald, L. C., Gertler, P. J., and Neufeld, L. M. (2009). 10-year effect of oportunidades, mexico's conditional cash transfer programme, on child growth, cognition, language, and behaviour: a longitudinal follow-up study. *The Lancet*, 374(9706):1997–2005.

Galor, O. (2012). The demographic transition: Causes and consequences. cliometrica, 6, 1-28.

Galor, O. and Weil, D. N. (1999). From malthusian stagnation to modern growth. *The American Economic Review*, 89(2):150–154.

Gertler, P. (2004). Do conditional cash transfers improve child health? evidence from progresa's control randomized experiment. *American economic review*, 94(2):336–341.

Jayachandran, S. and Lleras-Muney, A. (2009). Life expectancy and human capital investments: Evidence from maternal mortality declines. *The Quarterly Journal of Economics*, 124(1):349–397.

Jones, L. E., Schoonbroodt, A., and Tertilt, M. (2008). Fertility theories: can they explain the negative fertility-income relationship? Technical report, National Bureau of Economic Research.

Kugler, A. D. and Rojas, I. (2018). Do ccts improve employment and earnings in the very long-term? evidence from mexico. Technical report, National Bureau of Economic Research.

Laszlo, S., Farhan Majid, M., and Renée, L. (2019). Conditional cash transfers, women's empowerment and reproductive choices.

Moav, O. (2004). Cheap children and the persistence of poverty. *The economic journal*, 115(500):88–110.

Parker, S. W. and Todd, P. E. (2017). Conditional cash transfers: The case of progresa/oportunidades. *Journal of Economic Literature*, 55(3):866–915.

Parker, S. W. and Vogl, T. (2018). Do conditional cash transfers improve economic outcomes in the next generation? evidence from mexico. Technical report, National Bureau of Economic Research.

Potter, J. E., Schmertmann, C. P., and Cavenaghi, S. M. (2002). Fertility and development: Evidence from brazil. *Demography*, 39(4):739–761.

Rubio-Codina, M. (2010). Intra-household time allocation in rural mexico: Evidence from a randomized experiment. In *Child labor and the transition between school and work*, pages 219–257. Emerald Group Publishing Limited.

Schultz, T. P. (2004). School subsidies for the poor: evaluating the mexican progresa poverty program. *Journal of development Economics*, 74(1):199–250.

Skoufias, E., Parker, S. W., Behrman, J. R., and Pessino, C. (2001). Conditional cash transfers and their impact on child work and schooling: Evidence from the progresa program in mexico [with comments]. *Economia*, 2(1):45–96.

Stecklov, G., Winters, P., Todd, J., and Regalia, F. (2007). Unintended effects of poverty programmes on childbearing in less developed countries: experimental evidence from latin america. *Population Studies*, 61(2):125–140.

Todd, P. E. and Wolpin, K. I. (2006). Assessing the impact of a school subsidy program in mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American economic review*, 96(5):1384–1417.

Urquieta, J., Angeles, G., Mroz, T., Lamadrid-Figueroa, H., and Hernandez, B. (2009). Impact of oportunidades on skilled attendance at delivery in rural areas. *Economic Development and Cultural Change*, 57(3):539–558.

Vogl, T. S. (2015). Differential fertility, human capital, and development. *The Review of Economic Studies*, 83(1):365–401.

A Appendix

A.1 Support for the Identification Strategy

This Section aims to present various arguments in support of the framing of the research question and the specification used. Specifically, I provide suggestive evidence that the addition of the 2005 enrollment measure and marginality index suggested by Parker and Vogl (2018) is convincing in accounting for pre-trends. I also provide evidence that schooling indeed increased in targeted areas. Finally, I point out the challenges with using within census variation when looking at fertility.

