CAF - WORKING PAPER #2021/16
First version: November 25, 2021 (current)

Teen childbearing in Latin America: the mother-daughter link

Matilde P. Machado[1] | **Ricardo Mora**[2] | **Karen Olivo**[3]

[1]Associate Professor, Department of Economics, Universidad Carlos III Madrid. matilde.machado@uc3m.es

[2]Associate Professor, Department of Economics, Universidad Carlos III Madrid. ricmora@eco.uc3m.es

[3]nuji01@gmail.com

Using DHS data for six countries in Latin America and the Caribbean region, we estimate the relation between a mother's teenage childbearing and that of her daughter's. Our results show that restricting the estimating sample to mother-daughter matches in the data leads to large negative selection bias in the estimated effect because missing matches are non-random and potentially affected by the teen childbearing status of mothers and daughters. We deal with this selection bias by developing a methodology that uses all available data, including incomplete mother-daughter pairs, and allows missing observations to be endogenous. Our preferred specification shows that being the daughter of a teen mother increases the chances of being a teen mother between 7.4 and 22.2 percentage points (between 42 and 138%). In general, it is also associated with other negative outcomes such as lower educational achievement, acceptance of risky sexual behavior and submissive gender roles in sexual relationships.

KEYWORDS
teen childbearing, teen pregnancy, intergenerational transmission, non-ignorable missingness, econometric methods for missing data, DHS data

CAF - DOCUMENTO DE TRABAJO #2021/16

Primera versión: Noviembre 25, 2021 (actual)

Maternidad adolescente en América Latina: el vínculo entre madre e hija

Matilde P. Machado[1] | **Ricardo Mora**[2] | **Karen Olivo**[3]

[1]Profesor Titular, Departamento de Economía, Universidad Carlos III Madrid. matilde.machado@uc3m.es

[2]Profesor Titular, Departmento de Economía, Universidad Carlos III Madrid. ricmora@eco.uc3m.es

[3]nuji01@gmail.com

Utilizando los datos de las Encuestas de Demografía y Salud (DHS, por sus siglas en inglés) de seis países de la región de América Latina y el Caribe, estimamos la relación entre la maternidad adolescente de una madre y la de su hija. Nuestros resultados muestran que restringir la muestra de estimación a los emparejamientos madre-hija en los datos conduce a un gran sesgo de selección negativo en el efecto estimado porque los emparejamientos perdidos no son aleatorios y están afectados por la situación de maternidad adolescente de madres e hijas. Tratamos este sesgo de selección desarrollando una metodología que utiliza todos los datos disponibles, incluidas las parejas madre-hija incompletas, y permite que las observaciones que faltan sean endógenas. Nuestra especificación preferida muestra que ser hija de una madre adolescente aumenta las posibilidades de ser madre adolescente entre 7,4 y 22,2 puntos porcentuales (entre el 42 y el 138%). En general, también se asocia con otros resultados negativos, como un menor rendimiento educativo, la aceptación de conductas sexuales de riesgo y los roles de género sumisos en las relaciones sexuales.

KEYWORDS

maternidad adolescente, embarazo adolescente, transmisión intergeneracional, información no disponible no ignorable, métodos econométricos para tratamiento de información no disponible, Encuestas de Demografía y Salud

1 | INTRODUCTION

The Latin America and the Caribbean region, LAC, has the second highest rate of adolescent fertility in the world, only second to Sub-Saharan Africa. In 2010-2015, for example, the region had 66.5 births per 1000 girls aged 15-19, compared to only 30.0 in the US and 11.3 in Canada (Caffe et al. 2016). More strikingly, these differences have been increasing in relative terms since at least the beginning of the 1980s. Understanding the causes of such a persistent phenomenon in the LAC region is of essence since teen childbearing has been linked to negative socioeconomic and health outcomes of teen mothers and their children.

We estimate the link between a mother's teenage childbearing and the probability of her daughter's teenage childbearing using Demographic and Health Survey data (DHS) for six countries in the LAC region: Bolivia, Colombia, Dominican Republic, Guatemala, Haiti and Peru.[1] A salient characteristic of these type of data is the relative high percentage of mother-daughter pairs that are incomplete i.e. for whom there is either information on the daughter or on her mother, but not both. Indeed in our data, only 52% of the mother-daughter pairs are observed or interviewed. Missing mother-daughter matches are mostly caused by mother and daughters not sharing the same household. We show that restricting the estimating sample to mother-daughter matches in the data leads to a large negative coresidential bias in the estimated effect because missing matches are non-random and are, potentially, affected by the teen childbearing status of mothers and daughters. Indeed, daughters who cannot be matched to their mothers are between 2 to 5 times more likely to be teen mothers, which suggests that the missing information of the mother is not random and, therefore, missingness is non-ignorable. We avoid this bias by developing a methodology that allows missing observations to be endogenous and uses all available data, including incomplete mother-daughter pairs. We complete this analysis by estimating a similar model for early sexual behavior, which is a much more prevalent phenomenon than teen childbearing. We also study the associations between having a teen mother and other outcomes such as educational achievement, knowledge of reproductive health, vulnerability to risky sexual behavior; and fertility preferences.

Our work is related to two different, although connected, literatures. From the point of view of the mothers in our sample, it relates to the extensive literature on the consequences of being a teen mother. Recent studies show that observed differences in outcomes between teen mothers and their peers arise mostly due to selection into teen motherhood and that causal effects of teen childbearing on education and wages, although negative, are relatively modest (Ashcraft et al. 2013, Hotz et al. 2005, Fletcher and Wolfe 2009). If that was the case, does it matter whether there is inertia from mothers to daughters? There are at least two reasons why it could matter. First, conclusions of modest effects of teen childbearing were drawn from studies using contemporaneous US and European data, countries where teen pregnancy rates are relatively low, contraception is widespread, abortion is legal under general conditions, and public programs attend to teen mothers and their babies. Should we expect these results to apply to LAC countries, where abortion is either very restricted or banned altogether (Guttmacher Institute, 2018), there is no easy access to contraception, and no public safety net? Historical evidence for the period 1940-1968 in the US shows that, under similar conditions to those in contemporaneous LAC, teenage childbearing had large negative effects on teen mothers (Lang and Weinstein 2015). Some studies using

[1] Throughout the paper, we define the teen childbearing status of the mother from her age when she gave birth to her first live child(ren), i.e. we classify a mother as a teen mother if the eldest of her children was born when she was a teenager. Other papers in the literature aim to identify the effects of the teen mother on her first child(ren) and, hence, compare outcomes of siblings born when the mother was a teenager to those born when the mother was an adult woman.

Mexican and Chilean data also find similar strong negative results (e.g. Arceo-Gómez and Campos-Vazquez 2014, Kruger and Berthelon, 2012).[2] Second, although scarce, there is some evidence that teen childbearing may have long-term effects on the child. Francesconi (2008), for example, uses differences between siblings in mother's age at birth to identify the effect of teen motherhood on young adults, and finds strong negative results on school attainment and wages using the British Household Panel.[3]

From the point of view of the daughters in our sample, the paper is related to the literature on the determinants of teen pregnancy (e.g. Carneiro et al., 2021; Chetty et al. 2011; Black et al., 2008). In contrast to the literature on the consequences of childbearing, this literature includes plenty of studies in Latin America, which reveal associations between poverty, poor family structure, poor family background, low educational inputs, low aspirational objectives, low sexual literacy, poor neighborhood, high levels of violence, and teen pregnancy and childbearing status (Estrada et al., 2021; Aguia-Rojas et al., 2020; Alzate et al., 2020; Drewry and Garcés-Palacio, 2020; Tsaneva and Gunes, 2020; Dongarwar and Salihu, 2019; Millán, 2019; Mohr et al., 2019; Neal et al., 2018; Alzate, 2014).

Finally, a few papers, including ours, lie in the intersection of both literatures and aim to estimate the intergenerational transmission of teenage childbearing from mothers to daughters. More broadly, these papers fall into the strand of the literature on intergenerational transmission beyond social status, earnings, and education (Black and Devereux 2011). Early evidence of mother-daughter transmission of teenage childbearing is found on survey data for the US (e.g. Card, 1981, and Kahn and Anderson, 1992). Later studies, using more sophisticated econometrics methods, also find that teens born to teen mothers are more at risk of pregnancy.[4] Francesconi (2008), for example, uses differences across sisters in mother's age at birth to account for family characteristics and, hence, control for unobservable drivers. His study finds that being born to a teen mother at time of their birth increases the probability of being a teen mother by 135%.

Survey data, however, may suffer from a sample selection problem caused by teen childbearing—the event under study—which affects the creation and structure of families by increasing the probability of marriage and the abandonment of the parental home. Consider, for example, a household survey where information is gathered only for the individuals living at the household. Teenage daughters who no longer live with their parents usually appear as household heads or spouses and information on their mothers is missing. Similarly, the daughter information is also missing in the households of mothers whose daughters have already left their parents home. Therefore, these two groups of women are not present when the analysis is carried out using only the matched mother-daughter pairs, i.e., the pairs of mothers and daughters living in the same household. Studies that focus on the intergenerational transmission of teen motherhood are subject to, at least, two sources of selection bias. The bias that arises from unobserved factors affecting both the likelihood of being a teen mother and her daughter's outcomes—teen childbearing included—, and the bias that arises from non-ignorable missing pairs, which occurs when restricting the estimating sample to the matched mother-daughter pairs. To our knowledge,

[2] On the contrary, Azevedo et al. (2012b), which also uses Mexican data but identifies the effect of childbearing using miscarriages as a natural experiment, finds non-negative and even positive results, which are in line with other results using this methodology on US data (e.g. Hotz et al. 2005).

[3] The medical literature has extensively documented negative health outcomes for the mother (e.g. labor complications) and the child (e.g. low birth weight) related to teen labor. Although the amount of evidence seems to indicate an effect, causality is not formally established. Azevedo et al. (2012a) is one example from the non-medical literature also showing supportive evidence of greater maternal mortality and labor related complications for teen births.

[4] Other studies find negative effects on behaviors that increase the likelihood of teen pregnancy, such as early sexual activity and risky sexual behavior (Levine et al. 2001).

the literature has focused exclusively on the first source of bias.[5] In this paper, we develop an estimation approach that corrects the selection bias arising from non-ignorable missing mother-daughter pairs which is, as we show, of first order of magnitude when using data from LAC region.

Longitudinal surveys, such as the *National Longitudinal Survey of Youth* (NLSY) or the *Panel Survey of Income Dynamics* (PSID), in which the children of the original respondents are followed regardless of their residence, or administrative data, such as the one used in Aizer et al. (2020) for Norway,[6] offer good alternatives to survey data. The former, however, usually suffer attrition, which among other factors is, again, likely affected by teen childbearing whereas the latter avoids this problem but is only available in those countries where teenage pregnancy and childbearing have very low incidence. Comparable longitudinal or administrative data are not available for countries in the LAC region.

