

TEENAGE PREGNANCY IN MEXICO: EVOLUTION AND CONSEQUENCES*

EVA O. ARCEO-GOMEZ**

RAYMUNDO M. CAMPOS-VAZQUEZ***

We analyze the consequences of a teenage pregnancy event in the short and long run in Mexico. Using longitudinal and cross-section data, we match females who became pregnant and those who did not based on a propensity score. In the short run, we find that a teenage pregnancy causes a decrease of 0.6-0.8 years of schooling, lower school attendance, fewer hours of work and a higher marriage rate. In the long run, we find that a teenage pregnancy results in a 1-1.2-year loss in years of education, which implies a permanent effect on education, and lower household income per capita.

JEL classification: I00, J10, J11, O54

Keywords: Teenage pregnancy, schooling, labor outcomes, propensity score, matching

1. INTRODUCTION

According to Geronimus and Korenman (1992), “Teenage childbearing has been described as a cause of persistent poverty, and poverty that is transmitted intergenerationally” (p. 1187). As the event of teenage pregnancy may lead to an intergenerational cycle of poverty, the causes and consequences of teenage childbearing have been widely studied among social scientists (see, for example, Hoffman and Maynard, (2008), for an analysis in the United States and Stern (2012), for a sociological analysis in the Mexican case). However, most of the literature on the topic estimates associations or correlations of teenage pregnancy and socioeconomic outcomes and most of the international literature focuses on developed countries.

In this paper, we attempt to fill this void in the literature by analyzing the Mexican case. This is important because teenage mothers are far

* We gratefully acknowledge funding from the World Bank and the comments of Luis F. López-Calva. We are especially grateful to the editor and two anonymous reviewers for their constructive comments. All remaining errors are our own.

** Email: eva.arceo@cide.edu. Address: Centro de Investigación y Docencia Económicas, Carretera Mexico-Toluca 3655, Col. Lomas de Santa Fe, 01210, Mexico DF. Phone: +52-55-57279800, ext. 2759. Fax: +52-55-57279878

*** E-mail: rmcampos@colmex.mx, <http://cee.colmex.mx/raymundo-campos>. Address: El Colegio de México, Centro de Estudios Económicos, Camino al Ajusco 20, Pedregal de Santa Teresa, 10740, Mexico DF. Phone: +52-55-54493000, ext. 4153. Fax: +52-55-56450464.

more common in Mexico than in the United States or other developed countries. According to World Bank data, in Mexico 69 of every 1,000 adolescents between 15 and 19 years old have children, whereas in the United States only 36 per 1,000 do. Compared to other countries in Latin America with similar development levels, Mexico's teenage childbearing rates are just above average: Brazil has a rate of 76 per 1,000 women, but Argentina and Chile have rates of 56 and 57 per 1,000, respectively. Pantelides (2004) reviews the evolution of the phenomenon in Latin America, pointing out that these rates have not decreased significantly in the last decades.

The problem has also been recognized by the Mexican government, which in 2007 implemented PROMAJOVEN, a program targeting teenage mothers who had not yet finished their primary education. In 2010, 41 percent of mothers between 12 and 19 years old had not completed their basic education. The problem is more serious for older cohorts than for younger cohorts. Perhaps due to this heterogeneity in school attainment of adolescent mothers, the program now also helps women to complete up to middle school (9th grade in Mexico). These figures do not take into account that teenage mothers may be systematically different from adolescents who do not have children, and hence the observed educational underachievement may not be entirely due to early motherhood. Our paper will provide additional evidence to justify these kinds of programs in Mexico.

In order to disentangle the effect of teenage childbearing on several socioeconomic outcomes, we match females who became pregnant during adolescence with those who did not, based on a propensity score. In other words, using several observable characteristics we are able to compare very similar individuals whose only difference is the pregnancy event. We find substantial evidence that there is balance and common support between the treatment and control groups after matching. Our analysis focuses on both short- and long-run outcomes. We find that the single most important effect of teenage childbearing is to lower the educational attainment of females by 0.6 to 0.8 years in the short run. Most importantly, we present evidence that this effect is permanent: Our long-run estimates suggest a loss of between 1 and 1.2 years of schooling. There does not seem to be any short-run effect on the household labor supply or household income per capita. However, and most likely due to their lower educational attainment, we find that in the long run teenage mothers live in households with lower income per capita as compared to females who did not become mothers in adolescence.

Determining the causal effects of teenage childbearing has proven to be very elusive. The main empirical challenge in the estimation of the causal effects is that teen mothers are systematically different than adolescents who do not have children. This selection bias suggests that even in the absence of a child, those females who ultimately raise a child during their teenage years would have had a lower socioeconomic status than those females who did not. The literature presents several approaches to identifying the effect of teenage childbearing in the case of the United States. For instance, Bronars and Grogger (1994) analyze the effect of out-of-wedlock motherhood by comparing twin first births to single first births using a couple of censuses. Although teenage mothers tend to be unwed, this identification strategy seems to answer a different empirical question: It estimates the effect of having an additional child in the first birth of single women rather than the effect of the first birth of single women (independently of whether it was a multiple birth or not).

Other more successful approaches have been used. Geronimus and Korenman (1992) compare teen mothers to their childless sisters using several longitudinal surveys, thus removing the unobserved heterogeneity coming from family background. Hotz, McElroy, and Sanders (2005) and Ashcraft and Lang (2006) use miscarriages as an instrumental variable of birth delays. In this way, they estimate the causal effect of age at first birth on several socioeconomic outcomes. Hotz, McElroy, and Sanders (2005) find statistically significant positive effects on the probability of earning a General Educational Development (GED) degree, on the number of hours of work per week, and on wages. In contrast, Ashcraft and Lang (2006) find adverse but modest effects. Finally, Levine and Painter (2003) implement propensity score matching within schools attended by treatment and control teenagers in the United States, finding that teenage mothers are 20 percent less likely to graduate from high school. Similarly, Chevalier and Viitanen (2003) estimate a propensity score matching model using data from Great Britain. They also find adverse effects of teenage childbearing on schooling attainment, labor market experience, and wages in adulthood.

In our view, the evidence on the consequences of teenage pregnancy is more limited for developing countries than for developed countries.¹

1. Another strand of the literature focuses on the determinants of teenage pregnancy and other risky behaviors. For literature on developing countries, see for instance Blunch (2011) on Ghana; Cardoso and Verner (2007) on Brazil; and Marteleto, Lam, and Ranchhod (2008) on South Africa.

Ferre, Gerstenblüth, Rossi, and Triunfo (2009) estimate the impact of childbearing only on educational outcomes using matching methods in Uruguay. These authors work with a cross-section with no retrospective data. As a consequence, they are only able to match females on a very limited number of observable characteristics. Kruger, Berthelon, and Navia (2009) study the effect of teenage pregnancies on high-school completion in Chile using an instrumental variable strategy. The instruments they use reflect the society's and household tolerance for teenage births. In order to measure social acceptance, they estimate the proportion of teenagers in the county who gave birth and the average county rate of unwed births. To measure household tolerance, they use a dummy of whether the mother also had a teenage pregnancy. As for the first set of instruments, we doubt that they meet the exclusion restriction because social acceptance of teenage births may reflect preferences for gender roles, which in turn affect educational attainment. The same is also true for the measure of household tolerance: if having a teen birth reduces the probability of high school completion, the same is true for the teen mother's mother; hence, high school completion of the teen today is affected through the intergenerational transmission of educational attainment.² A paper that is more similar to ours is Ranchhod, Lam, Leibbrandt, and Marteleto (2011) who use the Cape Area Panel Study to estimate propensity-score weighted regression in South Africa. They find a negative effect of a teenage birth on educational attainment, but the effect tends to diminish over time, suggesting that teenage moms catch up with childless teenagers. Unlike our study, Ranchhod, Lam, Leibbrandt and Marteleto (2011) do not exploit the longitudinal nature of their data by estimating a difference-in-difference estimator. None of the studies cited above contrast the short- and long-run effects of teenage births as we do in this paper.

In the case of Mexico, most of the studies analyze the association of pregnancy with outcomes, but lack a clear control group to measure the impact of teenage pregnancy in later outcomes. For example, Stern (2012) conducts an excellent sociological review of the evolution of teenage pregnancy in Mexico. Using qualitative work, Stern (2007) finds that teenage pregnancy occurs in stable couples, and is not due to random encounters. Echarri Cánovas and Pérez Amador (2007)

2. For instance, Navarro Paniagua and Walker (2010) find that children of teenage mothers in Europe have lower educational attainment and are more likely to be teenage mothers themselves.

construct event histories of teenagers, finding that events such as dropping out of school, first consensual union, and leaving the parental home occur before the childbearing event. Menkes and Suárez (2003) find that a low schooling level is associated with lower contraceptive knowledge and a lower age at the first sexual encounter. These two factors, in turn, lead to a higher propensity of less educated women to become pregnant during adolescence. Furthermore, Menkes and Serrano (2010) find that women in poor families have higher rates of teenage pregnancy. Although these studies are relevant and important to increasing our understanding of the teenage pregnancy phenomenon, they only estimate associations of the pregnancy event with different outcomes. These studies also indicate that female teenagers with a pregnancy event are very different from females without the event. Hence, in order to estimate the impact of teenage pregnancy on outcomes like education, income, and work, we apply a novel strategy to the Mexican case in order to compare similar women in terms of observable characteristics.

Our identification strategy follows Levine and Painter (2003) and Chevalier and Viitanen (2003) in the sense that we match females who became mothers during adolescence to females who did not based on a propensity score. Due to data limitations, we are not able to match females within schools or families. However, we exploit two different databases to estimate short- and long-run effects. For the short-run effects we use the Mexican Family Life Survey (MxFLS), which is a longitudinal survey for which there are currently two waves publicly available (2002 and 2005). For the long-run effects, we use the 2011 Social Mobility Survey (EMOVI for its acronym in Spanish), which is a cross-section with socioeconomic information for the individuals when they were 14 years old.

