

# College Opportunity and Teen Fertility: Evidence from *Ser Pilo Paga* in Colombia\*

Michael D. Bloem<sup>†</sup>

Jesús Villero<sup>‡</sup>  
(Job Market Paper)

February 2023

[Most recent version here](#)

## Abstract

We study the effects of an increase in post-secondary educational opportunities on teen fertility by exploiting policy-induced variation from *Ser Pilo Paga* (SPP), a generous college financial aid program in Colombia that dramatically expanded college opportunities for low-income students. Our preferred empirical approach uses a triple difference design that leverages variation in the share of female students eligible for the program across municipalities and the fact that the introduction of SPP should not affect the education and fertility decisions of older women not targeted by the program. We find that after the introduction of SPP, fertility rates for women aged 15-19 years old decreased in more affected municipalities by about 6 percent relative to less affected municipalities. This effect accounts for approximately one-fourth of the overall decrease in teen fertility observed in the years following the program's announcement. Our results suggest that increasing economic opportunities through expanding college access can contribute to lowering teen fertility rates.

**JEL Codes:** H52, I22, I23, I24, I25, J13, O15.

\*We thank Daniel Kreisman, Jonathan Smith, Lauren Hoehn-Velasco, Thomas A. Mroz, Matias Busso, Juliana Londoño-Vélez, and Lasse Brune for generous feedback on earlier versions of this manuscript. We also thank seminar participants at Georgia State University, Universidad del Norte, Universidad Alberto Hurtado, the 2022 Economic Demography Workshop, the 15th Annual Meeting of the Impact Evaluation Network of LACEA, and the Carolina Region Empirical Economics Day for helpful comments.

<sup>†</sup>Department of Economics, Andrew Young School of Policy Studies, Georgia State University. 55 Park Place, 6th Floor, Atlanta, GA 30303. E-mail: [mbloem1@gsu.edu](mailto:mbloem1@gsu.edu).

<sup>‡</sup>Department of Economics, Andrew Young School of Policy Studies, Georgia State University. 55 Park Place, 6th Floor, Atlanta, GA 30303. E-mail: [jvilleroarocal1@gsu.edu](mailto:jvilleroarocal1@gsu.edu).

# 1 Introduction

Teen childbearing is associated with worse outcomes for mothers and their children, including lower educational attainment and poorer labor market outcomes, and with large public costs, including greater reliance on social programs (Kearney and Levine, 2012; Azevedo, Favara, Haddock, López-Calva, Muller and Perova, 2012; Aizer, Devereux and Salvanes, 2022). Not surprisingly, reducing its incidence is a persistent goal among national governments and international agencies. Furthermore, rates of teen childbearing are higher among low-income communities and in places with greater levels of income inequality (Kearney and Levine, 2012, 2014). Youth may be more likely to engage in risky behaviors when chances of economic mobility are low and opportunities to make investments in their own economic progress are limited. Thus, one possible way to break this cycle of early childbearing and poverty is to focus policies on reducing inequality in opportunities for youth to make investments in their own economic progress.

In this paper, we investigate how teen behavior responds to increases in post-secondary educational opportunities by studying the effects on teen fertility of Colombia’s 2014 introduction of *Ser Pilo Paga* (roughly translated as “Being a Good Student Pays Off”), a college financial aid program covering full tuition costs at high-quality institutions for high-achieving, low-income students. The Colombian setting is characterized by high teenage fertility rates and high economic inequality, as indicated by the cross-country comparison in Figure A1, with large income-based gaps in college enrollment, high college tuition costs, and little existing access to credit before *Ser Pilo Paga* (SPP).

This setting is suitable for studying this topic because SPP had large educational effects. Londoño-Vélez, Rodríguez and Sánchez (2020) show that SPP dramatically increased college enrollment on the eligibility margin (57 to 87 percent increases depending on the complier population), virtually eliminating the income-based gap in college enrollment among high-achieving students. Importantly, since colleges increased supply to capture the additional demand, SPP also increased college enrollment among low-income, aid-ineligible students by 14 percent.