In summary, we develop a ML procedure that uses matched and unmatched mother-daughter pairs to estimate the relation of mothers' teenage childbearing status at first birth (TCS) with that of their daughters' free of coresidential bias. Using detailed comparable individual level data for six LAC countries, we show the existence of a very high intergenerational transmission of TCS; Estimated average marginal effects predict that daughters of teen mothers are between 7.4 and 22.2 percentage points (or between 42 and 138%) more likely to be teen mothers themselves. Our results also reveal that estimates based on matched mother-daughter pairs alone are severely negatively biased because the prevalence of teenage childbearing among unmatched teenage girls, i.e. those who do not cohabit with their mothers, is 2 to 5 times higher than among matched girls. We estimate a similar model for early sexual initiation and find congruent results. Having a mother who had an early initiation to sex increases the probability of early sexual behavior among teens between 20.6 and 27.2 percentage points (or between 61.8 and 104.7%). Both sets of results lead us to conclude that sexual behavior, either initiation of sexual activity or its consequences in the form of teen childbearing, is prone to high levels of intergenerational inertia from mothers to daughters. We also study the relationship between mothers' TCS and other outcomes of the daughter, such as educational achievement, submissive gender roles in sexual relationships, knowledge of reproductive health, and preferences regarding number of children. We find that having a teen mother is associated with: lower educational achievement in the non Andean countries (between 45 and 65%); increased knowledge of reproductive health (between 59% and 201%); accepting risky sexual behavior from partners in four of the six countries (between 25% and 164%); a preference for larger families in four of the six countries (between 34% and 163%). Results from these exercises suggest that high teen pregnancy rates amongst daughters of teen mothers may in part be explained by submissive gender roles that lead to low rates of condom use and preference for larger families. Surprisingly, being the daughter of a teen mother increased educational achievement in Colombia and Peru, the two countries in our sample with the highest educational achievement among teenage girls. We speculate the positive effect is founded on a social support network for teen mothers and their children together with widespread rates of primary education, which brings a message of hope to the dim picture of teenage childbearing in the LAC region.

The remainder of this paper is structured as follows. We present the data in Section 2, the econometric model and the empirical strategy in Section 3, and the results in Section 4.

[5] Francesconi (2008) reports that as much as twenty-seven percent of young adults in his sample were not matched to their siblings because the siblings were not been interviewed, but does not discuss whether these missing pairs are random and, hence, ignorable.

[6] Aizer et al. (2020) uses variance in sister's age at first birth collected from administrative data to establish that those children whose mother was a teenager are 3 percentage points or 32% more likely to become teen mothers.

Section 5 concludes. Appendix A presents additional tables and results.

2 | DATA

We use comparable individual-level data from the standard *Demographic and Health Surveys* (DHS) for all Latin American countries for which the available DHS allow us to obtain the teen childbearing rates (TCR) for mothers' teenage years. The selected countries are: Bolivia, Colombia, Dominican Republic, Guatemala, Haiti, and Peru.

DHS are nationally-representative household surveys comprising independent cross-sections conducted at about every five years. An important feature of the DHS is that all women in the household aged between 15 and 49 answer a specific questionnaire that provides, among other items, birth information—such as the birth date—for all of their live children. Therefore, it is possible to count the number of women who gave birth during their adolescence and compute consistent and homogeneous annual TCRs statistics by country.

2.1 | Early and sample DHSs

Our empirical strategy requires external information on the mother's cohort-specific TCRs in the estimating sample. Lacking an external source of a long and homogeneous series of TCRs, we compute them ourselves using DHSs. To do so, we partition all available DHSs between those that we use to estimate the model—the "sample" surveys—and those that can only be used in the computation of cohort-specific TCRs—the "early" surveys.

To illustrate how we distinguish early from sample surveys, consider the case of Colombia, a country that participates in 1986 (wave I), in 1990, and, then subsequently every five years until 2015. As the vast majority of girls who become mothers do so after their tenth birthday, we consider as teens those girls aged between 10 and 18. We define the TCR in a given year as the proportion of teens in that year who eventually become teen mothers. We compute retrospective TCRs using the available birth histories of interviewed women (i.e. women aged between 15 and 49). How far back we go depends on the first survey available. For Colombia, 1955 is the year when a woman who is 49 in the first survey (1986) was 18 years old. Hence, 1955=1986-49+18 is the first year when TCRs can be computed.

We cannot use all available surveys for estimation. To understand why, let us consider again the Colombian case. For estimation purposes, we group mothers' age in five-year intervals (and, consequently, also compute five-year averages of annual TCRs). Mothers of teen daughters are most likely aged 25 to 64. Hence, the eldest mothers in the 1986 Colombia DHS, i.e. women who are between 60 and 64, were aged 10 in between 1951 and 1955. Hence, the first Colombia survey that we can use for estimation is the one in which TCRs for these women are available, i.e., 1955. Therefore, when women aged 10 in 1955 become 60 is the first Colombian survey year we can use for our estimation: 2005=1955+60-10. We cannot use the 2000 Colombian DHS because we would need TCRs for women aged 60-64, i.e., TCRs from the interval 1946-1950. Yet, if we were to use external information from other sources (such as official statistics on age-specific fertility rates), the estimation sample would be even more constrained as these statistics start in the 1970s at the earliest.

TABLE 1 Available Standard Demographic Health Surveys

Country	Early DHS	Sample DHS
Bolivia	1989, 1994 1998, 2003	2008
Colombia	1986, 1990 1995, 2000	2005, 2010, 2015
Dominican Republic	1986, 1991, 1996, 1999, 2002	2007, 2013
Guatemala	1987, 1995 1998-99	2014-15
Haiti	1994-95, 2000 2005-06, 2012	2016-17
Peru	1986, 1991-92, 1996 2000, 2004-06	2007-08:2012

Notes: For each country, we distinguish early Demographic Health Surveys (DHSs) (under the heading "Early DHS") from later DHSs (under the heading "Sample DHS"). Sample surveys make up our estimation sample. Early surveys comprise all available DHSs prior to the first sample DHS. For each sample DHS we compute retrospective TPRs using all available DHSs but that one.

In Table 1 we summarize all available Standard DHSs for each of the six countries, distinguishing sample surveys that form the estimation sample from early surveys that cannot be included in the estimation sample.

2.2 | Annual Teen Childbearing Rates

The top panel of Figure 1 plots the evolution from 1955 to 2008 of annual TCRs by country using the birth history information from all available DHS surveys.

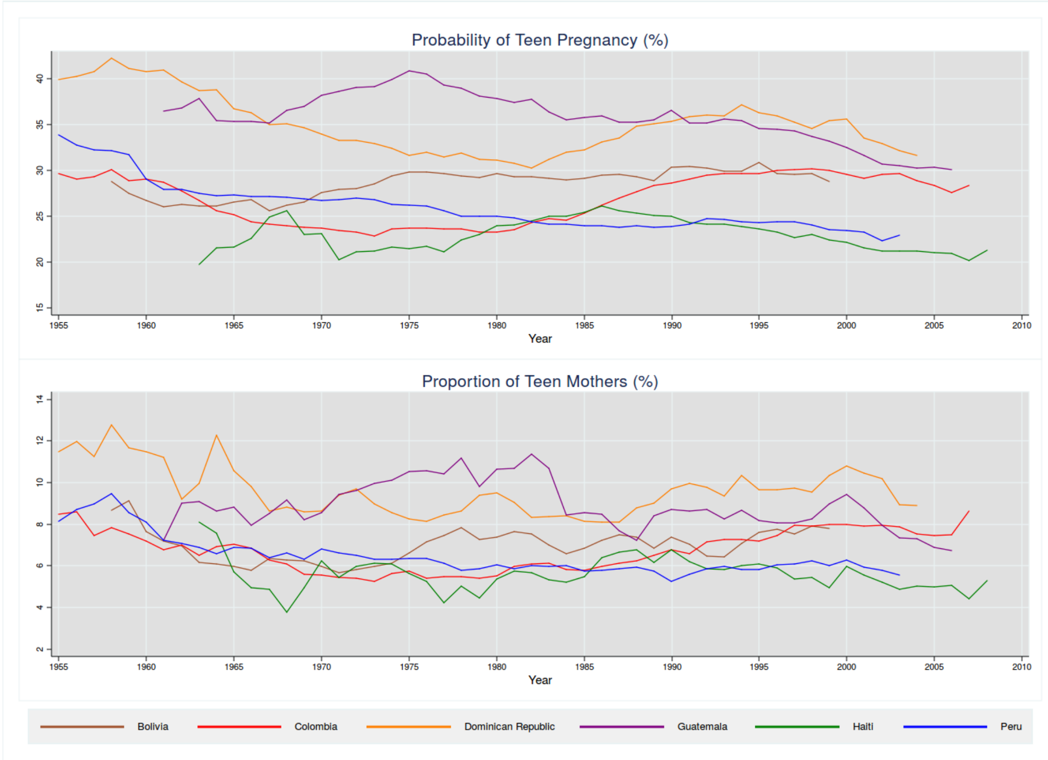

FIGURE 1 Teen Motherhood.

Notes: The countries and years contemplated were: (i) Bolivia: 1989, 1994, 1998, 2003, and 2008; (ii) Colombia: 1986, 1990, 1995, 2000, 2005, 2010 and 2015; (iii) Dominican Republic: 1986, 1991, 1996, 1999, 2002, 2007, and 2013; (iv) Guatemala: 1987, 1995, 1998-99, and 2014-15; (v) Haiti: 1994-95, 2000, 2005-06, 2012, and 2016-17; and (; and (vi)vi) Peru: 1986, 1991-92, 1996, 2000, 2004-06, 2007-08, 2009, 2010, 2011, and 2012. The Probability of Teen Pregnancy (top panel) shows, for each year, the percentage of girls aged 10 to 18 who become mothers before age 18. The Proportion of Teen Mothers (bottom panel) shows the percentage of girls aged 10 to 18 who are mothers.

Source: Own elaboration using birth histories from all available *Standard Demographic Health Surveys* (DHS).

The general trend is one of either stagnation or decrease. However, some countries suffer large increases for long periods—for example, Colombia and the Dominican Republic between 1980 and 1995. We also show in the bottom panel of Figure 1 the percentage of girls aged 10 to 18 who are mothers in a given year. TCRs correlate very strongly with the proportion of teen mothers in the bottom panel ($\rho = 0.899$) but, naturally, TCRs are higher and less volatile as they measure the cumulative probability of having a child throughout the teen years. Opposing trends between the top and bottom panels (for example, Haiti between 1963 and 1968) reflect important changes in the age at which adolescent mothers have their first child.

TCRs shown in Figure 1 cannot be used in the estimation because they are computed with all available observations, including those of the mothers of the sample surveys. Hence, they do not constitute external information. The solution is, for each DHS sample, to calculate retrospective TCRs using all the other DHS available. For example, for the Colombia 2005 sample, we compute cohort-specific TCRs with all other DHSs, including not only Colombia's early DHSs but also Colombia 2010 and 2015.

2.3 | Estimation Sample

Unique household and individual identifiers in each sample DHS permit the matching of teenage daughters aged 15 to 18 to their mothers aged 49 or younger living in the same household. However, some interviewed daughters cannot be matched with their mothers because either (i) their mothers are older than 49 or deceased, or (ii) they live in a different household. Similarly, some interviewed mothers cannot be matched to their teenage daughters because (i) the daughter was not interviewed (exceptional), or (ii) she does not live with her mother.[7]

TABLE 2 Sample DHSs: Sizes and missing observations

	Number of observations	Both observed (%)	Mother missing (%)	Daughter missing (%)
Bolivia 2008	2982	46.38	50.34	3.29
Colombia 2005	6101	51.60	43.62	4.79
Colombia 2010	7906	54.91	41.17	3.92
Colombia 2015	5716	50.59	44.16	5.25
Dominican Republic 2007	5112	46.19	48.92	4.89
Dominican Republic 2013	1532	46.02	48.96	5.03
Guatemala 2015	4845	56.57	40.72	2.70
Haiti 2017	2789	39.30	59.59	1.11
Peru 2008	5771	53.13	42.87	4.00
Peru 2009	3992	55.69	41.31	3.01
Peru 2010	3776	52.60	43.67	3.73
Peru 2011	3654	53.75	42.36	3.89
Peru 2012	3797	52.88	42.56	4.56
All sample DHSs	57973	51.60	44.44	3.96

Notes: DHS samples include all matched and unmatched pairs of mothers and their teenage daughters aged 15 to 18. "Number of observations" refers to the total number of pairs by country and wave. The proportion of pairs of mothers, aged 19 to 49, and their teenage daughters who are both interviewed is shown under the heading "Both observed (%)". Column "Mothers missing (%)" shows the percentage of mother-daughter pairs where the mother is not interviewed and, hence, cannot be matched to her daughter. Column "Daughters missing (%)" shows the percentage of mother-daughter pairs where the daughter is not interviewed.