Our results show that the most important effect of teenage childbearing is the permanent, lower educational attainment of the teenage mother. As a result, we find that in the long run, the households of those females who had their first child as teenagers tend to have lower income per capita. We also find that in the short run, teenage mothers reduce their school attendance (hence the lower educational attainment), and their labor supply. Finally, and in contrast with the literature in the United States, we find that having a child during adolescence has a positive effect on the probability of being married. This is most likely a result of cultural differences between Mexico and the United States.

The remainder of the paper is organized as follows. Section 2 shows the aggregate trends in teenage childbearing in Mexico. Section 3 describes the sources of data used in this paper and presents some descriptive statistics. Section 4 explains the empirical strategy that we implement. Section 5 presents the estimations of short- and long-run effects, and finally Section 6 provides concluding remarks and discussion of some policy implications.

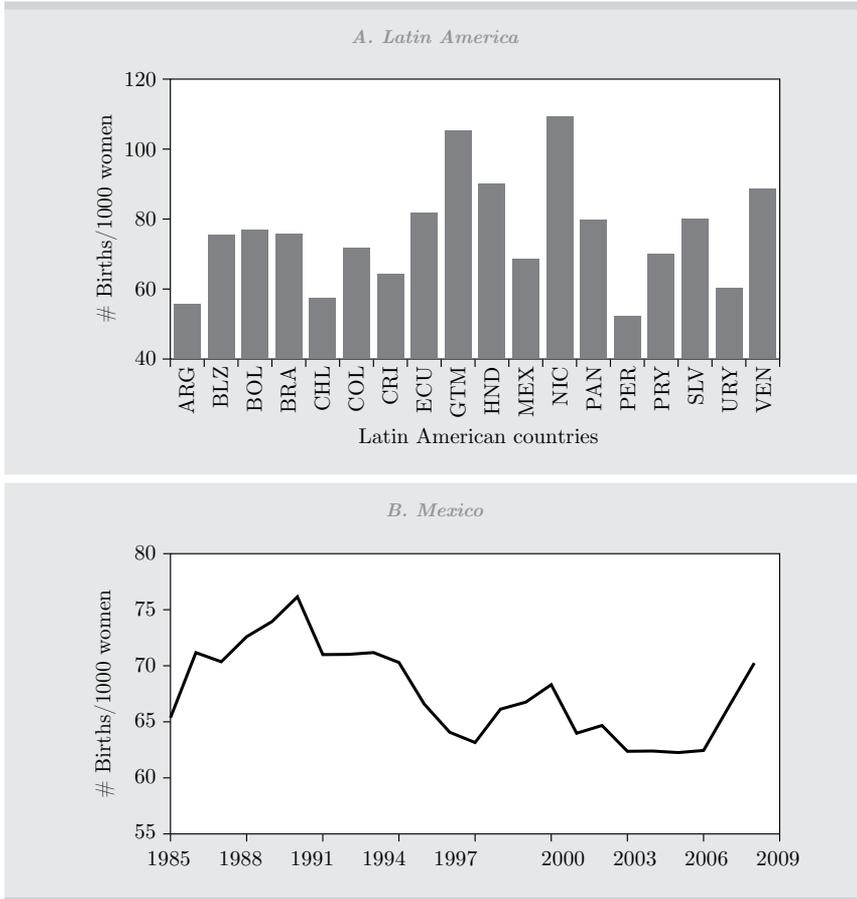
2. AGGREGATE TRENDS

In this section we discuss the aggregate trends for teenage births. The data of this section comes from the World Bank, the Mexican Population Census (1990, 2000, and 2010), and administrative birth records.³ Figure 1, Panel A shows the number of births per 1,000 women among teenagers aged 15-19 in 2009 for a sample of Latin American countries. The unweighted average number of births per 1,000 women for this sample of countries is 75.8, whereas Mexico has a rate equal to 68.6. Among those 18 countries, Mexico has the 6th-lowest rate in the number of births per 1,000 women after Argentina, Chile, Costa Rica, Peru and Uruguay. However, using the same data source for all available countries results in an unweighted world average of 50 births per 1,000 women. Hence, although Mexico shows a slightly lower teenage pregnancy rate as compared to other Latin American countries, its rate is still higher than that of the rest of the world. Panel B shows the evolution of the number of births per 1,000 women among teenagers, based on administrative records.⁴ The number of births per 1,000 women shows a decline from 1990 to 1997, then a relatively stable path from 1998-2006 at around 65 births per 1,000 women, and finally an increase in the 2007-2008 period to almost 70 births per 1,000 women.

3. Census data provides information on the numbers of childbearing women. Our results are very similar to those presented in Menkes and Serrano (2010), even though they use a different survey.

4. Administrative birth records are published by the National Statistical Institute (INEGI) in Mexico and the Ministry of Health. The data include all births registered in order to obtain a birth certificate. These administrative records include age of mother at birth, education, marital status and location of birth (county and state). We use these records in order to provide a broad picture of the evolution of teenage pregnancies. Data can be downloaded from the web sites of INEGI, <http://www.inegi.org.mx/> and the Ministry of Health, <http://www.sinais.salud.gob.mx/basesdedatos/index.html>. We use information from the year of birth rather than year of birth registry. To calculate a series without the problem of right-censoring (births that occurred in the past may be registered at any time in the future), we restrict the data to births registered only in the same year and the year following the birth year, which represents approximately 93% of births.

Figure 1. Number of births per 1,000 women aged 15-19, Latin America and Mexico



Source: Authors' calculations.

Notes: Panel A uses World Bank data for 2009; data available at <http://data.worldbank.org>. ARG=Argentina, BLZ=Belize, BOL=Bolivia, BRA=Brazil, CHL=Chile, COL=Colombia, CRI=Costa Rica, ECU=Ecuador, GTM=Guatemala, HND=Honduras, MEX=Mexico, NIC=Nicaragua, PAN=Panama, PER=Peru, SLV= El Salvador, URY=Uruguay, VEN=Venezuela. Panel B uses information from the Statistical Institute (INEGI). To construct teenage births per 1,000 people, we interpolate population rates using Census data from 1990, 2000, and 2010. We use year of pregnancy rather than year of registry of birth. Due to right-censoring of the data, we limit the calculation to births registered in the same year or year following occurrence (93% of the cases on average).

Panel A in Figure 2 exhibits the fraction of births to teenage mothers, of total births. The percentage of births among teenage mothers is stable at around 16%. In contrast, the percentage of births to single mothers among all births to teenage mothers has increased in the period. As a result, the proportion of births to married women or

Table 1. Aggregate statistics, females aged 15-19, 1990-2010

	Proportions			% Childbearing		
	1990	2000	2010	1990	2000	2010
National	100.0	100.0	100.0	12.3	12.5	13.0
Rural	25.8	25.7	26.0	17.4	16.0	14.9
Urban	74.2	74.3	74.0	10.5	11.3	12.3
Education						
Primary or less	50.1	38.9	28.7	18.3	19.5	17.7
Secondary	45.0	49.1	55.4	6.4	8.3	12.0
More than secondary	5.0	12.0	15.9	4.1	4.9	7.4
Civil Status						
Single	82.5	82.3	82.1	1.3	1.7	2.5
Married	10.8	8.5	4.7	65.3	64.6	63.2
Cohabiting	5.8	8.2	11.7	60.4	60.1	60.0
Other	0.9	1.1	1.5	70.2	71.5	65.7
School Attendance						
Not attending	59.4	54.6	42.9	19.9	22.1	28.0
Attending	40.6	45.4	57.1	1.1	1.1	1.8
Source: Authors' calculations using census data.						
Notes: Sample is restricted to females aged 15-19 years old with a valid answer for the number of own children. The last three columns indicate the percentage of women with at least one child born alive given the condition in the first column.						

women cohabiting has decreased. These findings could be a result of a lower marriage rate triggered by teen pregnancies or a higher age at first marriage that results in fewer married teen mothers. Also, Panel B shows that while in 1985 a teenage mother was more likely to have a primary degree or less (equal to or less than 6 years of schooling), by 2002 that had changed, and a teenage mother was more likely to have a secondary degree (9 to 11 years of schooling). This last finding could be a result of higher educational achievement, and not necessarily due to a decrease in the teen childbearing rate for those with primary schooling or less.

Table 1 provides statistics for females aged 15-19 years old in Mexico for the period 1990-2010 using Census data.⁵ The first three columns show the proportion of each group in the population and the last three

5. Census data are available at the web site of the National Statistical Institute (INEGI) of Mexico, <http://www.inegi.org.mx>.

columns show the percentage of women in that age group with at least one child born alive. The table shows that the percentage living in rural areas (less than 2,500 inhabitants) has remained relatively constant at 25%. On the other hand, education and school attendance has improved during the period of study. An interesting fact is that the proportion of single females is stable at 82% and the proportion of either married or cohabitating is stable at 16-17%. However, the percent of females who are married has decreased substantially over time, from 10.8% in 1990 to 4.7% in 2010. At the same time, the percentage of females who are cohabitating has increased from 5.8% in 1990 to 11.7% in 2010.

When examining data on childbearing teenagers only (columns 4 to 6 in the table), we find that the percentage of females with at least one child born alive has increased from 12.3% in 1990 to 13% in 2010. The increase in childbearing rates is mostly within the urban sector, as females in the rural sector have become less likely to be teenage mothers. Within education groups, the highest childbearing rate is among women with primary schooling or less (less than 8 years). Hence, the trends shown in Panel B of Figure 2 are a result of higher school attainment over time. However, the rate is decreasing slightly for the group of women with primary education and increasing for women with more education such as secondary (9-11 years of schooling) or more than secondary (more than 12 years of schooling). In terms of school attendance, if a woman is attending school the probability that the woman has children is small. When we disaggregate by marital status we find that the childbearing rate is very small (1.3-2.5%) among single women, although this rate doubled in the 1990-2010 period. In Mexico, childbearing is associated with marriage or cohabitation.⁶ Moreover, the childbearing rate among married women has remained stable over time, which indicates that the increase in childbearing has been borne by single women.