In Section 2.2, I motivated the impact of *Progresa* on fertility by the changes it imposed on education investment by the *older cohort* for the *younger cohort*. In order to provide evidence that this investment in education is indeed occurring, I keep the methodology from Parker and Vogl (2018) in equation (11) but instead use within census comparisons using the data from the 2010 census:

$$\begin{aligned} y_{imj} = \alpha_j + \delta_m + \sum_{j=13}^{59} (enroll_m^{1999} \times d_j)\beta_j \\ + \sum_{j=13}^{59} (enroll_m^{2005} \times d_j)\gamma_j + \sum_{j=13}^{59} (X_m^{1990} \times d_j)\Psi_j + \epsilon_{imj} \end{aligned} \tag{A.1}$$

In this model, each coefficient β_j multiplies the interaction of the enrollment measure in 1999 with a specific age dummy d_j where the omitted category is 60 in 2010 (47 in 2000). These age dummies are also interacted with the 2005 enrollment ratio and the marginality index to account, once more, for the time varying effect of being in a poor municipality.

The orange line in Figure A.4 plots the estimates $\hat{\beta}_j$ and their associated 95% confidence intervals from the estimation of equation (A.1) in which the dependent variable is number of years of education. As expected if the identification strategy is successful, these estimates are close to zero and insignificant for women aged 29 to 59 in the 2010 census. Recall that women aged 16 or higher in 1997 were not eligible for the program and should therefore not have benefited from the program in terms of their own education. It can also be seen in Figure A.4 that the effect on schooling starts to go up for women between the age of 11 and 15 in 1997 and becomes positive and significant for women age 10 or younger in 1997.[35] These are the expected results.[36] The *younger cohort* seem to have indeed benefited from the program in terms of schooling while the *older cohort* did not. Also, the stability of the estimates before the 16 year old threshold is encouraging vis-à-vis the ability of the identification strategy to account for pre-trends.[37]

Figure A.4 also plots the $\hat{\beta}_j$ and their associated 95% confidence interval for the estimation of equation (A.1) when the dependent variable is instead the total number of children. This corresponds to the purple line. The results seem to confirm two key elements of the model presented in equation (11). First, they seem to indicate that the age 36 in 1997 is a relevant threshold for the definition of the *older cohort*. Indeed, past this age we can see that there is no significant difference with the omitted category which is in line with the fact that, by that age, the number of children is far less likely to change.

Second, the shape of the estimates as decreases indicates the issue with using a within census strategy for fertility, in which the pre-treatment group could for instance be defined by women aged 46 or higher in 1997. To see why, consider the simple case where exposure to the treatment is defined by a dummy variable $T = \{0, 1\}$ and let $J = o$ represent the omitted category composed of older women (in our example women aged 60 in 2010). Then by definition of the difference-in-difference estimator, β_j captures the following differences in conditional expectations:

$$\hat{\beta}_j = E\left[Y|T=1, J=j\right] - E\left[Y|T=0, J=j\right] - \left(E\left[Y|T=1, J=o\right] - E\left[Y|T=0, J=o\right]\right)$$

[35]Recall from Section 2.1 that the 11 to 15 in 1997 cohort corresponds to a partially treated cohort in terms of schooling.

[36]These results are also quantitatively similar to those found in Parker and Vogl (2018).

[37]Results from estimation of equation (A.1) without the inclusion of the 2005 enrollment ratio display a pattern in which schooling is also increasing for the *older cohort*.

Note that the fertility expectations conditional on treatment are the same as the unconditional expectation and converge towards zero when age tends to zero, so that:

$$\lim_{j \to 0} E\left[Y|J=j\right] = \lim_{j \to 0} E\left[Y|T=1, J=j\right] = \lim_{j \to 0} E\left[Y|T=0, J=j\right] = 0$$

This means that the within census estimation of a difference-in-difference estimator for fertility will tend towards the single difference represented by the omitted category:

$$\lim_{j \to 0} \hat{\beta}_j = 0 - \left(E\left[Y|T=1, J=o\right] - E\left[Y|T=0, J=o\right]\right)$$