In Table 2, DHS samples sizes include all matched and unmatched pairs of mothers and their teenage daughters. The number of observations refers to the total number of

[7] We consider only women whose teenage daughters are alive. There is a potential bias if the daughter's death are related to labor or terminations. If it exists, this bias must be negligible because the proportion of women aged 25-49 who had a daughter who died in adolescence in the original data is only 0.37 percent.

mother-daughter pairs by country and wave. These totals vary mostly by country and are fairly stable by year. Overall, our sample consists of 57,973 observations. Mother-daughter matches (reported as percentage under the heading "Both observed (%)") represent on average 51.6 percent of the observations and, with the exception of Haiti 2017, they vary little across countries (ranging from 46.02 to 56.69 percent). Pairs where the mother is not interviewed are also common, on average 44.44 percent of all pairs. It is the most common situation only in Haiti 2017 (59.59 percent). Lastly, although the proportion of pairs in which the daughter is not interviewed is residual, ranging from 1.11 percent in Haiti 2017 to 5.25 in Colombia 2015, we do include them in our econometric approach.

Each annual survey provides demographic characteristics for every member of the household—such as current age, gender, education, and relation to the household head—and basic information on the characteristics of the household—such as its regional location and whether it is located in a rural or urban area. In addition, the DHS contains information to compute teen childbearing status of interviewed mothers and daughters and detailed information on i) sexual behavior, such as age at first sexual intercourse, number of sex partners, use of contraceptives; ii) personal beliefs regarding gender roles in sexual relations; and iii) knowledge of contraceptive health, such as fertility status along the menstrual cycle, contraceptives, and sexually transmitted diseases.

Importantly, abortion is very restricted or banned altogether in our estimation sample.[8] Research has shown that women that undertake abortion are of higher socioeconomic stratus than other pregnant women (Fletcher and Wolfe, 2009; Ashcraft et al., 2013) and, hence, high abortion rates could compromise the representativeness of our data. In the DHS surveys women are only asked whether they ever had a pregnancy that terminated in a miscarriage, abortion, or still birth. Thus, we can neither differentiate between abortion and the other reasons for terminations nor count the actual number of terminations per woman (nevertheless, we find it unlikely that the average number of terminations during teenage years differs much from the proportion of teenagers with at least one termination). In Panel A of Table A.1 we show that the percent of women that had at least one termination before age 19 is between 1.7 and 5.0% (between 2.1 and 6.2% among women who were sexually active). As a percentage of pregnant women, women with at least one termination lie between 7.4 and 14.4%.[9] Lang and Weinstein (2015) consider that in their US sample for the years 1948-1968, the percent of miscarriages among teens was around 6% and the percentage of abortions was not higher than 3%. Based on these figures, they discard the possibility that abortion rates bias their estimates. Our termination rates lead us to think that abortion rates in our sample are not higher than those reported in Lang and Weinstein

[8] Abortion is prohibited altogether (no explicit legal exception) for Dominican Republic, Haiti, and Colombia 2005. Allowed only if to save the life of the mother in Guatemala. It is allowed: (*i*) to save the life of the mother and preserve physical health in Bolivia and Peru; and (*ii*) to save the life of the mother and preserve physical and mental health in Colombia 2010 and 2015. Finally, exceptions such as rape or incest are allowed in Bolivia and Colombia 2010 and 2015.

[9] In our sample of teenage daughters, those who state having had a termination in the past are either mothers or currently pregnant. Hence, there is no difference between teen childbearing and teen pregnancy status. There is, however, a legitimate concern regarding under-reporting of abortions in LAC as most abortion cases are not allowed. There are reasons to think that this problem is minor in our sample: (i) As women are only asked about terminations, they may safely include an illegal abortion in their answers; (ii) termination rates during adolescence reported by women aged 23—shown in Panel A in Table A.1—are comparable with miscarriage rates reported elsewhere (Lang and Weinstein, 2015, Lang and Nuevo-Chiquero, 2012), (iii) they are smaller for no-teen mothers than for teen_mothers—which may justify why we find no difference between teen childbearing and teen pregnancy status in our sample of teenage daughters—, and they are lower than for women aged 19 to 23, suggesting that these terminations are mostly miscarriages (see Panel B in Table A.1); finally (iv) we checked that the presence of adults during the interview did not affect the rate of reported terminations.

(2015) and, hence, should not seriously bias our estimates.

3 | EMPIRICAL STRATEGY

3.1 | Non-ignorable missing observations

In Figure 2 we show the probability of teen childbearing by TCS of the mother for the six countries in our data. Figure 2 suggests the existence of a substantial TCS mother effect. The unconditional probability of the daughter being a teen mother increases from 2 to 6 percentage points (a 33 to 120 percent increase) when they are children of teen mothers.

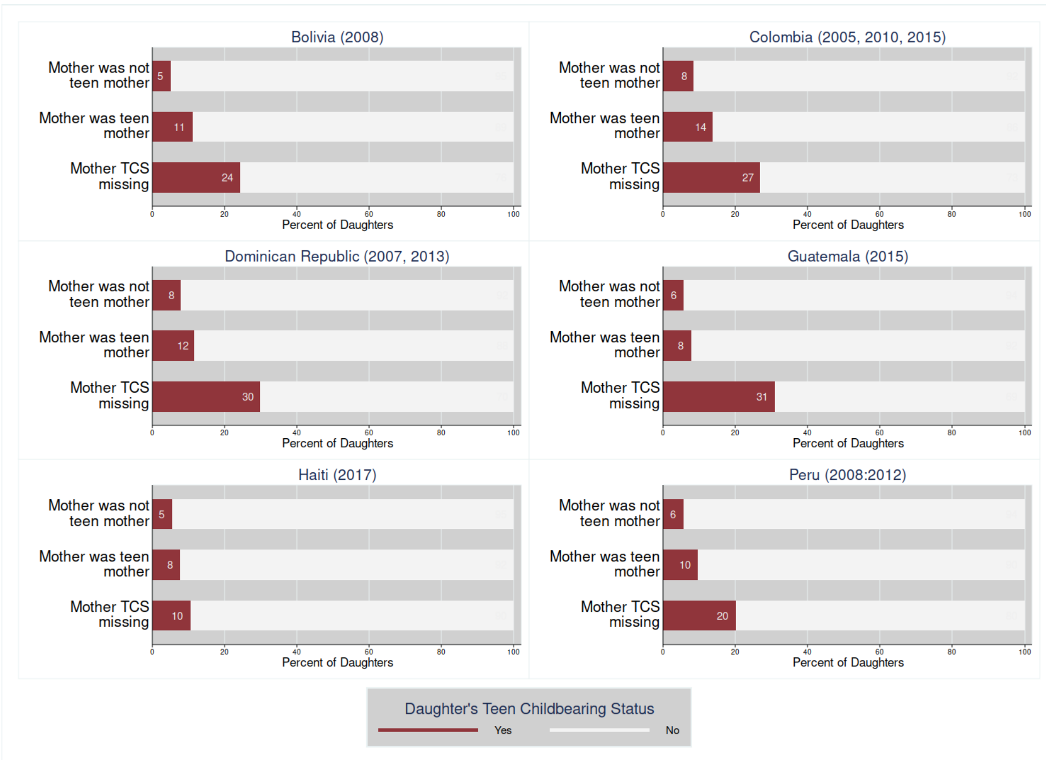

FIGURE 2 Daughter's Teen Pregnancy Status by her Mother's.

Notes: The countries and years contemplated were: (i) Bolivia: 2008; (ii) Colombia: 2005, 2010 and 2015; (iii) Dominican Republic: 2007, and 2013; (iv) Guatemala: 2014-15; (v) Haiti: 2016-17; and Peru: 2007-08, 2009, 2010, 2011, and 2012.

Source: Own elaboration using birth histories from sample *Standard Demographic Health Surveys* (DHS).

The figure also shows that unmatched daughters, i.e. those daughters whose mother does not live in the household and, hence, their mother's TCS is missing, are between 2 to 5 times more likely to be teen mothers than matched daughters. In fact, the majority of unmatched daughters who are mothers themselves abandoned their parents' home after conceiving and are currently married/cohabiting with a partner. That is to say, the missing information for the mother is a consequence of the daughter's TCS. This suggests that the missing information of the mother is not random and, therefore, missingness is non-ignorable.

The standard approach followed in the literature would be to ignore missing observations. But statistical analysis using only observations without missing values when the missing process is not random suffers from sample selection bias. Instead, we will deal with non-ignorability by developing a ML procedure in the spirit of the GMM approach first proposed by Ramalho and Smith (2013). This methodology involves Maximum Likelihood estimation using all observations under the assumption that missing values are generated by the same missing process and allowing the missing values to be endogenous.

In the empirical model, TCS of the daughter and her mother as well as observability of their TCS are dummy variables taking values 0 and 1. Identification follows from the assumption that observability of daughter and mother depends on the value of their TCSs. The procedure also allows for a test of ignorability. We improve the identification of the parameters of the model by incorporating external information regarding teen childbearing rates (TCRs) from independent external sources.[10]

3.2 | The econometric model

We model a daughter's TCS as dependent on her mother's TCS. Daughter i is a teen mother if $y_i = 1$, otherwise $y_i = 0$. We assume that the discrete choice is expressed in the following linear specification:

$$y_i = \mathbf{1}\{\alpha y_i^m + x_i\beta + x_i^m\beta^m + z_i\gamma + \epsilon_i > 0\} \tag{1}$$

where dummy variable y_i^m indicates the TCS of the mother.

Control vectors x_i, x_i^m, and z_i are discrete, which is the case in our data. Vector x_i includes the daughter's characteristics; all its components are missing when the daughter is not interviewed. Similarly, x_i^m includes mother's characteristics; all its components (together with y_i^m) are missing when the mother is not interviewed. Note that some controls, such as country and year of interview, are always observable. We denote these controls by vector z_i.