3. DATA AND DESCRIPTIVE STATISTICS

We are interested in the effects of teenage pregnancy on individual outcomes of the teenage mother and also on family outcomes. Most of the previous literature has focused on short-run outcomes, given

6. Census data cannot provide information about the timing of the events, so we cannot know whether pregnancy occurred before marriage or cohabitation took place.

the data availability. In this paper, we attempt to measure the consequences of teenage pregnancy both in the short and long run. For the short-run analysis, we use the Mexican Family Life Survey (MxFLS), a nationally representative longitudinal study, for the period 2002-2005.⁷ In the baseline year, the MxFLS was applied to 8,440 households and approximately 35,000 individuals. The follow-up survey was applied in several months between 2005 and 2006 with an attrition rate of approximately 10% at the household level. The survey includes information on demographics, work, and health.

In the short-run analysis, we restrict the MxFLS data to females aged 14 to 18 in 2002 who are childless and not pregnant. Moreover, we further restrict the sample to females who are not married or cohabitating in 2002. Then, we follow those females into the 2005 survey. Hence, we are interested in females who became pregnant between 2002 and 2005 while still a teenager, which represents the treatment variable. Under these restrictions, the final dataset includes 1,003 females with 131 observations in the treatment group.^{8, 9} The teenage pregnancy rate is around 13% in our sample, which is similar to our findings in the previous section. Due to the small sample size, we do not focus on teenage out-of-wedlock childbearing specifically, but we do present some results in the extensions section.¹⁰ The variables in the analysis include age, years of schooling, school attendance indicators, work status, indigenous language, dropout age,¹¹ knowledge of contraceptives, previous sexual activity, Raven test score (percent of correct answers in the test), having been born in rural areas (i.e., localities with less than 2,500 inhabitants), and father absent from the household. We also use information about the head of household: age, years of schooling, and dummies for gender and work status. Finally, we use variables at the household level: household size, number of members ages 0 to 5, 6 to 18, and older than 65, average hours of work for members older

7. Data available at <http://www.ennvih-mxfls.org>.

8. This definition has an important caveat: In the post-treatment year (2005) some of the women in the control group are still teens and could potentially become adolescent mothers. If we had all the completed histories of pregnancies in our sample, then the estimates would likely be higher using the right controls.

9. The attrition rate in this sample is 9.7%. We run t-tests on those women with missing information in 2005 (either missing or lost due to attrition) and women with complete information and there are no significant differences in age, working status, schooling, Raven's test score, previous sexual experience, and knowledge of contraceptives. Hence, there is no evidence that there is sample selection bias.

10. We include cohabitation in the definition of marriage.

11. Some women have not dropped out of school, and in those cases we replace the missing value for dropout age with the age of the individual. Since we are also controlling for age, this has no effect on the estimates.

than 18, average age, income per capita, number of rooms in dwelling, and dwelling characteristics (asset ownership).

In order to measure long-run impacts, we use data from the 2011 Social Mobility Survey (EMOVI).¹² This survey is representative at the national level for both males and females between 25 and 64 years old. The main goal of the survey is to estimate intergenerational mobility. The survey not only records current characteristics, but also collects information about characteristics of the household of origin when the individual was 14 years old. For example, the survey asks about the educational level of both parents and characteristics of the dwelling. The survey includes a question on the age of the individual when he or she had his or her first child. Hence, we define the treatment variable as females who had their first child when they were 15-19 years old. We do not include teenagers who became pregnant when they were 14 years old, in order to include pre-treatment characteristics of the household of origin. This allows us to capture long-run effects because, for example, we can analyze outcomes of females from 6 to 45 years after the teenage pregnancy. However, an important drawback of this survey is that it does not include extensive information about the women when they were teenagers as the MxFLS does, which precludes us from estimating difference-in-differences effects.¹³

Table 2 presents some descriptive statistics for both samples. The MxFLS sample is restricted to the baseline year. Age is relatively similar across samples. In the MxFLS, females who became pregnant between 2002 and 2005 had less education than other females, but the difference is not statistically significant at the 5% level. On the other hand, women in the treatment group had lower school attendance levels and were more likely to work before the pregnancy event. In the case of EMOVI, schooling and proportion working refer to current outcomes. They show that after a teenage pregnancy, women have lower schooling levels and a lower probability of being employed than women without a teenage pregnancy. The subsequent rows show that women who became pregnant come from more disadvantaged backgrounds, as measured by years of schooling of the head of the household (MxFLS) or parents (EMOVI). Also, in the case of the

12. For more information, visit <http://www.ceey.org.mx>.

13. As we discuss below, difference-in-differences estimates are more reliable than a simple difference between treatment and control because they control for individual fixed effects and common trends between treatment and control groups.

Table 2. Descriptive statistics, MxFLS and EMOVI

	MxFLS in baseline (2002)			EMOVI		
	Control	Treatment	Diff	Control	Treatment	Diff
N	872	131		3378	1030	
Age	15.69 [0.047]	15.92 [0.112]	0.23 [0.122]	39.11 [0.210]	39.61 [0.359]	0.49 [0.416]
Yrs school	8.29 [0.074]	8.00 [0.210]	-0.29 [0.223]	8.63 [0.072]	6.83 [0.113]	-1.81 [0.134]*
Working	0.12 [0.011]	0.20 [0.035]	0.08 [0.037]*	0.48 [0.009]	0.44 [0.015]	-0.04 [0.018]*
Attendance	0.72 [0.015]	0.49 [0.044]	-0.23 [0.046]*			
HH size	5.79 [0.067]	5.42 [0.179]	-0.37 [0.192]	5.81 [0.042]	6.29 [0.079]	0.49 [0.089]*
Yrs school (head household)	5.85 [0.146]	5.12 [0.359]	-0.73 [0.388]			
Father: Yrs school				3.99 [0.075]	2.83 [0.106]	-1.15 [0.130]*
Mother: Yrs school				3.70 [0.070]	2.65 [0.103]	-1.05 [0.125]*
Knowledge contraceptives	0.90 [0.009]	0.91 [0.025]	0.00 [0.027]			
Previous sexual exp.	0.02 [0.005]	0.08 [0.023]	0.06 [0.024]*			

Source: Authors' calculations using MxFLS and EMOVI data.

Notes: . Sample is restricted to females aged 15-19 years old with a valid answer for the number of own children in the case of MxFLS. In MxFLS: Treatment is defined as women with a pregnancy event (only 3 women report a pregnancy but no child alive). In EMOVI, treatment is defined as child was born when woman was a teenager. HH Size in EMOVI refers to household size when female was 14 years old. Standard errors in brackets. * denotes significance at 5%.

MxFLS, women who became pregnant were already more sexually active than women in the control group. On the other hand, the Raven test score does not show significant differences between the treatment and control groups. In general, these results show the importance of controlling for selection bias.

4. EMPIRICAL STRATEGY

Our goal in this paper is to estimate the effect of teenage pregnancy on outcome variables such as years of schooling, school attendance, working status, and marriage status. The ideal experiment would be to randomly assign pregnancies to teenagers (treatment) and then compare the outcomes. Obviously, such an experiment would be unethical and unfeasible. We define Y_{1i} as the potential outcome in the treatment state and Y_{0i} as the potential outcome in the control state for individual i and define treatment as $D_i = 1$. The parameter of interest is the average treatment on the treated (*ATT*) defined as the mean difference in outcome variables given treatment, $ATT = E[Y_{1i} - Y_{0i} | D_i = 1]$. However, the term cannot be estimated given that it is not possible to observe the same individual in the treatment and the control group at the same time. This is the “fundamental problem of causal inference” (Holland, 1986). The problem is that the term $E[Y_{0i} | D_i = 1]$ is not observed and has to be estimated (from this point forward we will omit the subscript i for notational simplicity).

We rely on the assumption of selection on observables in order to construct a valid counterfactual. In particular, we assume that conditioning on observable characteristics before the treatment occurs removes differences in the untreated state between teenagers who became pregnant and those who did not. In other words, we assume that $(Y_0 \perp D) | X$, which is commonly referred to in the literature as the conditional independence assumption (CIA) or the unconfoundedness assumption. This assumption means that the outcome for teenagers who did not become pregnant (untreated state), for example years of schooling, is independent of treatment conditional on observable characteristics.

In order to identify the *ATT*, the common support also needs to hold, $\Pr(D = 1 | X) < 1$. This assumption means that for every X there are individuals who do not get the treatment. Ideally, we would like to match individuals in the treatment and control

groups within cells of observable characteristics. However, this is not possible due to the multidimensionality problem. In order to overcome this issue, Rosenbaum and Rubin (1983) propose to estimate propensity scores. These can be easily estimated using a logit or probit of the probability of treatment on observable characteristics, $\Pr(D = 1 | X) = P(X)$. Rosenbaum and Rubin (1983) show that under the CIA:

$$(Y_0 \perp D) | X \Rightarrow (Y_0 \perp D) | P(X) \quad (1)$$

Instead of comparing treatment and control groups within the same set of X , we compare individuals based on an index that summarizes the observable characteristics information. If the assumptions of the model are satisfied, the *ATT* using a propensity score is estimated as:

$$\theta_{ATT}^{PSM} = E_{P(X)|D=1} \{E[Y_1 | D = 1, P(X)] - E[Y_0 | D = 0, P(X)]\} \quad (2)$$

The *ATT* is merely the difference in mean outcomes for treated individuals and mean outcomes of individuals in the control group but reweighted or readjusted by the propensity score, $P(X)$, such that they are as similar as possible to the treatment group in the common support region.^{14, 15}

We estimate the impact of teenage pregnancy in the short and long run. For the long-run estimates, we apply Equation (2). For the short-run estimates, we can improve our estimates by taking advantage of the panel structure of the data. If there is unobserved heterogeneity

14. This is the procedure we follow to estimate the effects: 1. Estimate the propensity score, 2. Match individuals based on the propensity score. In other words, we compare individuals with similar propensity scores and take the difference in outcomes for those individuals.