A.2 Supplemental Figures

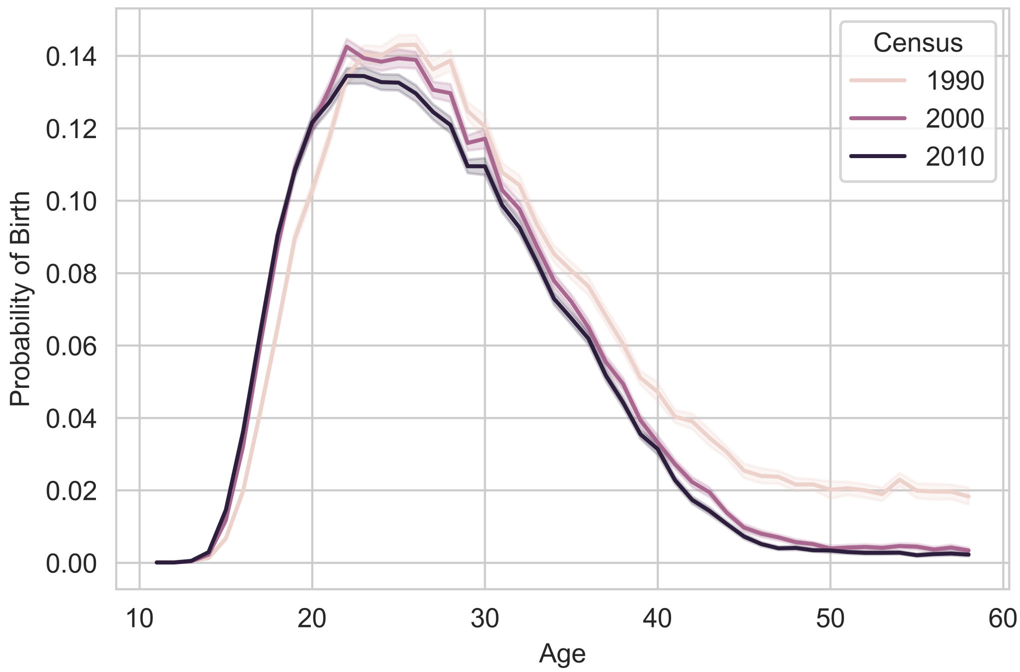

Figure A.1: Probability of Giving Birth by Age in Each Census

This figure displays the probability, by age, for a woman, between 12 and 60 to have a child aged 0. Each line represents the same exercise for a different census.
Source: IPUMS

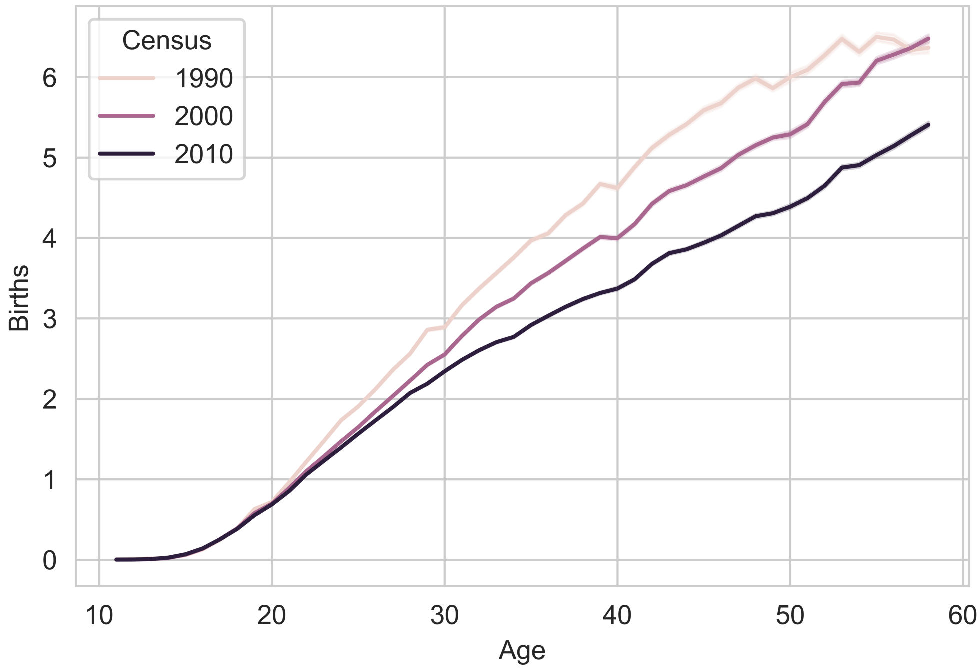

Figure A.2: Cumulative Fertility by Age and Census
This figure displays the cumulative number of children on average by age from 12 to 60. Each line represents the same exercise for a different census.
Source: IPUMS