Let us define a binary indicator I_i, which takes value 1 if the daughter is interviewed and 0 otherwise. Similarly, let I_i^m take value 1 if the mother is interviewed and 0 otherwise. The aim is to estimate parameter vector $\theta \equiv \{\alpha, \beta, \beta^m, \gamma\}$ where:

$$\Pr\{y_i | y_i^m, x_i, x_i^m, z_i\} = F(y_i, y_i^m, x_i, x_i^m, z_i; \theta). \tag{2}$$

Assuming normality we have the conditional probit model:

$$F(y_i, y_i^m, x_i, x_i^m, z_i; \theta) \equiv \begin{cases} \Phi\left(\alpha y_i^m + x_i\beta + x_i^m\beta^m + z_i\gamma\right) & \text{if } y_i = 1 \\ 1 - \Phi\left(\alpha y_i^m + x_i\beta + x_i^m\beta^m + z_i\gamma\right) & \text{otherwise.} \end{cases}$$

For an observation with non-missing information, the joint probability of non-missingness, i.e., $I_i = I_i^m = 1$, and the vector variables $\{y_i, y_i^m, x_i, x_i^m, z_i\}$ is:

[10] Our methodological approach is closest to Carro, Machado and Mora (2021), which estimates a mother-daughter transmission model of female labor force participation using historical data. The most important difference between the two models is that, contrary to their's, in our model when the TCS of the mother is not observed, none of her characteristics are observed because she was not interviewed. Similarly, when the TCS of the daughter is unobserved, none of her characteristics are observed. This feature of our data is likely to be more relevant for other applications.

$$\Pr\left\{I_i = I_i^m = 1, y_i, y_i^m, x_i, x_i^m, z_i\right\} = \Pr\left\{I_i = I_i^m = 1 \left| y_i, y_i^m, x_i, x_i^m, z_i \right.\right\} \times F\left(y_i, y_i^m, x, x_i^m, z_i; \theta\right) \times \Pr\left\{y_i^m, x_i, x_i^m, z_i\right\}. \tag{3}$$

There are two situations in which a given observation may have missing information: when the mother's information is missing but the daughter's is not and when the daughter's information is missing but the mother's is not. Consider the second case. The joint probability for observation $\left\{I_i = 0, I_i^m = 1, y_i^m, x_i^m, z_i\right\}$ decomposes into several event probabilities:

$$\Pr\left\{I_i = 0, I_i^m = 1, y_i^m, x_i^m, z_i\right\} = \sum_{\{y,x\}} \Pr\left\{I_i = 0, I_i^m = 1, y_i = y, y_i^m, x_i = x, x_i^m, z_i\right\}. \tag{4}$$

The treatment of the first case, i.e., when the mother's information is missing but the daughter's is not, is similar:

$$\Pr\left\{I_i = 1, I_i^m = 0, y_i, x_i, z_i\right\} = \sum_{\{y^m, x^m\}} \Pr\left\{I_i = 1, I_i^m = 0, y_i, y_i^m = y^m, x_i, x_i^m = x^m, z_i\right\}. \tag{5}$$

Define p_i as the probability of observation i. Then:

$$\begin{aligned} p_i = \; & \left(\Pr\left\{I_i = I_i^m = 1, y_i, y_i^m, x_i, x_i^m, z_i\right\}\right)^{I_i I_i^m} \times \\ & \left(\Pr\left\{I_i = 0, I_i^m = 1, y_i^m, x_i^m, z_i\right\}\right)^{(1-I_i) I_i^m} \times \\ & \left(\Pr\left\{I_i = 1, I_i^m = 0, y_i, x_i, z_i\right\}\right)^{I_i (1-I_i^m)} \end{aligned} \tag{6}$$

The previous considerations warrant the following:

Assumption 1 *(Response Conditional Independence, RCI) Non-response in y_i and y_i^m is conditionally independent of x_i, x_i^m, and z_i; i.e.,*

$$\Pr\{I_i, I_i^m | y_i, y_i^m, x_i, x_i^m, z_i\} = \Pr\{I_i, I_i^m | y_i, y_i^m\}. \tag{7}$$

Assumption 1 is a sufficient condition to identify the parameter vector θ. Let $H_{y_i y_i^m}^{I_i I_i^m} \equiv \Pr\left\{I_i, I_i^m \left| y_i, y_i^m \right.\right\}$, with $I_i, I_i^m, y_i,$ and $y_i^m \in \{0,1\}$, and $\Pi_{y_i^m, x_i, x_i^m, z_i} \equiv \Pr\left\{y_i^m, x_i, x_i^m, z_i\right\}$, where $\Pi_{y_i^m, x_i, x_i^m, z_i} \in [0,1)$. For observations with information for mother and daughter, assumption 1 implies that:

$$\Pr\{I_i = I_i^m = 1, y_i, y_i^m, x_i, x_i^m, z_i\} = H_{y_i y_i^m}^{11} F\{y_i, y_i^m, x_i, x_i^m, z_i; \theta\} \Pi_{y_i^m, x_i, x_i^m, z_i}. \tag{8}$$

When only the daughter's information is observed, the joint probability is:

$$\Pr\{I_i = 1, I_i^m = 0, y_i, x_i, z_i\} = \sum_{y^m, x^m} \left(H^{10}_{y_i y^m} F\{y_i, y^m, x_i, x^m, z_i; \theta\} \Pi_{y^m, x_i, x^m, z_i} \right)$$
$$= \sum_{y^m} \left(H^{10}_{y_i y^m} \sum_{x^m} \left(F\{y_i, y^m, x_i, x^m, z_i; \theta\} \Pi_{y^m, x_i, x^m, z_i} \right) \right), \tag{9}$$

where $F\{y_i, y^m, x_i, x^m; \theta\}$ and Π_{y^m, x_i, x^m, z_i} are evaluated at values y_i, x_i, and z_i and all potential combinations of running values y^m and x^m.

Finally, the joint probability of an observation without daughter's information is:

$$\Pr\{I_i = 0, I_i^m = 1, y_i^m, x_i^m, z_i\} = \sum_{y} \left(H^{01}_{y y_i^m} \sum_{x} \left(F\{y, y_i^m, x, x_i^m, z_i; \theta\} \Pi_{y_i^m, x, x_i^m, z_i} \right) \right) \tag{10}$$

The model parameters are θ, conditional missing process parameters $\left\{ H^{II^m}_{yy^m} \right\}$ and marginal probabilities $\{\Pi_{y^m, x, x^m, z}\}$. From Equation (6), the conditional likelihood $\mathcal{L}_i$ for any given observation i is:

$$\begin{aligned} \mathcal{L}_i = \; & \left(H^{11}_{y_i y_i^m} F\left\{ y_i, y_i^m, x_i, x_i^m, z_i; \theta \right\} \Pi_{y_i^m, x_i, x_i^m, z_i} \right)^{I_i I_i^m} \times \\ & \left(\sum_{y^m} \left(H^{10}_{y_i y^m} \sum_{x^m} \left(F\{y_i, y^m, x_i, x^m, z_i; \theta\} \Pi_{y^m, x_i, x^m, z_i} \right) \right) \right)^{I_i (1 - I_i^m)} \times \\ & \left(\sum_{y} \left(H^{01}_{y y_i^m} \sum_{x} \left(F\left\{ y, y_i^m, x, x_i^m, z_i; \theta \right\} \Pi_{y_i^m, x, x_i^m, z_i} \right) \right) \right)^{(1 - I_i) I_i^m} . \end{aligned} \tag{11}$$

The log-likelihood function results from the sum of the log of $\mathcal{L}_i$, $\log(\mathcal{L}) = \sum_{i=1}^{N} \log(\mathcal{L}_i)$ and is maximized subject to the following constraints:

$$H^{II^m}_{yy^m}, \Pi_{y^m, x, x^m, z} \in [0, 1) \;\text{ for all } I, I^m, y, y^m, x, x^m, z \tag{12}$$

$$H^{00}_{yy^m} = 0 \text{ for all } y, y^m \tag{13}$$

$$\sum_{I, I^m} H^{II^m}_{yy^m} = 1 \text{ for all } y, y^m \tag{14}$$

$$\sum_{y^m, x, x^m, z} \Pi_{y^m, x, x^m, z} = 1 \; . \tag{15}$$

Maximum Likelihood estimation will yield consistent and asymptotically efficient estimates of θ. The number of parameters in $\Pi_{y^m, x, x^m, z}$ grows exponentially with the number of controls and rapidly becomes computationally intractable. In our application, we may reduce the number of parameters $\Pi_{y^m,x,x^m,z} = \Pi_{y^m,x^m|x,z}\Pi_{x,z}$ by assuming that $\Pi_{y^m,x^m|x,z} = \Pi_{y^m,x^m|z}$, i.e., mother characteristics are predetermined with respect to those of the daughter. The number of parameters can be further reduced by using restrictions on $\Pi_{y^m,x^m|z}$ so that the probability only varies along two subsets $v \subseteq x^m$ and $w \subseteq z$, i.e. $\Pi_{y^m,x^m|z} = \Pi_{y^m,v(x^m)|w(z)}$. Altogether, $\Pi_{y^m,x,x^m,z} = \Pi_{y^m,v(x^m)|w(z)}\Pi_{x,z}$ does not increase with the number of mother's controls. Finally, sample identification can also be improved with the use of external information that eliminates the need to estimate these matrices. In our application, we take from external information the matrix $\Pi_{y^m,v(x^m)|w(z)}$.

4 | RESULTS

4.1 | Teen Childbearing

Table 3 shows coefficient estimates and the average marginal effect (AME) of the mother TCS status under ignorability, i.e using only observations where both the mother and the daughter are interviewed. In this basic specification, we control for a small set of dummy variables for rural residence, the age of the mother, and the age of the daughter, which have the expected sign and direction i.e. older age and rural residence are associated with higher probability of being a teen mother. The AMEs are positive and statistically significant for all countries, except Haiti. Their magnitudes imply that being a daughter of a teen mother increases her chances of being a teen mother herself between 2.1 to 5.2 percentage points, or between 33% and 69%.

TABLE 3 Probit estimates under ignorability. Observed pairs in Estimation Sample

	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
y^m	0.403***	0.306***	0.237***	0.201**	0.180	0.268***
	(0.109)	(0.036)	(0.067)	(0.079)	(0.128)	(0.038)
Dcohort1	0.069	0.009	0.082	0.008	0.037	-0.022
	(0.119)	(0.042)	(0.085)	(0.091)	(0.134)	(0.044)
Dage16	0.361**	0.314***	0.399***	0.384***	0.138	0.317***
	(0.171)	(0.054)	(0.111)	(0.132)	(0.216)	(0.059)
Dage17	0.687***	0.670***	0.721***	0.756***	0.586***	0.655***
	(0.165)	(0.052)	(0.107)	(0.127)	(0.196)	(0.057)
Dage18	0.817***	0.981***	1.136***	1.009***	0.951***	0.973***
	(0.170)	(0.052)	(0.105)	(0.127)	(0.190)	(0.058)
Drural	0.638***	0.180***	0.038	0.185**	0.148	0.253***
	(0.109)	(0.038)	(0.069)	(0.081)	(0.134)	(0.038)
AME(y^m)	0.051***	0.052***	0.037***	0.024**	0.021	0.034***
	(0.014)	(0.006)	(0.010)	(0.009)	(0.015)	(0.005)
$\bar{y}$	0.074	0.106	0.095	0.066	0.063	0.072
No. obs.	1383	10381	3066	2741	1096	11247

Notes: *** p<0.01, ** p<0.05, * p<0.1. Dependent variable is daughter's TCS. All models include survey dummies. Standard errors in parenthesis. Variable y^m is the dummy for the mother's teen childbearing status. Dcohort1 is a dummy variable for mother 45 and older. Dage16-Dage18 are age dummies (reference category is age 15). Drural takes value 1 when the household is located in a rural area. AME(y^m) is the estimated Average Marginal Effect of y^m. $\bar{y}$ is the average value of the daughter's teen childbearing status in the estimating sample.

We repeat the previous estimation but using our non-ignorable ML approach developed in Section 3.2.