15. For comparison purposes we also include the results of a simple ordinary least squares (OLS) regression. However, we emphasize that PSM is preferred over regression for several reasons. First, PSM only takes into account observations with very similar values in the propensity score. Observations with dissimilar values are not taken into account to calculate the *ATT*. Second, we show balance in the covariates and common support tests in order to be transparent about the estimation. Third, OLS estimators also suffer from the curse of multidimensionality. And finally, the PSM estimator does not impose as many restrictions on the functional form as OLS.

that is fixed over time for individuals in the sample or common trends between the treatment and control groups, then we can eliminate this bias by estimating difference-in-differences effects:¹⁶

$$\theta_{ATT}^{PSM} = E_{P(X)|D=1} \left(\left\{ E[Y_{t1} | D = 1] - E[Y_{t0} | D = 1] \right\} - \left\{ E[Y_{t1} | D = 0, P(X)] - E[Y_{t0} | D = 0, P(X)] \right\} \right) \quad (3)$$

Hence, the matching difference-in-differences estimator also relies on the assumption of parallel counterfactual trends of the outcomes between the treatment and control groups.

Before estimating the *ATT*, three key aspects need to be considered. First, it is important to question the conditional independence assumption. Of course, the assumption is untestable, but we do have possible checks to investigate whether the assumption is likely to hold. Second, there are no strict rules as to what variables should be included in the propensity-score estimation (Caliendo and Kopeinig, 2008). Third, it is possible that the *ATT* is sensitive to the matching method (Smith and Todd, 2005).

The main assumption of matching on the propensity score is that observable characteristics are balanced between the treatment and control groups. In other words, within some specified values of the propensity score there should be no differences in observable characteristics between the treatment and control group. If there are differences in observable characteristics, then it is likely that there are differences in unobservable characteristics, making the estimation of the *ATT* unfeasible. Below, we present different tests in order to provide evidence of balance in the propensity score. However, it is important to point out that the estimate is correct only if there is no selection on unobservables bias present. If there are unobservable characteristics that differ between treatment and control, then the *ATT* estimate will be biased. We attempt to control for this bias by including a rich set of control variables as well as non-linear effects.

16. The procedure for estimating this equation is similar to Equation (2), however instead of taking differences in the post-treatment period, we also take differences in the pre-treatment period. For example, for the effect on schooling we take the difference after the teenage pregnancy event between treatment and control minus the difference before the teenage pregnancy event.

One of the main advantages of the propensity score is that the information on all observable characteristics is summarized in a single index. There is a trade-off of bias versus efficiency in the number of explanatory variables. On the one hand, Caliendo and Kopeinig (2008), Dehejia and Wahba (1999, 2002), and Heckman, Ichimura, and Todd (1997) mention that omitting important variables that determine treatment could bias the *ATT* estimate. On the other hand, Bryson, Dorsett, and Purdon (2002) point out that including irrelevant variables increases the variance of the *ATT* estimate. Moreover, the assumption of balance needs to hold not only for linear terms but also for non-linear terms. This implies that the propensity score may include interactions and higher-order terms (Dehejia and Wahba, 1999, 2002). This could potentially increase the variance in the *ATT* estimate. Instead of relying on the statistical significance of observable characteristics on the propensity score, we include variables in order to achieve balance. Nonetheless, in the robustness checks section we compare models with variations in the set of observable characteristics included in the propensity score estimation in order to compare the *ATT* and its standard errors.

Smith and Todd (2005) show that the *ATT* estimate may be sensitive to the matching method. Also, Heckman, Ichimura, and Todd (1997) suggest that the matching may be done on the log odds ratio ($\log(P(X)/(1 - P(X)))$) instead of on the propensity score $P(X)$. This is especially recommended when there is choice-based sampling in the survey. We include both recommendations in our analysis.

5. RESULTS¹⁷

As previous literature has pointed out that the *ATT* may vary according to the matching method, we present our results for three different matching methods: (1) matching with a kernel Epanechnikov and a bandwidth of 0.1; (2) matching to the three nearest neighbors within a radius of 0.01; and (3) in order to restrict even further the comparison group, we match treatment and control individuals within urban/rural, age, and school attendance status (for the long-run estimates, we only restrict to urban/rural and age). We also present the results using other matching methods as a robustness check.

17. All our matching results use the ado-file `psmatch2` in Stata provided by Leuven and Sianesi (2003). We employ a logistic regression to estimate the propensity score.

The main results are presented using a propensity score that includes linear, squares and interaction terms. The model using the MxFLS data uses 108 variables and the model using the EMOVI data employs 57 variables.¹⁸ The robustness section includes results for different specifications of the propensity score. Also, we present robustness checks with the log odds ratio as the matching score instead of the propensity score. In general, our results are stable across specifications and matching methods.

5.1. Balance of the propensity score

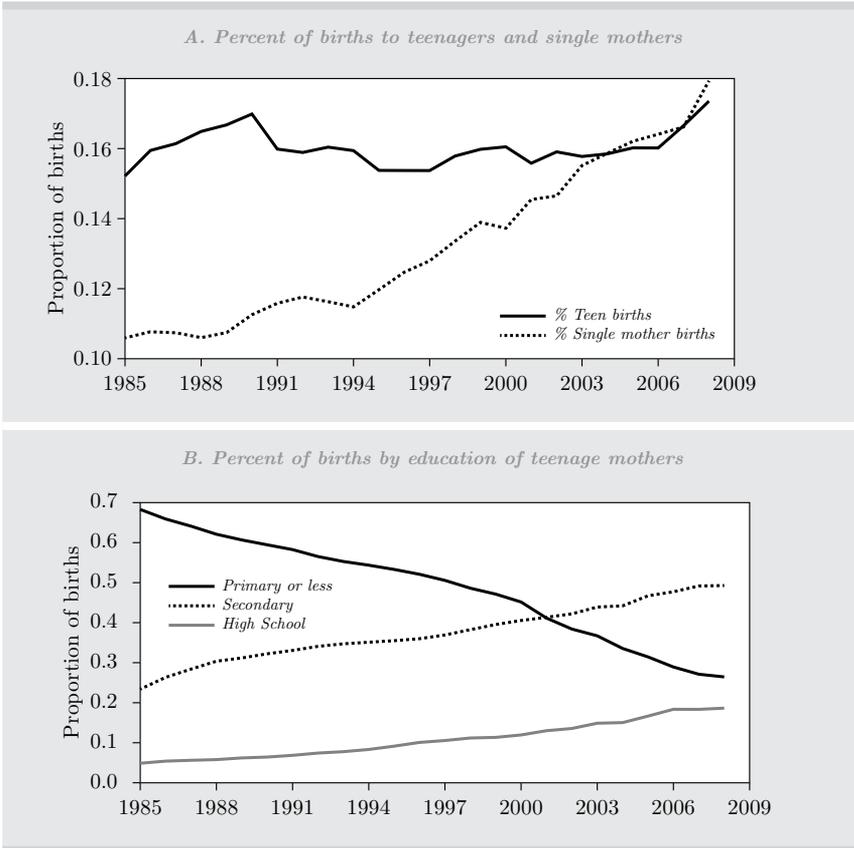
We estimate different tests to corroborate balance in the propensity score. First, we provide graphical evidence based on results by Dehejia and Wahba (1999, 2002) before and after matching to corroborate the balancing and the common support assumptions. We also include the stratification test before and after matching proposed by Dehejia and Wahba (1999, 2002).¹⁹ Second, we include the standardized bias measure proposed by Rosenbaum and Rubin (1985) before and after matching.²⁰ We report only the median standardized bias. According to Caliendo and Kopeinig (2008), a median standardized bias less than 5% is “sufficient.” Third, as proposed by Sianesi (2004), we report the p-value of the joint significance test of the propensity score model

18. MxFLS 2002: age, years of schooling, school attendance, work status, indigenous language, dropout age, Raven’s test score, knowledge of contraceptives, previous sexual activity, rural status, and father absent from the household. The variables included related to the head of the household are: years of education, age, female, and work status. We also include household size; number of members 0-5, 6-18, and older than 65; average hours worked in the household; mean age and income per capita of the household; number of rooms in the household; and several dummies for household assets such as indicator variables for no vehicle, no stove, no public water service and no sewage service. We also include 72 interaction terms between individual variables (age, schooling, work, indigenous, dropout age, Raven’s test score, knowledge of contraceptives and previous sexual activity) and household variables and squares of age and years of schooling. We include 57 variables in the estimation of the propensity score for EMOVI: age and age squared, born in rural area, and information about both parents when individual was 14 years old, namely: education, work status, formal sector job, indigenous language, and what parent the individual was living with. The variables included about the household are: number of siblings, household size, number of rooms and cars, household assets such as no stove, no washing machine, no refrigerator, no television, no public water service, no sewage service, and no electricity. Finally, we include interactions of individual variables with household characteristics as well as squares and interactions of years of education of both parents, and work status of both parents.

19. However, they only present the stratification test before matching, while we believe the result of the test after matching is also informative. The stratification test relies on dividing observations in the treatment and control groups into quintiles or deciles. Then, within each quintile or decile, t-tests are employed for difference in means between treatment and control groups. If we have 10 variables and 5 quintiles, we have 50 tests. We report the percentage of the total tests that fail to reject the null of equal means. Dehejia and Wahba (1999, 2002) point out that this test can be used to select the variables included in the propensity score.

20. The Standardized Bias (SB) is defined as $100 \times (\bar{X}_1 - \bar{X}_0) / \sqrt{0.5(V(X_1) + V(X_0))}$, where the subscript refers to treatment (1) and control (0).

Figure 2. Teenage pregnancy (ages 15-19) in Mexico, 1985-2008



Source: Authors' calculations.