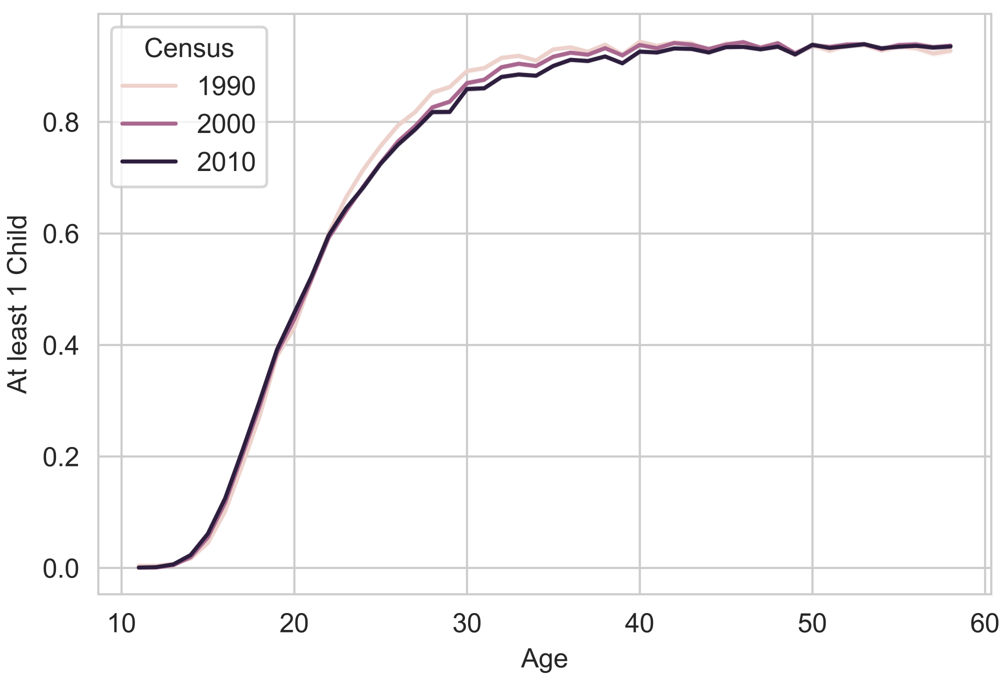

Figure A.3: Probability of First Child
This figure displays the probability of already having at least one child, on average by age, from 12 to 60. Each line represents the same exercise for a different census.
Source: IPUMS