TABLE 4 Non-ignorable ML estimates. Full Estimation Sample

	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
y^m	1.194***	0.473***	1.238***	1.305***	1.078***	0.519***
	(0.171)	(0.048)	(0.066)	(0.079)	(0.228)	(0.052)
Dcohort1	0.824***	0.381***	1.056***	1.223***	0.497***	0.432***
	(0.094)	(0.039)	(0.058)	(0.060)	(0.120)	(0.042)
Dage16	0.355***	0.400***	0.328***	0.332***	0.484***	0.358***
	(0.114)	(0.037)	(0.071)	(0.090)	(0.160)	(0.041)
Dage17	0.805***	0.782***	0.737***	0.850***	0.847***	0.815***
	(0.107)	(0.036)	(0.068)	(0.085)	(0.151)	(0.039)
Dage18	1.079***	1.130***	1.119***	1.066***	1.314***	1.153***
	(0.108)	(0.035)	(0.067)	(0.085)	(0.148)	(0.039)
Drural	0.489***	0.297***	0.136***	0.293***	0.171**	0.351***
	(0.068)	(0.025)	(0.045)	(0.056)	(0.084)	(0.025)
AME(y^m)	0.228***	0.108***	0.247***	0.182***	0.155***	0.095***
	(0.015)	(0.006)	(0.006)	(0.006)	(0.013)	(0.005)
$\bar{y}$	0.162	0.178	0.199	0.168	0.088	0.129
N. obs.	2982	19723	6644	4845	2789	20990
logL	-1.16e+04	-9.82e+04	-2.96e+04	-2.46e+04	-1.01e+04	-9.36e+04

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Dependent variable is daughter's TCS. All models include survey dummies. Standard errors in parenthesis. Variable y^m is the dummy for the mother's teen childbearing status. Dcohort1 is a dummy variable for mother 45 and older. Dage16-Dage18 are age dummies (reference category is age 15). Drural takes value 1 when the household is located in a rural area. AME(y^m) is the estimated Average Marginal Effect of y^m. $\bar{y}$ is the average value of the daughter's teen childbearing status in the estimating sample. logL stands for the loglikelihood function evaluated at the ML estimates.

Table 4 shows how the signs and direction of the estimates are similar to the ignorability estimates but now having an older mother increases significantly the probability of being a teen mother in all countries including Haiti. Moreover, the AMEs are now between 2.1 to 7.6 times larger. Thus, being a daughter of a teen mother increases the changes of being a teen mother between 9.5 percentage points in Peru to 24.7 percentage points in Dominican Republic. Comparing these AME estimates to those of Table 3, we conclude that ignoring the missing observations leads to a substantial negative selection bias in AMEs of the mother TCS.

TABLE 5 Non-ignorable ML estimates. Full Estimation Sample. Additional controls

	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
y^m	0.750***	0.364***	1.184***	1.460***	0.887***	0.493***
	(0.229)	(0.047)	(0.071)	(0.082)	(0.250)	(0.051)
Dcohort1	0.608***	0.228***	0.877***	1.228***	0.386***	0.272***
	(0.109)	(0.041)	(0.062)	(0.061)	(0.124)	(0.044)
Dprimary	0.158	-0.044	0.366***	0.085	0.317	-0.114**
	(0.152)	(0.038)	(0.076)	(0.125)	(0.194)	(0.047)
Dage16	0.384***	0.415***	0.340***	0.328***	0.475***	0.372***
	(0.114)	(0.038)	(0.074)	(0.092)	(0.157)	(0.043)
Dage17	0.880***	0.797***	0.726***	0.854***	0.817***	0.824***
	(0.109)	(0.037)	(0.070)	(0.087)	(0.150)	(0.041)
Dage18	1.101***	1.141***	1.146***	1.068***	1.300***	1.162***
	(0.110)	(0.036)	(0.070)	(0.087)	(0.148)	(0.041)
Drural	0.160*	0.192***	0.076	0.253***	0.097	0.310***
	(0.096)	(0.032)	(0.049)	(0.060)	(0.098)	(0.030)
Dhsize2	0.268**	0.608***	0.562***	0.118*	0.031	0.551***
	(0.115)	(0.036)	(0.063)	(0.069)	(0.124)	(0.039)
Dpoor	0.609***	0.183***	0.046	0.065	0.122	0.160***
	(0.150)	(0.045)	(0.067)	(0.071)	(0.146)	(0.054)
Dwork	-0.383***	-0.115***	0.017	0.118*	-0.122	-0.403***
	(0.113)	(0.036)	(0.061)	(0.067)	(0.126)	(0.040)
N. obs.	2979	19723	6635	4845	2789	20987
logL	-9905.745	-8.95e+04	-3.08e+04	-2.20e+04	-7732.049	-8.39e+04
AME(y^m)	0.127***	0.074***	0.222***	0.195***	0.121***	0.078***
	(0.019)	(0.005)	(0.006)	(0.006)	(0.014)	(0.005)
$\bar{y}$	0.162	0.178	0.199	0.168	0.088	0.129
AME under	0.042***	0.035***	0.015	0.018*	0.014	0.021***
ignorability	(0.014)	(0.006)	(0.011)	(0.010)	(0.015)	(0.005)

Notes: *** p<0.01, ** p<0.05, * p<0.1. Dependent variable is daughter's TCS. All models include survey dummies. Standard errors in parenthesis. Variable y^m is the dummy for the mother's teen chilbearing status. Dcohort1 is a dummy variable for mother 45 and older. Dprimary is a dummy variable that takes value 1 if the mother has no more than primary education. Dage16-Dage18 are age dummies (reference category is age 15). Drural takes value 1 when the household is located in a rural area. Dhsize2 is a dummy variable for parents households with at least seven members. Dpoor takes value 1 if the parents' household belongs to the two poorest quintiles of the country's population based on a continuous wealth measure produced by the *DHS*. Dummy Dwork takes value 1 if the mother works. logL stands for the loglikelihood function evaluated at the ML estimates. AME(y^m) is the estimated Average Marginal Effect of y^m. $\bar{y}$ is the average value of y in the estimating sample. AME under ignorability stands for the average marginal effect of y^m using the probit estimates under ignorability for the same variable specification.

Because mothers' TCS may be correlated with relevant omitted variables, we enrich our basic specification by including a few more mother characteristics as controls in our estimation, such as a dummy for whether the mother has no more than primary education (Dprimary), a dummy for whether the mother's household size was larger than 6 people (Dhsize2), a dummy for whether the mother's household belongs to the two poorest quintiles in the country (Dpoor), and a dummy for whether the mother works (Dwork).

As predicted, these variables are correlated with the mother's TCS status and, hence, the AME is now slightly lower than with the basic specification as can be seen in Table 5, but still very large and statistically significant. A poorer mother and a large household size increase the probability of teen age childbearing significantly in most countries. Having a mother who works, however, tends to decrease significantly the probability of TCS but this is only observed for the Andean countries. Having a mother with only primary education decreases the probability of TCS in Peru and Colombia, the two countries with highest share of mothers with more than primary education (only in Peru the estimate is significantly different from zero). Having a mother with only primary education increases the probability of TCS in all other countries (only in the Dominican Republic the estimate is significant).

Table 6 includes estimates related to the missing process as well as a likelihood ratio test for the null hypothesis that the missing process is ignorable, i.e. that the conditional probabilities $H^{1,1}_{y_i y_i^m}$=$\Pr(I=1, I^m=1|y, y^m)$ are invariant to daughter and mother TCS, y and y^m, respectively. The probability estimates of having a mother-daughter match vary considerably by y^m and y and we can strongly reject the null hypothesis of ignorability in all countries. As expected, the probability of a match is higher when daughters are not teen mothers, i.e. $y=0$, but, somewhat surprisingly, it increases when their mothers were teen mothers, i.e. $y^m=1$. The latter result may be the result of two features of our data: i) the sample of daughters are between the ages of 15 and 18; ii) mothers older than 49 are not interviewed. Notice that by selecting the sample of daughters to be between the ages of 15 and 18, their mothers, ceteris paribus, would be more likely observed if they are young mothers and they are more likely young if they were teen mothers.

TABLE 6 Missing process estimates & ignorability tests. Full Estimation Sample

	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
$\Pr(I=1, I^m=1\|y=0, y^m=0)$	0.417	0.456	0.392	0.477	0.323	0.455
$\Pr(I=1, I^m=1\|y=0, y^m=1)$	0.776	0.908	0.892	0.951	0.670	0.924
$\Pr(I=1, I^m=1\|y=1, y^m=0)$	0.177	0.232	0.326	0.418	0.357	0.233
$\Pr(I=1, I^m=1\|y=1, y^m=1)$	0.259	0.506	0.182	0.172	0.232	0.437
Ignorability tests	224.8	7193.7	2660.1	3458.0	76.1	8544.6
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: ML estimates for the probabilities of the missing process using the non-ignorable ML estimates with additional controls reported in Table 5. Ignorability tests are likelihood ratio tests for the null hypothesis that conditional probabilities $\Pr(I=1, I^m=1|y, y^m)$ are invariant to mother and daughter TCS, y^m and y, respectively. p-values in parenthesis.

4.2 | Early Sexual Behavior

We use our ML approach to estimate the effect of a mother sexual behavior during adolescence on the probability that her teenage daughter(s) had already had sex. Accordingly, the dependent variable y takes value one if the daughter reports to have had sex and zero otherwise and y^m takes value 1 if her mother had sex for the first time during adolescence, and zero otherwise.

Despite having higher sample means, the new definitions of y and y^m are related to teen childbearing and, hence, the results of the two models are very aligned and produce AMEs of the same order to magnitude. For example, having a mother with early initiation to sex increases the probability of having had sex between 20.6 and 27.2 percentage points (see Table 7), although these figures represent a lower effect in relative effect (between 61.8 and 104.7%) because early sexual behavior is more prevalent than teen childbearing in the sample. As in the case of teenage childbearing, the AMEs under ignorability (shown in the bottom of Table 7 for comparison purposes) are much lower although still significantly positive in all countries.

Results so far lead us to conclude that sexual behavior, either initiation of sexual activity or its consequences in the form of teen childbearing, has high intergenerational inertia.

TABLE 7 Sexual behavior: Ever had sex. Non-ignorable ML estimates.

	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
y^m	0.934***	0.847***	0.964***	1.965***	0.781***	1.054***
	(0.208)	(0.031)	(0.118)	(0.096)	(0.189)	(0.040)
Dcohort1	0.556***	0.206***	0.605***	1.302***	0.151*	0.317***
	(0.096)	(0.029)	(0.060)	(0.053)	(0.084)	(0.032)
Dprimary	0.040	-0.150***	0.183***	0.008	0.175*	-0.207***
	(0.114)	(0.027)	(0.061)	(0.094)	(0.106)	(0.033)
Dage16	0.381***	0.484***	0.404***	0.302***	0.445***	0.441***
	(0.094)	(0.029)	(0.057)	(0.076)	(0.082)	(0.034)
Dage17	0.833***	0.912***	0.805***	0.631***	0.843***	0.904***
	(0.090)	(0.029)	(0.056)	(0.074)	(0.082)	(0.034)
Dage18	1.131***	1.324***	1.170***	0.902***	1.297***	1.317***
	(0.092)	(0.030)	(0.058)	(0.075)	(0.084)	(0.034)
Drural	0.279***	0.012	0.065	0.121**	-0.053	0.150***
	(0.081)	(0.026)	(0.042)	(0.055)	(0.066)	(0.027)
Dhsize2	0.404***	0.311***	0.506***	-0.140**	-0.001	0.399***
	(0.091)	(0.027)	(0.056)	(0.055)	(0.081)	(0.029)
Dpoor	0.362***	0.041	0.048	0.058	-0.122	0.209***
	(0.116)	(0.030)	(0.057)	(0.058)	(0.096)	(0.037)
Dwork	-0.339***	-0.111***	0.071	0.081	-0.195**	-0.370***
	(0.092)	(0.026)	(0.052)	(0.054)	(0.084)	(0.030)
N. obs.	2925	19662	6494	4807	2786	20982
logL	-9721.052	-9.17e+04	-3.04e+04	-2.16e+04	-8426.477	-8.59e+04
AME(y^m)	0.206***	0.272***	0.255***	0.266***	0.228***	0.226***
	(0.018)	(0.005)	(0.012)	(0.005)	(0.019)	(0.004)
$\bar{y}$	0.267	0.440	0.352	0.254	0.356	0.244
AME under	0.087***	0.153***	0.079***	0.078***	0.072**	0.078***
ignorability	(0.023)	(0.010)	(0.017)	(0.016)	(0.034)	(0.008)

Notes: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable y is a dummy variable that takes value one if the daughter reports to have had sex and zero otherwise. All models include survey dummies. Standard errors in parenthesis. Variable y^m takes value one if the mother reports to have had sex before age 19 and zero otherwise. Dcohort1 is a dummy variable for mother 45 and older. Dprimary is a dummy variable that takes value 1 if the mother has no more than primary education. Dage16-Dage18 are age dummies (reference category is age 15). Drural takes value 1 when the household is located in a rural area. Dhsize2 is a dummy variable for parents households with at least seven members. Dpoor takes value 1 if the parents' household belongs to the two poorest quintiles of the country's population based on a continuous wealth measure produced by the *DHS*. Dummy Dwork takes value 1 if the mother works. logL stands for the loglikelihood function evaluated at the ML estimates. AME(y^m) is the estimated Average Marginal Effect of y^m. $\bar{y}$ is the average value of y in the estimating sample. AME under ignorability stands for the average marginal effect of y^m using the probit estimates under ignorability for the same variable specification.