Notes: Panels use information from the Statistical Institute (INEGI). To construct teenage births per 1,000 people, we interpolate population rates using Census data from 1990, 2000, and 2010. We use year of birth rather than year of registry of birth. Due to right-censoring of the data, we limit the calculation to births registered in the same year or year following occurrence (93% of the cases on average). In panel A, the percentage of births reported by single women excludes the percentage of women with invalid information on civil status. % Teen births refers to the percentage of teen births of total births. % Single mother births refers to the percentage of teen births with a single mother (excludes cohabitation). In panel B, around 3-5% of females have invalid education information. Primary or less refers to 8 years of schooling or less, Secondary refers to 9-11 years of schooling, High School refers to 12-15 years of schooling.

before and after matching.²¹ Fourth, we report the percentage of variables that fail to reject the null hypothesis of equal means before and after matching. Finally, we report the number of observations in

21. In other words, we estimate $P(X) = \beta X$ and test the joint hypothesis that $H_0 : \beta = 0$ before and after matching. The procedure after matching includes the weights for each control.

the treatment and control for each matching method. With all these tests, our aim is to provide evidence in favor of the balancing and common support assumptions.

Figure 3 shows box plots and histograms before and after matching. To present the results, we use 3 nearest neighbors within a radius of 0.01. The figure, which includes the results for both the MxFLS and EMOVI, shows that even before matching, the treatment and control groups are not substantially different. Before matching in the MxFLS (Panel A), the mean value of the propensity score for the control group is approximately 0.10 and for the treated group it is approximately 0.25. For the EMOVI (Panel B), the mean values are even closer. Panels E and F show the box plots after matching. The box plots do not show differences in the range of the propensity score between treatment and control. Panels C and D show the number of observations in the treatment and control by deciles of the propensity score. The histograms illustrate that there is a sufficient number of observations

Figure 3. Balance in the propensity score, MxFLS and EMOVI

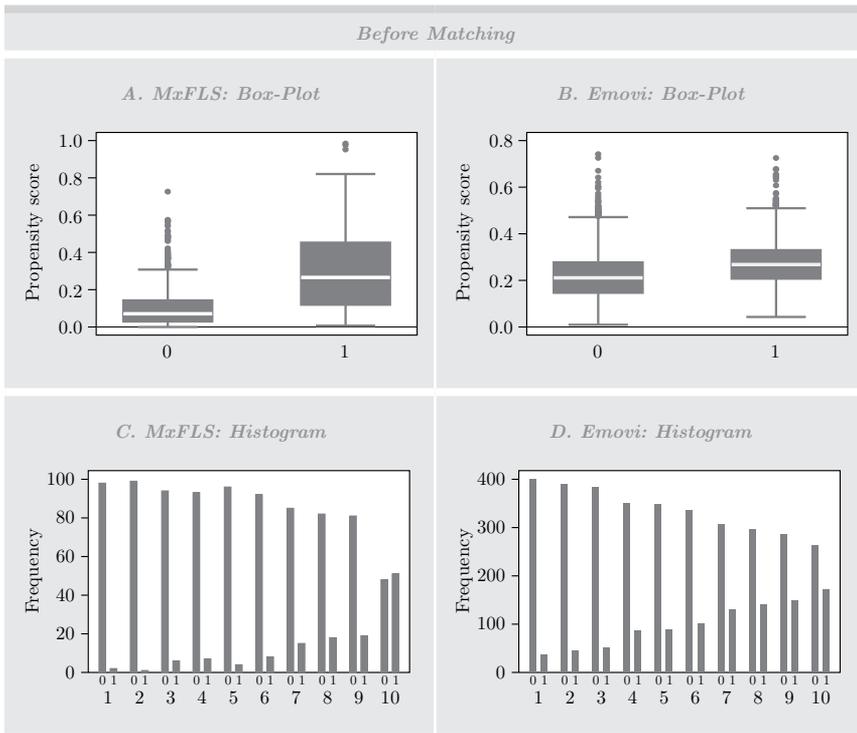
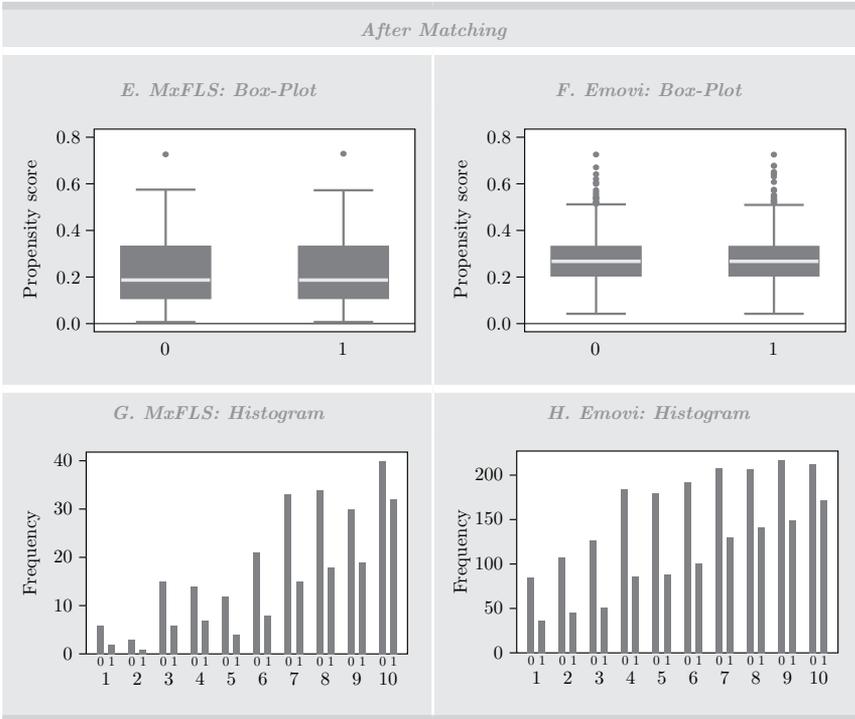


Figure 3. (continued)

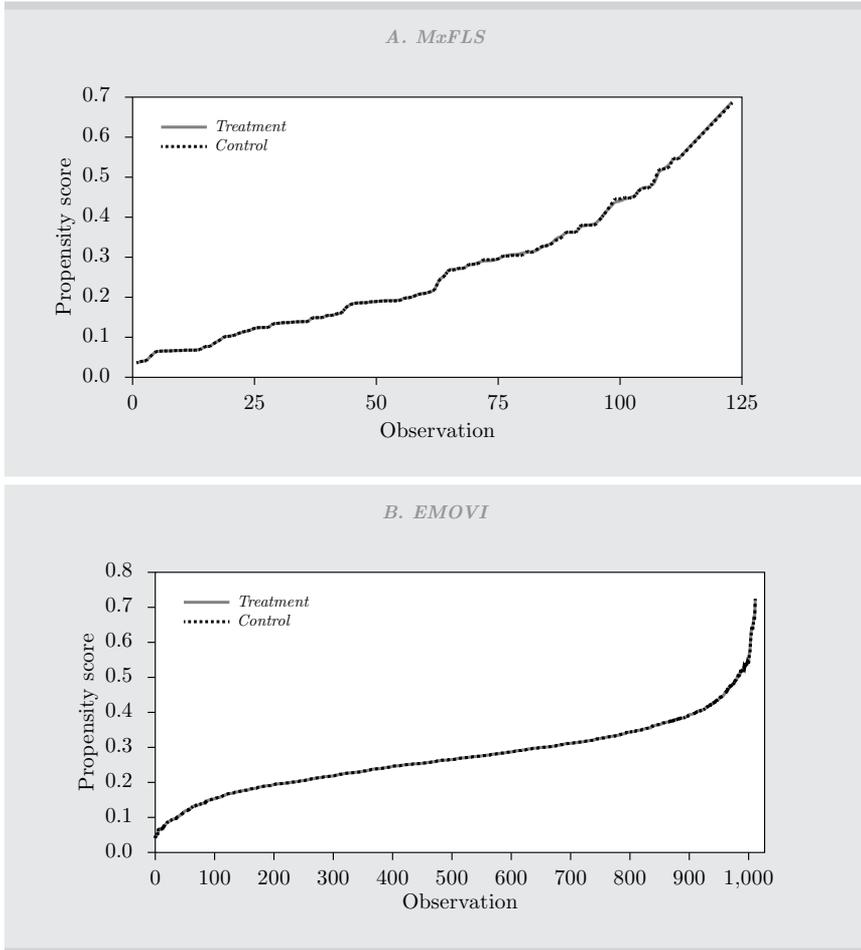


Source: Authors' calculations.

Notes: After matching, figures use the method of 3 nearest neighbors within a radius of 0.01. In the box plots, 0 refers to control observations and 1 to treatment observations. In the histograms, the x-axis has two rows: the first row refers to control and treatment, and the second row to deciles of the estimated propensity score. We include 86 variables in 2002 for the estimation of the propensity score for MxFLS: age, years of schooling, school attendance, work status, indigenous language, knowledge of contraceptives, previous sexual activity, rural status, and father absent from the household. The variables included about the head of the household are: years of education, age, female, and work status. We also include household size, and household members 0-5, 6-18, older than 65, average hours worked in the household, mean age, and income per capita of the household, number of rooms in the household, and dummy variables for household assets, such as: no vehicle, no stove, no public water service and no sewage service. We also include 72 interaction terms between individual variables (age, schooling, work, indigenous, knowledge of contraceptives, and previous sexual activity) and household variables and squares of age and years of schooling. We include 57 variables in the estimation of the propensity score for EMOVI: age and age squared, born in rural areas, and information about both parents when individual was 14 years old, such as: education, work status, formal sector job, indigenous language, what parent the individual was living with. We also include information about the household: number of siblings, household size, number of rooms and cars, and dummies of household assets such as: no stove, no washing machine, no refrigerator, no television, no public water service, no sewage service, and no electricity. Finally, we include the interactions of individual variables with household characteristics as well as squares and interactions of years of education of both parents, and work status of both parents.

in the control group to match the treated group. The after matching histograms show that for each decile we have more observations in the control than in the treated group, with the exception of the top decile in the MxFLS.