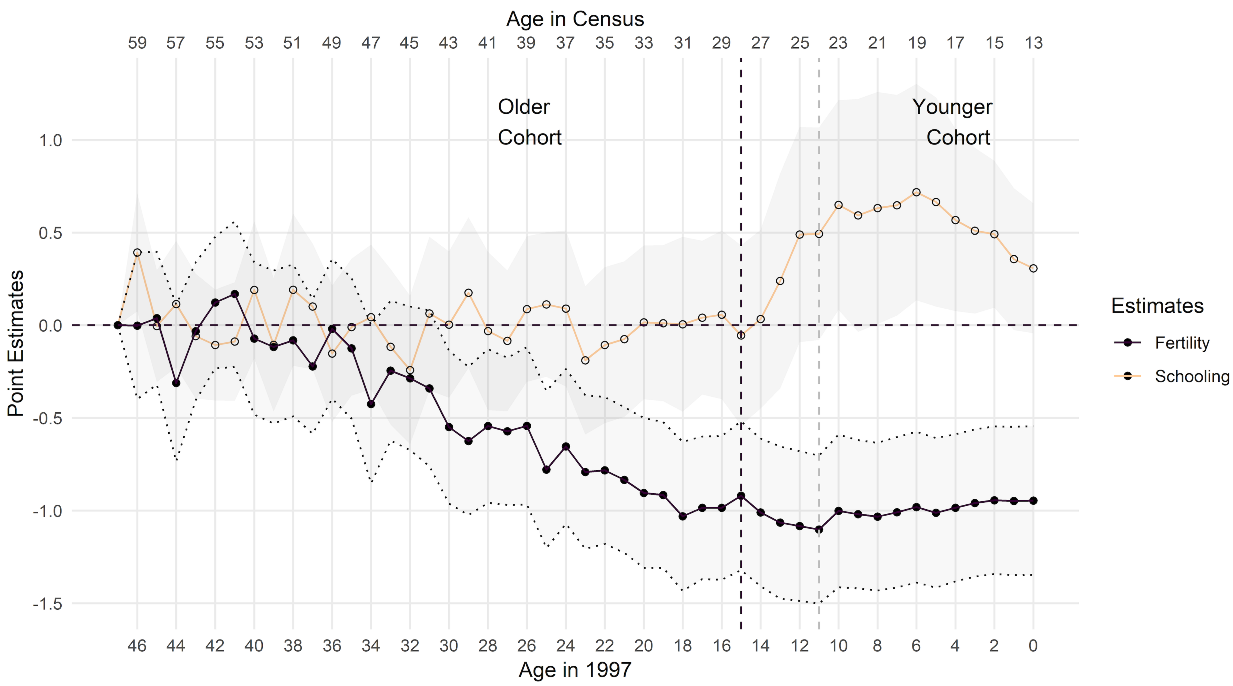

Figure A.4: Within Census Comparisons, Schooling and Fertility

This figure displays the coefficients on interactions between age cohort dummies and the 1999 enrollment intensity, from the estimation of the model detailed in equation (A.1), for the two dependent variables: years of education (purple) and fertility (orange). Data includes women aged 13 to 60 in the 2010 census. Women aged 60 constitute the omitted category. The 95% confidence intervals are calculated using of standard errors clustered at the municipality level.

A.3 Supplemental Tables

Cash Benefits of Progresa/Oportunidades: Monthly Pesos, 2003 Second Semester		
	Boys	Girls
Primary School		
Grade 3	105	105
Grade 4	120	120
Grade 5	155	155
Grade 6	210	210
Middle School		
Grade 7	305	320
Grade 8	320	355
Grade 9	335	390
High School		
Grade 10	510	585
Grade 11	545	625
Grade 12	580	600
Fixed monthly nutrition grant per household	155	
Maximum household monthly transfer with no children in senior high school	950	
Maximum household monthly transfer with children in senior high school	1,610	

Table A.1: Amount of Transfers
(100 pesos is approximately 6.5 CAD in 2018)

	1997-1999		1999-2005		2005-2010	
	Univariate	Multivariate	Univariate	Multivariate	Univariate	Multivariate
	(1)	(2)	(3)	(4)	(5)	(6)
Marginality index	0.183		0.0487		0.0168	
(mean=0, s.d.=1)	(0.00325)		(0.00360)		(0.00203)	
Illiterate	1.111	0.411	0.210	-0.146	0.139	0.317
(mean=0.22, s.d.=0.15)	(0.0283)	(0.119)	(0.0275)	(0.126)	(0.0145)	(0.0623)
No toilet	0.559	0.0589	0.188	0.0251	0.0236	-0.0284
(mean=0.47, s.d.=0.26)	(0.0157)	(0.0233)	(0.0160)	(0.0257)	(0.00815)	(0.0125)
No electricity	0.600	0.119	0.106	-0.0142	0.0704	0.0301
(mean=0.24, s.d.=0.22)	(0.0181)	(0.0254)	(0.0181)	(0.0280)	(0.00929)	(0.0138)
No running water	0.542	0.124	0.0858	-0.0752	0.0300	-0.0155
(mean=0.34, s.d.=0.25)	(0.0167)	(0.0218)	(0.0168)	(0.0235)	(0.00804)	(0.0115)
With dirt floor	0.607	0.129	0.196	0.0740	0.0639	0.0607
(mean=0.41, s.d.=0.27)	(0.0138)	(0.0312)	(0.0143)	(0.0327)	(0.00737)	(0.0155)
Share $\leq 2\times$ min. wage	1.211	0.238	0.606	0.672	0.0676	-0.0733
(mean=0.79, s.d.=0.12)	(0.0319)	(0.0481)	(0.0293)	(0.0513)	(0.0156)	(0.0256)
Share below primary school	1.114	-0.0756	0.184	-0.158	0.135	-0.130
(mean=0.24, s.d.=0.14)	(0.0296)	(0.111)	(0.0289)	(0.117)	(0.0154)	(0.0588)
Household crowding	-0.856	-0.216	-0.255	-0.105	-0.0259	0.103
(mean=0.70, s.d.=0.18)	(0.0234)	(0.0305)	(0.0201)	(0.0313)	(0.0118)	(0.0180)
Municipalities	2392	2392	2392	2392	2392	2392