4.3 | Other Outcomes

The literature on the effects of teen pregnancy on outcomes of the children is somewhat inconclusive, with some studies finding close to zero effects or effects that vanish in the relative short term (e.g. Angrist and Lavy, 1996) whereas others find large and significant effects (e.g. Francesconi, 2008). In this section, we look at the effects of having a teen mother on human capital formation, knowledge of contraceptive health, gender roles in relationships with sexual partners, and fertility preferences i.e. ideal number of children. All outcomes which are not naturally dummy variables, such as for example ideal number of children, must be transform into a dummy variable for our ML method to work. Although this restriction may be seen as a shortcoming of our procedure, we argue that this discretization is capable of capturing a lot of the mother-daughter inertia.

Importantly, applying our ML methodology to different outcomes implies assuming that the missing process is dependent on the new y and y^m variables (see 1). Whereas, that may be a reasonable assumption for some variables (for example, in the case of early initiation to sex shown in Section 4.2) it may be less convincing for other variables, such as, knowledge of the ovulatory cycle. Given this potential shortcoming, the results shown in Figure 3 should be taken with a grain of salt. We are particularly attentive to the ignorability tests that show us whether the probabilities of observing the mother-daughter match vary according to the values of y and y^m. For all outcomes, reassuringly we strongly reject the null that the conditional probabilities are independent of the new y and y^m (see Table 8).

When analyzing the effects of mother TCS on human capital, we use as dependent variable a dummy that takes value one if the daughter has obtained at most primary education and zero otherwise. Contrary to the previous outcomes, we now obtain a wide range of AMEs, from negative to zero effects in the Andean countries, to large and significantly positive effects in Dominican Republic, Guatemala and Haiti as shown in the top left graph of Figure 3 and in Table A.2 in the Appendix. The negative AMEs for Colombia and Peru, which indicate that teen mothers have positive effects on their daughters educational achievement, are likely the result of social support networks for teen mothers and their children together with widespread rates of primary education (only 12% of the teenagers have at most primary education compared to more than 30% in the other three countries). Positive AMEs, on the contrary, reveal TCS as a potential poverty trap to daughters of teen mothers. In the Dominican Republic and Guatemala—countries with low education rates and amongst the highest rates of teen pregnancy— these negative effects on human capital may be working through the effect on daughters' teen childbearing which forces teenagers to school interruption. This explanation for the high AMEs is, however, less plausible for the case of Haiti where AME on childbearing are lower.

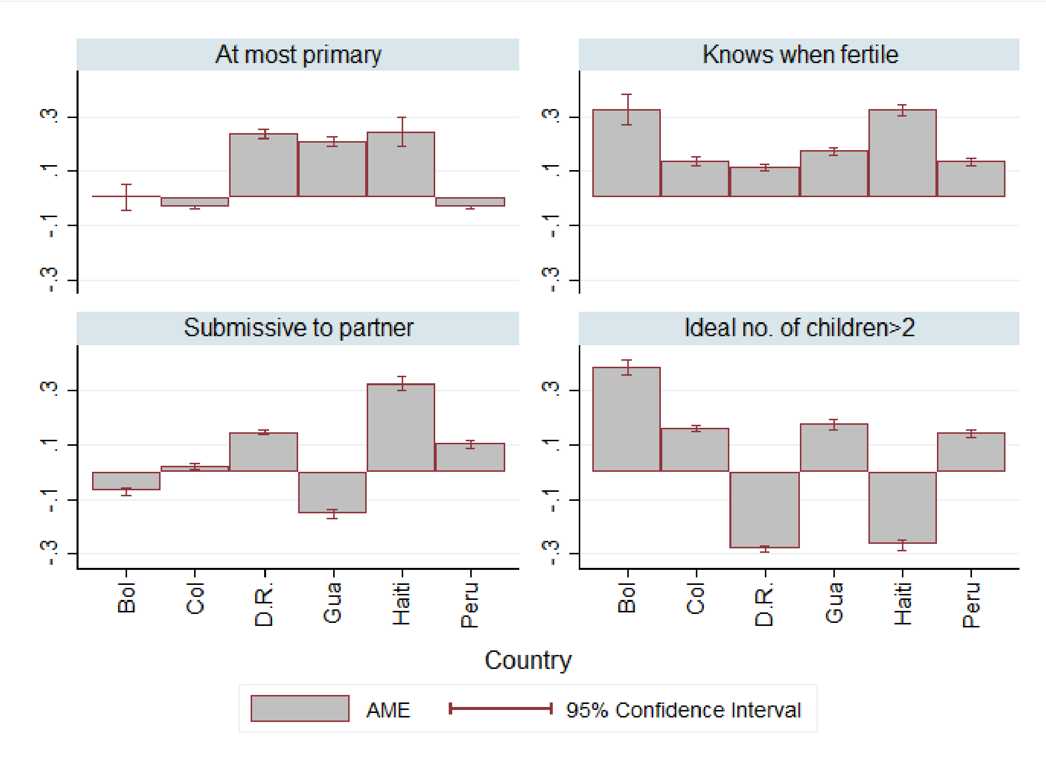

FIGURE 3 Average Marginal Effects of Mothers TCS.

Notes: For "At most primary" we use as dependent variable a dummy that takes value one if the daughter has obtained at most primary education and zero otherwise. For "Ideal no. of childen > 2", the outcome variable takes value one if the daughter declares her ideal number of children to be larger than 2 and zero otherwise. . For "Knows when fertile" the dependent variable takes value one if the daughter correctly states the moment during the menstrual cycle in which a woman is most fertile and zero otherwise. For "Submissive to partner", the dependent variable that takes value one if the teenager supports at least one of the following statements: (i) "the partner having a sexually transmitted disease is no reason to use a condom"; (ii) "the partner having a sexual affair with another individual is no reason to refuse sex". It takes value zero otherwise.

TABLE 8 Ignorability tests. Other outcomes

	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
At most primary	180.607	1931.676	692.655	974.997	45.709	1856.566
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Knows when fertile	41.006	1769.039	178.662	4547.844	3099.614	7012.239
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Submissive to partner	106.906	613.434	8143.863	4260.404	933.307	2399.411
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ideal no. of children>2	884.693	20512.414	9493.012	1117.754	1429.425	6512.533
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Ignorability tests are likelihood ratio tests for the null hypothesis that conditional probabilities $\Pr(I = 1, I^m = 1|y, y^m)$ are invariant to mother and daughter TCS, y^m and y, respectively. p-values in parenthesis.

Our next outcome is knowledge of the women's ovulatory cycle at the time of the survey. The knowledge of ovulatory cycle amongst teenagers is low (albeit correlated with that of the mothers at the country level), from 7.9% of teenagers in Dominican Republic to 25.9% in Bolivia. Given the higher rates of TCS amongst daughters of teen mothers, one could argue that daughters of teen mothers would on average show less knowledge of the ovulatory cycle. However there are at least two effects that should lead to higher averages among daughters of teen mothers. First, early sexual activity is more likely among daughters of teen mothers and that may lead to an increase in knowledge. And second, daughters of teen mothers are more likely to be pregnant or mothers, which, at the time of the survey, likely increases both their's and their mother's knowledge of women's ovulatory cycle. We construct a dependent variable that takes value one if the daughter correctly states the moment during the menstrual cycle in which a woman is most fertile and zero otherwise. Results in the top right graph of Figure 3 and in Table A.3 in the Appendix show that daughters of teen mothers are significantly much more likely to correctly state when a woman is more fertile during her menstrual cycle. For example, in Dominican Republic where only 7.9% of the daughters correctly report when a woman is more fertile, the daughters of teen mothers are 1.46 times more likely to know the correct answer. Notice that the results under ignorability are close to zero and not statistically significant except for Peru.

Next, we explore whether mothers' TCS is related to the adoption of certain gender roles in sexual relationships. In particular, whether having a teen mother makes teenagers more submissive and tolerant to certain behaviors of their sexual partners. The dependent variable in this case takes value one if the teenager supports at least one of the following statements: (i) "the partner having a sexually transmitted disease is no reason to use a condom"; (ii) "the partner having a sexual affair with another individual is no reason to refuse sex". It takes value zero otherwise. The bottom left graph of Figure 3 and Table A.4 in the Appendix reveal the following: First, the acceptance of submissive gender roles in sexual relations runs surprisingly high amongst teenagers, from 6.3% of teenagers in Peru to 24.6%

in Guatemala and 34.6% in Haiti. Secondly, with the exception of Bolivia and Guatemala, having a teen mother reinforces these negative gender roles in sexual relationships. Results under ignorability are small and most are not statistically significant. Results from this exercise suggest that high teen pregnancy rates amongst daughters of teen mothers may in part be explained by submissive gender roles that lead to low rates of condom use.

Finally, we assess whether being a daughter of a teen mother may affect her preferences regarding fertility outcomes. Higher rates of teen pregnancy amongst daughters of teen mothers may be affected by preferences. For example, daughters of teen mothers may have a preference for larger families and that leads them to start young. We assess to what extent teenage preferences depend on their mothers' TCS by defining an outcome which takes value one if the daughter declares her ideal number of children to be larger than 2 and zero otherwise. Results shown in the bottom right graph of Figure 3 and in Table A.5 in the Appendix show that, for all countries except Dominican Republic and Haiti, being the daughter of a teen mother is related to a preference for larger families which could in part explain higher rates of teen childbearing. However, this is not the case in Dominican Republic and Haiti, two countries with very high preferences for large families amongst teenagers, where daughters of teen mothers prefer smaller families. AME results under ignorability are small and poorly correlated with the AMEs obtained with our ML approach.

5 | CONCLUSIONS

We find very strong evidence of intergenerational transmission of sexual behavior from mothers to daughters in the LAC region. Using DHS data from Bolivia, Colombia, Dominican Republic, Guatemala, Haiti and Peru, we find that being a daughter of a teen mother increases very significantly the probability of teen childbearing between 7.4 and 22.2 percentage points (or between 42 and 138%). Similarly, having a mother who had an early initiation to sex increases the probability of early sexual behavior among teens between 20.6 and 27.2 percentage points (or between 61.8 and 104.7%). Importantly, our results show that restricting the sample to matched mother-daughter pairs which, for most surveys and census data means restricting the sample to those teenagers who live with their mothers, leads to a large negative bias. The reason for such large coresidential bias lies in the fact that teenage daughters leave their parents home (and, consequently, become unmatched) when they are pregnant or when they become mothers and, hence, the missing or incomplete mother-daughter pairs are not random. We develop a Maximum Likelihood methodology to avoid the coresidential bias that enables us to use all data in the estimation, including the incomplete mother-daughter pairs, by modeling the missing process together with the childbearing process. Our methodology is easily adaptable to other contexts where missing matches are non-random.