Figure 4. Average propensity score in treatment and control, MxFLS and EMOVI



Source: Authors' calculations.

Notes: Matching uses the 3-nearest neighbors method within a radius of 0.01. We sort the treated observations with the propensity score (solid line) and then take the average of the propensity score for the matched controls of each treated observation (dotted line).

Figure 4 shows the estimated propensity score for each treated observation and the average propensity score for the matched controls. What it indicates is that the matching method succeeds in finding very similar observations between the treatment and control groups. In general, Figures 3 and 4 show that the common support condition for the estimation of *ATT* holds.

Table 3. Balance in the propensity score

	DW test		Median bias		LR test		Diff means		# Treat	# Control
	Before	After	Before	After	Before	After	Before	After		
A. MxFLS										
Epanechnikov kernel, bw=0.01	0.03	0.00	13.97	4.69	0.00	0.99	0.32	0.01	118	865
NN-3, radius 0.01	0.03	0.00	13.97	5.65	0.00	0.99	0.32	0.02	118	224
Exact match + NN-3, radius 0.01	0.03	0.00	13.97	6.85	0.00	0.98	0.32	0.02	99	195
B. EMOVI										
Epanechnikov kernel, bw=0.01	0.040	0.00	12.16	0.87	0.00	0.99	0.68	0.00	1024	3376
NN-3, radius 0.01	0.042	0.00	12.16	0.91	0.00	0.99	0.68	0.00	1024	1691
Exact match + NN-3, radius 0.01	0.042	0.00	12.16	1.38	0.00	0.99	0.68	0.00	956	1637

Source: Authors' calculations.

Notes: The first column indicates the matching method. NN refers to nearest neighbor matching. The exact matching method restricts the sample to individuals within rural or urban areas and exact age for EMOVI, and for ENNVIIH it also restricts it to individuals with the same school attendance status. "DW test" refers to the Dehejia and Wahba (1999) stratification test using quintiles of the estimated propensity score. The "Median Bias" column shows the median standardized bias. The "LR test" column shows the p-value of the likelihood ratio test that all coefficients in the regression are equal to zero. The column "Diff Means" shows the percentage of tests out of total possible tests in which the null hypothesis of equal means between treatment and control is rejected. The last two columns indicate the number of observations in treatment and control after matching. For details on the variables included in the propensity score estimation, please refer to the text.

Table 3 provides the balance tests for stratification (Dehejia and Wahba 1999, 2002), standardized bias (Rosenbaum and Rubin 1985), likelihood ratio (Sianesi, 2004), the difference in means, and the number of observations after matching. We include only three matching methods for each survey (a full set of results can be found in Table A1 of the appendix). The matching method is successful in balancing treatment and control groups. After matching, there are no significant differences in observable characteristics between treatment and control. However, balance is relatively more difficult to achieve with MxFLS than with EMOVI as measured by the standardized median bias and the difference in means. Nonetheless, the values are very small and fall within the region of “sufficient” balance mentioned by Caliendo and Kopeinig (2008). In the appendix, we show that balance is more successfully achieved in a model in which the propensity score excludes interaction terms and only includes linear terms. But since excluding important variables may bias the *ATT* estimates, we present the main results using the estimated propensity score with interactions and squares, and as a robustness exercise we show the *ATT* results using the model with linear terms.

5.2. Short-run impacts

Table 4 shows the main results using MxFLS with a difference-in-difference *ATT*. For a simple comparison, we also include the estimate using regression analysis, although as previously mentioned, the PSM estimates are more reliable due to a similar comparison between treatment and control. The table includes the individual outcomes of years of schooling, school attendance, marriage, working, hours of work, and whether the individual left the household by 2005. The table also includes outcomes at the household level, restricting the sample to females who did not leave the household during the period of study.

The results provide evidence that a teenage pregnancy reduces school attainment. Females who had a child between 2002 and 2005 or 2006 have 0.6-0.8 years less of schooling than a female who did not have a child. The estimate is statistically significant, although with relatively large standard errors. If they drop out of school permanently, we should expect the gap to grow; if they drop out temporarily, we should observe a reduction in the gap in the long run, or that the gap remains constant if women select the age to drop out of school. We also find that school attendance decreases. However, it is important to point out that not all teenagers who became pregnant dropped

Table 4. Short-run results, MxFLS

Individual outcomes	Yrs school	Attendance	Married	Working	Hrs work	Left HH
Epanechnikov kernel, bw=0.01	-0.652 [0.220]	-0.283 [0.074]	0.573 [0.059]	-0.176 [0.071]	-10.587 [3.503]	0.417 [0.061]
NN3, Radius 0.01	-0.685 [0.260]	-0.318 [0.082]	0.574 [0.062]	-0.139 [0.078]	-9.127 [3.841]	0.423 [0.063]
Exact match + NN-3, radius 0.01	-0.866 [0.333]	-0.267 [0.076]	0.548 [0.084]	-0.088 [0.105]	-6.475 [4.734]	0.394 [0.082]
Regression	-0.819 [0.153]	-0.281 [0.039]	0.614 [0.027]	-0.168 [0.041]	-10.443 [2.013]	0.429 [0.025]
Household outcomes	Total hours of work	Parents hours of work	Income per capita	HH size		
Epanechnikov kernel, bw=0.01	-1.63 [14.19]	2.31 [8.35]	69.71 [283.37]	1.10 [0.28]		
NN3, Radius 0.01	2.29 [15.89]	1.69 [9.21]	9.31 [324.64]	1.23 [0.29]		
Exact match + NN-3, radius 0.01	3.99 [18.45]	-0.33 [11.25]	117.85 [415.82]	1.32 [0.32]		
Regression	3.99 [6.57]	-0.33 [4.13]	117.85 [146.78]	1.32 [0.11]		

Source: Authors' calculations.

Notes: The model includes linear and interaction terms; in total, the estimated propensity score includes 108 variables. Exact matching restricts the sample to individuals within the same rural/urban, age, and school attendance cells. The first panel includes outcomes at the individual level. The second panel restricts the sample to females who did not leave the household between 2002-2005 and analyzes outcomes at the household level. Standard errors are estimated using 500 bootstrap replications. The "Regression" row includes the estimates of teenage pregnancy controlling for 26 variables: individual characteristics, rural status, parental and household characteristics, and assets at the household level.

out of school by 2005-2006. The estimate implies that between 27 and 32% of teenagers who became pregnant are not attending school after pregnancy, compared to similar teenagers in the control group.

A key difference from results in the literature on the United States is that teenage pregnancy does not reduce the likelihood of marriage. In fact, a larger share of childbearing teenagers are married as compared to similar childless teenagers. These results are very possibly due to cultural differences between Mexico and the United States. In general, Mexican females tend to marry more frequently and teenage out-of-wedlock pregnancies are severely stigmatized by Mexican society. In the extension section, we analyze outcomes for teenage out-of-wedlock pregnancy.

Additionally, there is some evidence that teenage pregnancy reduces the probability of working by 9-18 percentage points. However, the standard errors are large and in the case of exact matching the results are not statistically significant. But there is statistical evidence that teenagers who became pregnant reduce their hours of work by 6-10 hours on average. Also, teenagers who became pregnant are 39-41% more likely to leave their household than teenagers who did not become pregnant. This latter finding is a result of marriage.

It is important to analyze not only the consequences of childbearing by teenagers themselves, but the consequences for the family of origin. This is interesting but hard to measure. As we analyze longitudinal data, we observe households in two periods. But if the teenager leaves the household, we are only able to observe information for the newly formed family. We could link the information to the family of origin, but in this case the interpretation of the treatment effect would not be clear, given that the treatment on the family of origin is somewhat lost. For these reasons, we focus on teenagers who did not leave the household of origin during the period of study. Thus, we are comparing how the family is affected in the short run when a teenager becomes pregnant.

The bottom panel in Table 4 includes the results at the household level. For females who did not leave the household of origin, we observe little changes at the household level. There is no evidence that the family reacts with more hours of work (this variable excludes the labor supply of the childbearing adolescent). The results are close to zero and not statistically significant. This effect may be due to more hours of work of parents and fewer hours of work of siblings. In order to test for this possibility, we estimate the effect on parents' labor

supply (as shown in the next column). However, the estimates are not statistically significant for hours of work of parents. There seems to be no adjustment in the labor supply of other household members. This could be due to the timing of data collection. We observe teenagers after the birth of their child, and it is possible that the household has already adjusted to previous levels of hours of work. We also do not find any significant effect on income per capita, but there is a clear increase in household size. The reason that the effect on household size is greater than one is that some teenagers became pregnant and their husband or partner moved in with her and her family. In sum, we find little evidence that a pregnancy for a teenager who stays in the household of origin has significant consequences for the family of origin itself. It is important to stress that we do not measure the immediate effects of pregnancy but rather an average of 1 to 2 years after pregnancy.

5.3. Long-run impacts

Table 5 presents the estimates using EMOVI, for both PSM and regression. Women who became pregnant when they were teenagers attain less schooling than females who did not become pregnant. We find that the difference is close to 1 year of education. Although the estimate is larger than the short-run results, it is not possible to reject the null hypothesis of equal effects. However, the results do not support the hypothesis that the gap in education is reduced in the long run. On the contrary, once a teenage pregnancy occurs, the difference in years of education will be maintained. The estimate using regression analysis is much lower, but this is likely due to a lack of adequate controls (in the PSM framework we are controlling for many more variables including interaction and non-linear terms).