All regressions are estimated using ordinary least squares. Standard errors clustered at the municipality level are in parentheses

Table A.2: Marginality Index and Roll-Out

This table displays the marginality index and its components and how they correlate with changes in exposure shares over the three waves. Odd numbered columns present the correlation between each variable (including the full marginality index) and the change in enrollment ratios. Even columns present the estimates for each component from a multivariate regression. The 'Marginality index' regressor varies at the municipality level and corresponds to the normalized first principal component of the other variables appearing in this table. These variables are shares of individual above 15 in a municipality with the following characteristics: 'Illiterate' (share of illiterates); 'No toilet' (share without access to a toilet); 'No electricity' (share without electricity); 'No running water' (share without access to running water); 'With dirt floor' (share with a dirt floor in their house); 'Share $\leq 2\times$ min. wage' (share earning less than twice the minimum wage); 'Share below primary school' (share with less than primary school education); 'Household crowding' (share with crowding as measured by number of rooms divided by household size).

	Teenagers	*Younger Cohort*			*Older Cohort*	*All Cohorts*	*Placebo Cohorts*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age in 1997	2-6	7-11	11-15	7-15	16-36	7-36	40-47
	-	-	-	-	-	-	-
Age in 2010	15-19	20-24	20-28	15-28	29-49	20-49	53-60
Fertility							
2010 Mean of Cohort	0.17	0.87	1.65	1.19	3.14	2.44	5.13
2010 SD of Cohort	(0.45)	(1.06)	(1.42)	(1.28)	(2.22)	(2.15)	(3.22)
2000 Mean of Cohort	0.17	0.93	1.78	1.29	3.69	2.74	6.31
2000 SD of Cohort	(0.48)	(1.14)	(1.52)	(1.38)	(2.63)	(2.51)	(3.74)
Enrollment 1999	0.00985	-0.0773*	-0.126**	-0.0946**	-0.227***	-0.209***	0.115
$\times Post$	(0.0142)	(0.0396)	(0.0581)	(0.0462)	(0.0872)	(0.0695)	(0.148)
Enrollment 2005	-0.0435**	-0.102**	-0.167***	-0.127***	-0.370***	-0.226***	0.409***
$\times Post$	(0.0176)	(0.0446)	(0.0611)	(0.0485)	(0.0957)	(0.0682)	(0.158)
Marginality	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes

All regressions are estimated using ordinary least squares. Standard errors clustered at the municipality level are in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

This table displays estimated effects of being exposed to *Progresa* on fertility for different cohorts. The dependent variable is the total number of children ever born and varies at the individual-census level. The effect of the treatment is captured by the coefficient on the interaction of a 2010 census dummy with the 'Enrollment 1999' share, which varies at the municipality level and corresponds to the cumulative sum of households enrolled in the program divided by the total number of households in that municipality. The variable 'Enrollment 2005' is the equivalent for the year 2005. The variable 'Marginality' is the index that determined the location of the initial treatment. Combined, these two variables aim to capture the time varying effect of belonging to a poor municipality. 'Municipality' represents fixed effects for the 2,392 municipalities. The 'Year' fixed effect corresponds to a dummy equal to 1 if an observation belongs to the 2010 census and 0 if it belongs to the 2000 census.

Table A.3: Effects on Fertility 2010 vs 2000