The prevalence of such high intergenerational transmission is at the core of persistent high teenage childbearing rates in the LAC region and suggests alternative public policy fixes. Concretely, complementing policies targeted at teenagers with policies targeted at their mothers, under the motto "when you educate the mother you educate the child," may increase effectiveness and break the intergenerational cycle.

ACKNOWLEDGEMENTS

We thank Lucila Berniell, Dolores De la Mata, Juliana Jaramillo, Ernesto Schargrodsky, and Jan Stuhler, for helpful comments.

REFERENCES

Aguia-Rojas, K., Gallego-Ardila, A. D., Bonilla, M. V. E. and Rodriguez-Nino, J. N. (2020) Individual and contextual factors associated with teenage pregnancy in Colombia: A multilevel analysis. *Maternal and Child Health Journal*, **24**, 1376–1386.

Aizer, A., Devereaux, P. and Salvanes, K. (2020) Grandparents, moms, or dads? why children of teen mothers do worse in life. *Journal of Human Resources*, 1019–10524R2.

Alzate, M. M. (2014) Adolescent pregnancy in colombia: The price of inequality and political conflict. In *International Handbook of Adolescent Pregnancy*, 241–255. Springer.

Alzate, M. M., Dongarwar, D., Matas, J. L. and Salihu, H. M. (2020) The effect of sexual literacy on adolescent pregnancy in Colombia. *Journal of Pediatric and Adolescent Gynecology*, **33**, 72–82.

Angrist, J. D. and Lavy, V. (1996) The effect of teen childbearing and single parenthood on childhood disabilities and progress in school. *NBER working paper*.

Arceo-Gómez, E. O. and Campos-Vazquez, R. M. (2014) Teenage pregnancy in Mexico: evolution and consequences. *Latin American Journal of Economics*, **51**, 109–146.

Ashcraft, A., Fernández-Val, I. and Lang, K. (2013) The consequences of teenage childbearing: Consistent estimates when abortion makes miscarriage non-random. *The Economic Journal*, **123**, 875–905.

Azevedo, J. P., Favara, M., Haddock, S. E., López-Calva, L. F., Muller, M. and Perova, E. (2012a) *Teenage pregnancy and opportunities in Latin America and the Caribbean: on teenage fertility decisions, poverty and economic achievement*. World Bank, Washington, DC.

Azevedo, J. P., Lopez-Calva, L. F. and Perova, E. (2012b) *Is the baby to blame? An inquiry into the consequences of early childbearing*. World Bank, Washington, DC.

Black, S. E. and Devereux, P. J. (2011) Chapter 16 - recent developments in intergenerational mobility**we would like to thank anders björklund, dan hamermesh, helena holmlund, kanikakapur, gary solon, alexandra spitz-oener, and steve trejo for helpful comments. prudence kwenda provided excellent research assistance. devereux thanks the irish research council for the humanities and social sciences (irchss) for financial support. this chapter was completed while black was on leave from the economics department at ucla. e-mail addresses: sblack@austin.utexas.edu (sandra e. black), devereux@ucd.ie (paul j. devereux). vol. 4 of *Handbook of Labor Economics*, 1487–1541. Elsevier. URL: https://www.sciencedirect.com/science/article/pii/S0169721811024142.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2008) Staying in the classroom and out of the maternity ward? the effect of compulsory schooling laws on teenage births. *The Economic Journal*, **118**, 1025–1054.

Caffe, S., Chandra-Mouli, V., Camacho, A. V., Brumana, L., Abdool, S., Gomez Ponce de Leon, R., Gordillo-Tobar, A. and Menard-Freeman, L. (2016) *Accelerating progress toward the reduction of adolescent pregnancy in Latin America and the Caribbean*. Pan American Health Organization/UNFPA/Unicef, Washington.

Card, J. J. (1981) Long-term consequences for children of teenage parents. *Demography*, **18**, 137–156.

Carneiro, P., Garcia, I. L., Salvanes, K. G. and Tominey, E. (2021) Intergenerational mobility and the timing of parental income. *Journal of Political Economy*, **129**, 757–788.

Carro, J., Machado, M. and Mora, R. (2021) The role of mothers on female labor force participation: an approach using historical parish records. http://www.eco.uc3m.es/~ricmora/CMM.pdf. Mimeo.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W. and Yagan, D. (2011) How does your kindergarten classroom affect your earnings? evidence from project STAR. *The Quarterly Journal of Economics*, **126**, 1593–1660.

Dongarwar, D. and Salihu, H. M. (2019) Influence of sexual and reproductive health literacy on single and recurrent adolescent pregnancy in Latin America. *Journal of Pediatric and Adolescent Gynecology*, **32**, 506–513.

Drewry, J. and Garcés-Palacio, I. C. (2020) Socio-demographic characteristics and female empowerment as determinants of adolescent pregnancy in colombia. *International Journal of Sexual Health*, **32**, 293–305.

Estrada, F., Atienzo, E. E., Cruz-Jimenez, L. and Campero, L. (2021) A rapid review of interventions to prevent first pregnancy among adolescents and its applicability to Latin America. *Journal of Pediatric and Adolescent Gynecology*.

Fletcher, J. M. and Wolfe, B. L. (2009) Education and labor market consequences of teenage childbearing evidence using the timing of pregnancy outcomes and community fixed effects. *Journal of Human Resources*, **44**, 303–325.

Francesconi, M. (2008) Adult outcomes for children of teenage mothers. *Scandinavian Journal of Economics*, **110**, 93–117.

Guttmacher Institute (2018) Abortion in Latin America and the Caribbean. Fact Sheet.

Hotz, V. J., McElroy, S. W. and Sanders, S. G. (2005) Teenage childbearing and its life cycle consequences exploiting a natural experiment. *Journal of Human Resources*, **40**, 683–715.

Kahn, J. R. and Anderson, K. E. (1992) Intergenerational patterns of teenage fertility. *Demography*, **29**, 39–57.

Kruger, D. and Berthelon, M. (2012) Education consequences of adolescent motherhood in Chile. *Washington, DC: Banco Mundial*.

Lang, K. and Nuevo-Chiquero, A. (2012) Trends in self-reported spontaneous abortions: 1970–2000. *Demography*, **49**, 989–1009.

Lang, K. and Weinstein, R. (2015) The consequences of teenage childbearing before Roe v. Wade. *American Economic Journal: Applied Economics*, **7**, 169–97.

Levine, J. A., Pollack, H. and Comfort, M. E. (2001) Academic and behavioral outcomes among the children of young mothers. *Journal of Marriage and Family*, **63**, 355–369.

Millán, J. (2019) Drugs, guns and early motherhood in Colombia. Mimeo.

Mohr, R., Carbajal, J. and Sharma, B. B. (2019) The influence of educational attainment on teenage pregnancy in low-income countries: A systematic literature review. *Journal of Social Work in the Global Community*, **4**, 2.

Neal, S., Harvey, C., Chandra-Mouli, V., Caffe, S. and Camacho, A. V. (2018) Trends in adolescent first births in five countries in latin america and the caribbean: disaggregated data from demographic and health surveys. *Reproductive health*, **15**, 1–10.

Ramalho, E. A. and Smith, R. J. (2013) Discrete choice non-response. *The Review of Economic Studies*, **80**, 343–364.

Tsaneva, M. and Gunes, P. M. (2020) The effect of violent crime on teenage pregnancy in Mexico. *Review of Economics of the Household*, **18**, 141–164.

A | APPENDIX

TABLE A.1 Terminations & live births

PANEL A: During adolescence

	% women with terminations					
	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
Only no teen-mothers	1.918	3.578	2.634	1.946	1.369	1.907
Only teen-mothers	3.694	7.879	8.707	7.246	2.795	4.935
All	2.498	4.947	5.010	3.496	1.677	2.724
	% women sexually active with terminations					
	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
Only no teen-mothers	3.074	4.728	3.898	2.927	1.869	2.988
Only teen-mothers	3.631	7.823	8.678	7.353	2.812	4.956
All	3.310	5.937	6.178	4.692	2.126	3.715
	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
Terminations rate	7.360	14.432	12.300	11.419	7.396	9.593
Chilbearing Rate	32.644	31.841	39.128	29.237	21.596	27.009

PANEL B: After adolescence

	% sexually active women with terminations					
	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
All	6.151	9.836	13.525	4.047	3.828	7.186

Notes: Pooled data from all surveys post 2000. Women aged 23 during the year of interview. Terminations include abortions, misscarriages, and still births. *Childbearing rate* is the percentage of women who have had at least one live birth during adolescence. *Terminations rate* is the percentage of women with at least one termination during adolescence over all women who were pregnant during adolescence.

TABLE A.2 Effect on education: At most Primary.Non-ignorable ML estimates.

	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
TCS^m	0.018	-0.218***	0.789***	0.737***	0.773***	-0.226***
	(0.199)	(0.042)	(0.058)	(0.056)	(0.218)	(0.039)
Dcohort1	0.180*	0.001	0.568***	0.564***	0.384***	0.136***
	(0.092)	(0.048)	(0.057)	(0.048)	(0.090)	(0.043)
Dprimary	0.736***	0.362***	0.716***	1.055***	0.933***	0.252***
	(0.113)	(0.046)	(0.059)	(0.128)	(0.138)	(0.047)
Dage16	-0.555***	-0.104***	-0.428***	-0.225***	-0.262***	-0.054
	(0.077)	(0.037)	(0.052)	(0.060)	(0.079)	(0.035)
Dage17	-0.589***	-0.216***	-0.700***	-0.258***	-0.585***	-0.044
	(0.076)	(0.038)	(0.054)	(0.061)	(0.080)	(0.037)
Dage18	-0.744***	-0.102***	-0.861***	-0.332***	-0.622***	-0.043
	(0.081)	(0.038)	(0.057)	(0.063)	(0.083)	(0.038)
Drural	0.707***	0.510***	0.113***	0.509***	0.398***	0.782***
	(0.068)	(0.033)	(0.041)	(0.047)	(0.069)	(0.029)
Dhsize2	0.179**	0.549***	0.445***	0.300***	0.105	0.443***
	(0.084)	(0.041)	(0.055)	(0.054)	(0.081)	(0.038)
Dpoor	0.332***	0.524***	0.307***	0.800***	0.404***	0.619***
	(0.100)	(0.056)	(0.054)	(0.056)	(0.095)	(0.053)
Dwork	-0.166*	-0.100**	0.040	-0.150***	-0.123	-0.314***
	(0.088)	(0.042)	(0.050)	(0.054)	(0.085)	(0.042)
N. obs.	2979	19723	6635	4845	2789	20987
logL	-1.03e+04	-8.79e+04	-3.19e+04	-2.32e+04	-8659.909	-8.38e+04
AME(TCS^m)	0.005	-0.033***	0.237***	0.210***	0.244***	-0.034***
	(0.024)	(0.003)	(0.008)	(0.009)	(0.027)	(0.003)
$\overline{y}$	0.312	0.116	0.365	0.463	0.382	0.124
AME under	0.068***	0.021***	0.023	0.053***	0.051*	0.012**
ignorability	(0.020)	(0.005)	(0.015)	(0.016)	(0.027)	(0.005)

Notes: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable y is a dummy that takes value one if the daughter has obtained, at most, primary education and zero otherwise. All models include survey dummies. Standard errors in parenthesis. Variable TCS^m is the dummy for the mother's teen chilbearing status. Dcohort1 is a dummy variable for mother 45 and older. Dprimary is a dummy variable that takes value 1 if the mother has no more than primary education. Dage16-Dage18 are age dummies (reference category is age 15). Drural takes value 1 when the household is located in a rural area. Dhsize2 is a dummy variable for parents households with at least seven members. Dpoor takes value 1 if the parents' household belongs to the two poorest quintiles of the country's population based on a continuous wealth measure produced by the *DHS*. Dummy Dwork takes value 1 if the mother works. logL stands for the loglikelihood function evaluated at the ML estimates. AME(TCS^m) is the estimated Average Marginal Effect of TCS^m. $\overline{y}$ is the average value of y in the estimating sample. AME under ignorability stands for the average marginal effect of TCS^m using the probit estimates under ignorability for the same variable specification.