Females who became pregnant while adolescents are more likely to be married, and in turn less likely to be single in the long run than their counterparts. At the same time, they are more likely to go through a divorce or separation. Hence, we do not find any evidence in the short or long run that a teenage pregnancy reduces the likelihood of marriage. Also, it seems that a teenage pregnancy is considered as an “extra child”, otherwise they would have had the same total number of children as the control females. Moreover, the increase in the number of children results in a larger household size. As for the impact on the labor supply, although the effect of teenage pregnancy on work is

Table 5. Long-run results, EMOVI

	Yrs school	Married	Single	Separated	# Children	Works	HH size	Income per capita
Epanechnikov kernel, bw=0.01	-1.065 [0.106]	0.050 [0.017]	-0.124 [0.012]	0.051 [0.013]	1.085 [0.049]	-0.043 [0.018]	0.555 [0.065]	-322.322 [72.89]
NN3, radius 0.01	-1.122 [0.169]	0.059 [0.022]	-0.132 [0.017]	0.048 [0.016]	1.120 [0.066]	-0.054 [0.024]	0.560 [0.086]	-338.132 [140.32]
Exact match + NN-3, radius 0.01	-1.159 [0.168]	0.045 [0.024]	-0.108 [0.018]	0.042 [0.020]	1.062 [0.067]	-0.043 [0.026]	0.528 [0.089]	-401.429 [128.12]
Regression	-0.414 [0.128]	0.022 [0.019]	-0.033 [0.015]	0.011 [0.014]	-0.163 [0.055]	-0.011 [0.019]	-0.217 [0.064]	286.60 [140.28]

Source: Authors' calculations.

Notes: The model includes linear and interaction terms; in total, the estimated propensity score includes 57 variables. Exact matching restricts the sample to individuals within the same rural/urban, age, and school attendance cells. Standard errors are estimated using 500 bootstrap replications.

negative, it is not statistically significant. Hence, there is no evidence that having children as an adolescent reduces the likelihood of working in the long run. However, there is some evidence of a lower income per capita in the household, which is most likely a consequence of a lower educational attainment.

5.4. Extensions and robustness checks

In the previous sections, we have not analyzed outcome for pregnancies out of wedlock. It is possible that out-of-wedlock pregnancies are more costly to teenagers. The MxFLS identifies the year of pregnancy and the year of marriage. We restrict the treatment sample to females who are not married in 2005 and females who had a birth before marriage, reducing it to 76 observations instead of 131.²² Table 6 shows the estimates for this sample.

There are no large differences between the estimates using the full sample and restricting it to out-of-wedlock pregnancies. Both the loss in years of education and the reduction in the percent working are similar to the full sample. Since we dropped pregnancies after marriage from the sample, the effect on marriage decreases but it is still high and close to 34%. Hence, there is no evidence that out-of-wedlock pregnancies are different than teenage pregnancies in a marriage. Table 6 also includes results for the EMOVI, restricting the sample to females between 25 and 39 years old. There is no evidence that the loss in years of education or the probability of working is different from the full sample. However, the percentage that is married is relatively higher than in the full sample, although we cannot reject the hypothesis of equal coefficients.

In Table 7 we provide robustness results using more matching methods and results employing a different estimated propensity score. Panel A shows the main propensity score, which includes the interactions and squares of many variables. Results are robust to changes in the matching method. Panel B modifies the estimated propensity score by including only linear terms. In total, we include only 27 and 26 variables for the MxFLS and EMOVI, respectively. The *ATT* are, on average, similar to previous estimations, but the standard error is lower, as suggested by Bryson, Dorsett, and Purdon (2002). Panel C

22. Of the 76 observations, 41 are unmarried and 35 are married in 2005.

Table 6. Extensions: Short-run effects of teenage, out-of-wedlock pregnancy and teenage pregnancy effects at 25-39 years old

	MxFLS: Unmarried teenagers			EMOVI: 25-39 years old		
	Yrs school	Works	Marriage	Yrs school	Works	Marriage
Epanechnikov kernel, bw=0.01	-0.616 [0.341]	-0.154 [0.115]	0.345 [0.081]	-1.177 [0.152]	-0.065 [0.025]	0.111 [0.026]
NN3, radius 0.01	-0.608 [0.401]	-0.157 [0.128]	0.343 [0.083]	-1.061 [0.204]	-0.061 [0.032]	0.115 [0.032]
Exact match + NN-3, radius 0.01	-0.813 [0.473]	-0.122 [0.141]	0.306 [0.115]	-1.226 [0.215]	-0.045 [0.034]	0.100 [0.035]

Source: Authors' calculations.
 Notes: The model includes linear and interaction terms; in total, the estimated propensity score includes 108 variables for MxFLS and 57 variables for EMOVI. Exact matching restricts the sample to individuals within the same rural/urban, age, and school attendance (for MXFLS) cells. Standard errors are estimated using 500 bootstrap replications.

Table 7. Robustness tests

	MxFLS			EMOVI			
	Yrs school	Works	Marriage	Yrs school	Works	Marriage	Income per capita
A. Propensity score with interactions and squares							
Epanechnikov kernel, bw=0.0025	-0.539 [0.327]	-0.153 [0.112]	0.624 [0.073]	-1.099 [0.118]	-0.047 [0.020]	0.051 [0.017]	-345.757 [82.12]
Gaussian kernel, bw=0.01	-0.615 [0.259]	-0.111 [0.073]	0.598 [0.062]	-1.084 [0.103]	-0.044 [0.018]	0.053 [0.017]	-305.646 [71.94]
NN3, radius 0.025	-0.610 [0.313]	-0.147 [0.093]	0.574 [0.069]	-1.129 [0.168]	-0.054 [0.024]	0.060 [0.022]	-334.045 [140.25]
B. Propensity score with linear terms							
Epanechnikov kernel, bw=0.01	-0.837 [0.177]	-0.177 [0.050]	0.622 [0.046]	-1.158 [0.145]	-0.043 [0.017]	0.059 [0.018]	-318.719 [75.10]
NN3, radius 0.01	-1.005 [0.238]	-0.217 [0.063]	0.616 [0.050]	-1.163 [0.303]	-0.039 [0.051]	0.069 [0.044]	-279.820 [165.54]
Epanechnikov kernel, bw=0.0025	-0.879 [0.210]	-0.177 [0.053]	0.609 [0.046]	-1.137 [0.156]	-0.043 [0.019]	0.065 [0.019]	-313.687 [165.57]
NN3, radius 0.025	-1.021 [0.232]	-0.217 [0.064]	0.619 [0.050]	-1.157 [0.302]	-0.040 [0.051]	0.070 [0.044]	-279.689 [80.72]

Table 7. (continued)

	MxFLS			EMOVI			
	Yrs school	Works	Marriage	Yrs school	Works	Marriage	Income per capita
C. (A) + matching in the log odds ratio							
Epanechnikov kernel, bw=0.01	-0.564 [0.213]	-0.123 [0.068]	0.589 [0.059]	-1.090 [0.110]	-0.045 [0.019]	0.054 [0.017]	-306.752 [76.69]
NN3, radius 0.01	-0.626 [0.257]	-0.104 [0.072]	0.591 [0.062]	-1.130 [0.171]	-0.055 [0.024]	0.060 [0.022]	-336.757 [141.56]
Epanechnikov kernel, bw=0.0025	-0.536 [0.257]	-0.147 [0.089]	0.618 [0.065]	-1.068 [0.128]	-0.042 [0.020]	0.050 [0.019]	-341.993 [104.20]
NN3, radius 0.025	-0.600 [0.264]	-0.169 [0.085]	0.565 [0.064]	-1.132 [0.171]	-0.055 [0.024]	0.060 [0.022]	-334.589 [140.64]

Source: Authors' calculations.

Notes: Panel A includes linear and interaction terms; in total, the estimated propensity score includes 108 variables for MxFLS and 57 variables for EMOVI. Panel B includes only linear terms in the estimation of the propensity score. We include 27 and 26 variables in the MxFLS and EMOVI, respectively. Standard errors are estimated using 500 bootstrap replications.

matches on the log of odds ratio of the main estimated propensity score as suggested by Heckman, Ichimura, and Todd (1997). In general, the impact on years of schooling in the short run varies from -0.56 to -1 and in the long run from -1.09 to -1.16. Both are within the standard errors obtained for the main estimates. The impact on income per capita in the long run is consistently negative and varies from -279 to -346. In sum, the main estimates are robust to the matching method and to the estimated propensity score.

6. CONCLUSIONS

In this paper we estimate the effect of teenage childbearing on several outcomes for the teenage mother and her family of origin in the short run, and also the long-run effects on the mother. The identification of the causal effect of teenage childbearing has proven to be very elusive due to selection bias: Those adolescents who give birth to a child are systematically different from adolescents who do not have children. For instance, we find that in the case of Mexico, treated teenagers tend to be more sexually active before pregnancy and come from more disadvantaged backgrounds.

We attempt to solve this selection problem by implementing a propensity-score matching model using two different data sources: a longitudinal survey (Mexican Family Life Survey, or MxFLS), and a cross-section survey designed to measure mobility in Mexico (Social Mobility Survey, or EMOVI). Therefore, we have information on the individual and her household when she was 14 years old. The MxFLS allows us to estimate the short-run effects on the teenage mother and her family of origin. On the other hand, the EMOVI enables us to estimate the long-run effects on the teenage mother. However, the estimates from MxFLS are more reliable as they allow us to estimate difference-in-differences models.

According to our results, the single most important effect of teenage childbearing is a lower educational attainment of the teenage mother, both in the short and long run. As a result, we find that in the long run the households of those females who had their first child as teenagers tend to have a lower income per capita. We also find that in the short run, teenage mothers reduce their college attendance (hence the lower educational attainment) and reduce their labor supply. We do not find any significant effects on labor supply of other household members in the short run, nor on the labor supply of the teenage

mothers themselves in the long run. Finally, and in contrast with the literature in the United States, we find that having a child during adolescence has a positive effect on the probability of being married. This difference is most likely a result of cultural differences between Mexico and the United States.

Although still greatly debated, there is evidence that teenage childbearing is associated with higher levels of poverty and welfare dependence in the United States. To our knowledge, there is not a large literature on the effects of teenage childbearing for developing countries. This paper contributes to fill that gap in the literature. Our findings provide evidence that teenage childbearing has adverse effects in the Mexican context. The fact that teenage childbearing prevents teenage mothers from continuing their human capital investments shows that teenage childbearing may have a deleterious effect on the probability of living in a poor household. Moreover, given that there is little social mobility in Mexico (Torche, 2010), teenage childbearing may be a gateway into an intergenerational poverty trap. As such, our work has two important policy implications. First, programs aimed at preventing teenage pregnancies, such as sexual education during primary and secondary education, should be expanded, as should access to contraceptives through public health systems. Second, once a teenager becomes pregnant, the state should provide support in the form of childcare and merit scholarships, to prevent the teenage mother from dropping out of school. The latter measure is partially being addressed by PROMAJOVEN. However, the program is still limited to the poorest population even though not all teenage mothers meet that criterion.

REFERENCES

- Ashcraft, A. and K. Lang (2006), “The consequences of teenage childbearing,” Working Paper 12485, National Bureau of Economic Research.
- Blunch, N. (2011), “Skills, schooling and non-marital teenage pregnancy in Ghana,” unpublished manuscript, Washington and Lee University.
- Bronars, S.G. and J. Grogger (1994), “The economic consequences of unwed motherhood: Using twin births as a natural experiment,” *The American Economic Review* **84**(5): 1141–56.
- Bryson, A., R. Dorsett, and S. Purdon (2002), “The use of propensity score matching in the evaluation of labour market policies,” Working Paper 4, Department for Work and Pensions.
- Caliendo, M. and S. Kopeinig (2008), “Some practical guidance for the implementation of propensity score matching,” *Journal of Economic Surveys* **22**(1): 31–72.
- Cardoso, A.R. and D. Verner (2007), “Youth risk taking behavior in Brazil: Drug use and teenage pregnancies,” Discussion Paper Series 3030, Institute for the Study of Labor.
- Chevalier, A. and T.K. Viitanen (2003), “The long-run labour market consequences of teenage motherhood in Britain,” *Journal of Population Economics* **16**(2): 323–43.
- Dehejia, R.H. and S. Wahba (1999), “Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs,” *Journal of the American Statistical Association* **94**(448): 1053–62.
- . (2002). “Propensity-score matching methods for nonexperimental causal studies,” *The Review of Economics and Statistics* **84**(1): 151–61.
- Echarri Cánovas, C. and J. Pérez Amador (2007), “En tránsito hacia la adultez: eventos en el curso de vida de los jóvenes en México,” *Estudios demográficos y urbanos* **22**(1): 43–77.
- Ferre, Z., M. Gerstenblüth, M. Rossi and P. Triunfo (2009), “The impact of teenage childbearing on educational outcomes,” unpublished manuscript, Society for the Study of Economic Inequality.
- Geronimus, A.T. and S. Korenman (1992), “The socioeconomic consequences of teen childbearing reconsidered,” *The Quarterly Journal of Economics* **107**(4): 1187–214.
- Heckman, J.J., H. Ichimura and P.E. Todd (1997), “Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme,” *The Review of Economic Studies* **64**(4): 605–54.
- Hoffman, S.D. and R.A. Maynard (2008), *Kids having kids: Economic costs and social consequences of teen pregnancy*, 2nd edition. Washington, D.C.: Urban Institute Press.
- Holland, P.W. (1986), “Statistics and causal inference,” *Journal of the American Statistical Association* **81**(396): 945–60.

- Hotz, V.J., S.W. McElroy, and S.G. Sanders (2005), "Teenage childbearing and its life cycle consequences: Exploiting a natural experiment," *The Journal of Human Resources* **40**(3): 683–715.
- Kruger, D., M. Berthelon and R. Navia (2009), "Adolescent motherhood and secondary schooling in Chile," Discussion Paper Series 4552, Institute for the Study of Labor.
- Leuven, E. and B. Sianesi (2003), "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing." Statistical Software Components, Boston College, Department of Economics.
- Levine, D.I. and G. Painter (2003), "The schooling costs of teenage out-of-wedlock childbearing: Analysis with a within-school propensity-score-matching estimator," *Review of Economics and Statistics* **85**(4): 884–900.
- Marteletto, L., D. Lam, and V. Ranchhod (2008), "Sexual behavior, pregnancy, and schooling among young people in urban South Africa," *Studies in Family Planning* **39**(4): 351–368.
- Menkes, C. and O. Serrano (2010), "Condicionantes sociodemográficos del embarazo adolescente en México," Working paper, *X Reunión Nacional de Investigación Demográfica en México*.
- Menkes, C. and L. Suárez (2003), "Sexualidad y embarazo adolescente en México," *Papeles de población* **35**: 233–62.
- Navarro Paniagua, M. and I. Walker (2010), "The impact of teenage motherhood on the education and fertility decisions of their children: Evidence for Europe," unpublished manuscript, XXIV Annual Conference of the European Society for Population Economics.
- Pantelides, E.A. (2004). "Aspectos sociales del embarazo y la fecundidad adolescente en América Latina," *Revista Notas de Población* **78**: 7–33.
- Ranchhod, V., D. Lam, M. Leibbrandt, and L. Marteletto (2011), "Estimating the effect of adolescent fertility on educational attainment in Cape Town using a propensity score weighted regression," in *Fifth Annual PopPov Conference on Population, Reproductive Health, & Economic Development*.
- Rosenbaum, P.R. and D.B. Rubin (1983), "The central role of the propensity score in observational studies for causal effects," *Biometrika* **70**(1): 41–55.
- . (1985) "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score," *The American Statistician* **39**(1): 33–8.
- Sianesi, B. (2004), "An evaluation of the Swedish system of active labor market programs in the 1990s," *The Review of Economics and Statistics* **86**(1): 133–55.
- Smith, J. and P.E. Todd (2005), "Does matching overcome Lalonde's critique of nonexperimental estimators?" *Journal of Econometrics* **125**(2): 305–53.
- Stern, C. (2007), "Estereotipos de género, relaciones sexuales y embarazo adolescente en las vidas de jóvenes de diferentes contextos socioculturales en México," *Estudios Sociológicos* **25**(73): 105–29.

- Stern, C. (2012), *El “problema” del embarazo en la adolescencia. Contribuciones a un debate*. México: El Colegio de México.
- Torche, F. (2010), “Cambio y persistencia de la movilidad intergeneracional en Mexico,” in Serrano, J. and F. Torche, eds., *Movilidad social en Mexico. Poblacion, desarrollo y crecimiento*: 71–134. Mexico City: Centro de Estudios Espinosa Yglesias.

APPENDIX

Table A1. Balance in the propensity score: robustness

	DW Test		Median Bias		LR Test		Diff Means		# Treat	# Control
	Before	After	Before	After	Before	After	Before	After		
	A. MxFLS									
Epanechnikov kernel, bw=0.0025	0.03	0.00	13.97	4.34	0.00	0.99	0.32	0.05	97	597
NN-1, radius 0.01	0.03	0.00	13.97	8.95	0.00	0.06	0.32	0.03	118	99
Gaussian kernel, bw=0.01	0.03	0.00	13.97	10.06	0.00	0.99	0.32	0.11	131	868
NN-3, radius 0.025	0.03	0.00	13.97	4.64	0.00	0.99	0.32	0.02	122	226
Propensity score with linear terms										
Epanechnikov kernel, bw=0.01	0.06	0.00	10.43	1.63	0.00	0.99	0.24	0.00	125	862
NN-3, radius 0.01	0.06	0.00	10.43	3.65	0.00	0.99	0.24	0.00	125	269
Exact match + NN-3, radius 0.01	0.06	0.00	10.43	6.71	0.00	0.98	0.24	0.00	114	236
Kernel Epanechnikov, bw=0.0025	0.06	0.00	10.43	3.48	0.00	0.99	0.24	0.00	120	674

Table A1. (continued)

	DW Test		Median Bias		LR Test		Diff Means		# Treat	# Control
	Before	After	Before	After	Before	After	Before	After		
B. EMOVI										
Epanechnikov kernel, bw=0.0025	0.042	0.00	12.16	0.45	0.00	0.99	0.68	0.00	1012	3302
NN-1, radius 0.01	0.042	0.00	12.16	1.43	0.00	0.99	0.68	0.00	1024	754
Gaussian kernel, bw=0.01	0.042	0.00	12.16	0.59	0.00	0.99	0.68	0.00	1030	3378
NN-3, radius 0.025	0.042	0.00	12.16	0.98	0.00	0.99	0.68	0.00	1030	1691
Propensity score with linear terms										
Epanechnikov kernel, bw=0.01	0.08	0.00	10.71	0.33	0.00	0.99	0.62	0.00	1025	3377
NN-3, radius 0.01	0.08	0.00	10.71	1.45	0.00	0.99	0.62	0.00	1025	1781
Exact match + NN-3, radius 0.01	0.08	0.00	10.71	0.93	0.00	0.99	0.62	0.00	972	1687
Epanechnikov kernel, bw=0.0025	0.08	0.00	10.71	0.75	0.00	0.99	0.62	0.00	1019	3327

Source: Authors' calculations.
 Notes: The first column indicates the matching method. NN refers to nearest-neighbor matching. The exact matching method restricts the sample to individuals within rural or urban areas and exact age for EMOVI, and for ENNVIH it also restricts the sample to individuals with the same school attendance status. "DW test" refers to the Dehejia and Wahba (1999) stratification test using quintiles of the estimated propensity score, the "Median Bias" column shows the median standardized bias, while the "LR test" column shows the p-value of the likelihood ratio test that all coefficients in the regression are equal to zero. The "Diff Means" column shows the percentage of tests, of the total possible tests, in which the null hypothesis of equal means between treatment and control is rejected. The last two columns indicate the number of observations in treatment and control after matching. We include 86 variables in 2002 in the estimation of the propensity score for MxFLS including interactions. We include 57 variables in the estimation of the propensity score for EMOVI. The models with the "Propensity score with linear terms" do not include interaction terms. In this case, MxFLS includes 27 variables and EMOVI includes 26 variables.