TABLE A.3 Effect on fertility knowledge. Non-ignorable ML estimates.

	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
TCS^m	1.017***	0.432***	0.711***	0.911***	1.340***	0.480***
	(0.199)	(0.055)	(0.111)	(0.059)	(0.096)	(0.045)
Dcohort1	0.122*	0.179***	0.218***	0.242***	0.350***	0.113***
	(0.072)	(0.032)	(0.078)	(0.055)	(0.084)	(0.031)
Dprimary	-0.147*	-0.137***	0.015	-0.218***	-0.209**	-0.129***
	(0.078)	(0.030)	(0.068)	(0.082)	(0.100)	(0.031)
Dage16	0.005	0.105***	0.184**	0.199***	0.167*	0.142***
	(0.080)	(0.029)	(0.075)	(0.072)	(0.099)	(0.030)
Dage17	0.196**	0.181***	0.276***	0.233***	0.230**	0.306***
	(0.078)	(0.030)	(0.073)	(0.074)	(0.098)	(0.030)
Dage18	0.412***	0.275***	0.265***	0.427***	0.505***	0.471***
	(0.078)	(0.030)	(0.075)	(0.073)	(0.095)	(0.030)
Drural	-0.382***	-0.094***	-0.024	-0.245***	-0.109	-0.188***
	(0.074)	(0.029)	(0.054)	(0.053)	(0.075)	(0.026)
Dhsize2	-0.201***	-0.174***	-0.182**	-0.226***	-0.330***	-0.121***
	(0.072)	(0.032)	(0.083)	(0.062)	(0.080)	(0.030)
Dpoor	0.094	-0.229***	-0.298***	-0.171**	-0.333***	-0.194***
	(0.090)	(0.032)	(0.072)	(0.068)	(0.095)	(0.035)
Dwork	-0.101	-0.009	0.202***	0.023	0.034	-0.024
	(0.073)	(0.028)	(0.065)	(0.059)	(0.083)	(0.029)
N. obs.	2975	19723	6612	4844	2789	20984
logL	-1.05e+04	-9.20e+04	-2.99e+04	-2.23e+04	-8168.374	-8.73e+04
AME(TCS^m)	0.324***	0.134***	0.115***	0.173***	0.325***	0.134***
	(0.029)	(0.009)	(0.006)	(0.007)	(0.010)	(0.007)
$\bar{y}$	0.259	0.229	0.079	0.133	0.162	0.200
AME under	0.014	-0.010	0.002	0.009	-0.032	-0.013*
ignorability	(0.025)	(0.009)	(0.011)	(0.013)	(0.023)	(0.008)

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The dependent variable y is a dummy that takes value one if the daughter correctly states the moment during the menstrual cycle in which a woman is likely most fertile and zero otherwise. All models include survey dummies. Standard errors in parenthesis. Variable TCS^m is the dummy for the mother's teen chilbearing status. Dcohort1 is a dummy variable for mother 45 and older. Dprimary is a dummy variable that takes value 1 if the mother has no more than primary education. Dage16-Dage18 are age dummies (reference category is age 15). Drural takes value 1 when the household is located in a rural area. Dhsize2 is a dummy variable for parents households with at least seven members. Dpoor takes value 1 if the parents' household belongs to the two poorest quintiles of the country's population based on a continuous wealth measure produced by the *DHS*. Dummy Dwork takes value 1 if the mother works. logL stands for the loglikelihood function evaluated at the ML estimates. AME(TCS^m) is the estimated Average Marginal Effect of TCS^m. $\bar{y}$ is the average value of y in the estimating sample. AME under ignorability stands for the average marginal effect of TCS^m using the probit estimates under ignorability for the same variable specification.

TABLE A.4 Submissive Gender Role in Sex. Non-ignorable ML estimates.

	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
TCS^m	-0.369***	0.136*	0.991***	-0.550***	0.887***	0.645***
	(0.081)	(0.072)	(0.056)	(0.053)	(0.099)	(0.094)
Dcohort1	-0.021	-0.042	0.275***	-0.093*	0.280***	0.216***
	(0.094)	(0.056)	(0.069)	(0.048)	(0.088)	(0.047)
Dprimary	0.231**	0.021	0.296***	0.474***	0.049	0.326***
	(0.115)	(0.051)	(0.069)	(0.105)	(0.125)	(0.044)
Dage16	-0.015	-0.040	-0.100	-0.007	-0.040	-0.087**
	(0.087)	(0.047)	(0.066)	(0.058)	(0.090)	(0.041)
Dage17	-0.128	-0.061	-0.173**	-0.176***	-0.045	-0.080*
	(0.087)	(0.048)	(0.067)	(0.060)	(0.090)	(0.043)
Dage18	-0.191**	-0.102**	-0.282***	-0.250***	0.112	-0.153***
	(0.091)	(0.050)	(0.072)	(0.062)	(0.092)	(0.046)
Drural	0.437***	0.248***	0.080	0.267***	-0.009	0.302***
	(0.076)	(0.042)	(0.052)	(0.048)	(0.074)	(0.035)
Dhsize2	0.295***	0.242***	0.013	0.270***	-0.057	0.021
	(0.088)	(0.050)	(0.066)	(0.056)	(0.089)	(0.041)
Dpoor	0.074	0.240***	0.063	0.483***	-0.045	-0.030
	(0.106)	(0.057)	(0.066)	(0.058)	(0.105)	(0.046)
Dwork	-0.036	-0.026	-0.061	-0.032		-0.024
	(0.095)	(0.050)	(0.061)	(0.055)		(0.045)
N. obs.	2979	11817	6632	4845	1902	20987
logL	-9949.205	-4.72e+04	-3.02e+04	-2.30e+04	-6660.531	-8.21e+04
AME(TCS^m)	-0.071***	0.022***	0.146***	-0.151***	0.326***	0.103***
	(0.007)	(0.006)	(0.004)	(0.008)	(0.013)	(0.008)
$\bar{y}$	0.133	0.086	0.090	0.246	0.346	0.063
AME under ignorability	-0.016	0.017**	0.024**	-0.001	-0.029	-0.003
	(0.017)	(0.007)	(0.011)	(0.016)	(0.036)	(0.005)

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The dependent variable y is a dummy that takes value one if the daughter supports at least one of the following statements: (i) "the partner having a sexually transmitted desease is no reason to use a condom"; (ii) "the partner having a sexual affair with another individual is no reason to refuse sex". It takes value zero otherwise. Standard errors in parenthesis. Variable TCS^m is the dummy for the mother's teen childbearing status. Dcohort1 is a dummy variable for mother 45 and older. Dprimary is a dummy variable that takes value 1 if the mother has no more than primary education. Dage16-Dage18 are age dummies (reference category is age 15). Drural takes value 1 when the household is located in a rural area. Dhsize2 is a dummy variable for parents households with at least seven members. Dpoor takes value 1 if the parents' household belongs to the two poorest quintiles of the country's population based on a continuous wealth measure produced by the *DHS*. Dummy Dwork takes value 1 if the mother works. logL stands for the loglikelihood function evaluated at the ML estimates. AME(TCS^m) is the estimated Average Marginal Effect of TCS^m. $\bar{y}$ is the average value of y in the estimating sample. AME under ignorability stands for the average marginal effect of TCS^m using the probit estimates under ignorability for the same variable specification.

TABLE A.5 Effect on fertility preferences. Non-ignorable ML estimates.

	Bolivia	Colombia	DR	Guatemala	Haiti	Peru
TCS^m	1.243***	0.603***	-0.749***	0.474***	-0.760***	0.507***
	(0.104)	(0.049)	(0.045)	(0.049)	(0.075)	(0.042)
Dcohort1	0.218***	0.169***	-0.203***	0.377***	-0.057	0.231***
	(0.068)	(0.035)	(0.048)	(0.041)	(0.074)	(0.031)
Dprimary	-0.198***	-0.003	-0.150***	0.320***	0.033	0.031
	(0.076)	(0.031)	(0.047)	(0.075)	(0.102)	(0.032)
Dage16	0.012	0.017	-0.017	0.035	0.119*	-0.027
	(0.081)	(0.032)	(0.046)	(0.053)	(0.072)	(0.028)
Dage17	0.129	0.041	0.088*	0.083	0.175**	-0.043
	(0.080)	(0.032)	(0.046)	(0.054)	(0.072)	(0.029)
Dage18	0.148*	0.129***	0.195***	0.119**	0.176**	0.037
	(0.082)	(0.032)	(0.048)	(0.055)	(0.073)	(0.030)
Drural	-0.181**	0.088***	0.063*	0.349***	0.267***	-0.032
	(0.073)	(0.029)	(0.035)	(0.041)	(0.063)	(0.025)
Dhsize2	-0.025	0.207***	0.192***	0.243***	0.197***	0.161***
	(0.069)	(0.031)	(0.046)	(0.049)	(0.075)	(0.029)
Dpoor	0.057	0.087**	0.198***	0.058	0.247***	-0.112***
	(0.084)	(0.035)	(0.045)	(0.052)	(0.087)	(0.035)
Dwork	-0.052	-0.023	0.018	-0.083*	0.117	-0.163***
	(0.070)	(0.030)	(0.042)	(0.048)	(0.077)	(0.029)
N. obs.	2937	19654	6616	4838	2788	20921
logL	-1.03e+04	-9.00e+04	-3.25e+04	-2.36e+04	-8830.901	-8.72e+04
AME(TCS^m)	0.385***	0.161***	-0.281***	0.176***	-0.267***	0.143***
	(0.014)	(0.007)	(0.007)	(0.010)	(0.010)	(0.007)
$\overline{y}$	0.236	0.158	0.490	0.523	0.414	0.199
AME under	-0.001	-0.009	0.039**	-0.053***	-0.035	0.017**
ignorability	(0.024)	(0.007)	(0.019)	(0.019)	(0.031)	(0.008)

Notes: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable y is a dummy that takes value one if the daughter declares her ideal number of children to be larger than 2 and zero otherwise. All models include survey dummies. Standard errors in parenthesis. Variable TCS^m is the dummy for the mother's teen chilbearing status. Dcohort1 is a dummy variable for mother 45 and older. Dprimary is a dummy variable that takes value 1 if the mother has no more than primary education. Dage16-Dage18 are age dummies (reference category is age 15). Drural takes value 1 when the household is located in a rural area. Dhsize2 is a dummy variable for parents households with at least seven members. Dpoor takes value 1 if the parents' household belongs to the two poorest quintiles of the country's population based on a continuous wealth measure produced by the *DHS*. Dummy Dwork takes value 1 if the mother works. logL stands for the loglikelihood function evaluated at the ML estimates. AME(TCS^m) is the estimated Average Marginal Effect of TCS^m. $\overline{y}$ is the average value of y in the estimating sample. AME under ignorability stands for the average marginal effect of TCS^m using the probit estimates under ignorability for the same variable specification.