There is also evidence that the introduction of SPP altered human capital investment decisions before college. Bernal and Penney (2019) and Laajaj, Moya and Sánchez (2022) both show that test scores on the national high school exit exam increased among low-income students immediately after the introduction of SPP,

particularly at the top of the score distribution. [Laajaj et al. \(2022\)](#) further show that test scores also increased for low-income students on the national *9th grade* exam, which the authors characterize as a “motivational” effect of SPP. Importantly, they show that these motivational effects on 9th grade test scores reached all the way down to the 19th percentile of the score distribution, illustrating that SPP’s behavioral effects extend far beyond the top of the academic distribution when students have more time to prepare to take the high school exit exam that determines their eligibility for SPP.

Given the evidence above, it is reasonable to expect that teens may also alter their behavior on other non-academic dimensions, like decisions about childbearing. We highlight a few notable potential mechanisms for how SPP could affect teenage fertility rates. First, by increasing college attendance, SPP could reduce fertility due to teens having less time to engage in risky behaviors (i.e., a pure incapacitation effect). Second, receiving the scholarship could grant teens greater access to contraceptives through an income effect. Third, SPP can decrease fertility by increasing the opportunity cost of becoming pregnant as a teenager or through a motivational effect that leads students to pursue college attendance opportunities unavailable before SPP. Since the first two mechanisms would largely derive from scholarship recipients after completing high school and enrolling in college, we refer to these potential channels as SPP’s “direct” effects on fertility. Conversely, since the third mechanism would mostly derive from students before determining eligibility for SPP, we refer to this channel as SPP’s “indirect” (or ex-ante, as in [Laajaj et al. \(2022\)](#)) effects on fertility.

Our preferred empirical approach uses a triple difference research design leveraging municipality-level variation in SPP eligibility rates determined prior to the introduction of the program and the fact that SPP should not affect the fertility decisions of older women aged 25-29. Eligibility for SPP was based on test scores on the national standardized high school exit exam and scores on a household wealth index. By the time SPP was initially announced, students had already taken the high school exit exam and there was not sufficient time to request a reevaluation of their household wealth index. Our empirical approach uses eligibility rates only from this first cohort of students, who could not influence their scores around the eligibility cutoffs.

We find that fertility rates for women aged 15-19 decreased by about 6 percent in more affected municipalities relative to less affected municipalities. This accounts

for approximately one-fourth of the overall decrease in teen fertility observed in the years following SPP's announcement. We rule out that our observed effects are entirely driven by the direct effects of receiving the scholarship upon finishing high school. The timing of the decline in fertility rates—and the fact that the number of fewer births implied by our estimates is larger than the number of actual female SPP scholarship recipients—suggests that incapacitation or income effects of receiving the scholarship cannot fully explain the results. We also show results on rates of teen fatherhood, using the more granular data available on father's age that is not available for mother's age, that document effects even for younger teens aged 15-17 who are not old enough to have received the scholarship. Thus, we interpret our findings as largely comprised of indirect effects of SPP, where the new college opportunities created by the program influenced teen fertility decisions before being able to benefit from the program directly. This is consistent with the ex-ante motivational effects on test scores documented by [Laajaj et al. \(2022\)](#).

We also find that the teen fertility impacts of SPP are larger in municipalities that, before the program, exhibited higher levels of income inequality and a higher share of female students reporting low expectations of enrolling in higher education after finishing high school.<sup>1</sup> These results are broadly consistent with inequality, and “economic hopelessness”, being an important determinant of teen childbearing rates ([Kearney and Levine, 2014](#)). In addition, we document that the relative reduction in teen births is more prominent in municipalities where female teenagers tend to have children with other teenagers—perhaps indicating a reinforcement of incentives to avoid parenthood when potential fathers also face enhanced college opportunities or increased bargaining power for female teens in such relationships.

Our results are robust to alternative specifications and empirical approaches. We show that the estimated effect of SPP on teen fertility increases in magnitude as a municipality's initial SPP eligibility rate increases, which illustrates that our preferred estimates are not dependent upon how we characterize municipalities as more or less affected by SPP. Furthermore, we see results that are consistent with our main estimates when using a simpler difference-in-differences design, and when using alternative triple difference approaches that rely on different sources of variation. For instance, triple difference results are similar when using a municipality's distance to the nearest SPP-eligible higher education institution instead

---

<sup>1</sup>[Figure A2](#) shows the pre-SPP correlation between adolescent fertility and income inequality and access to higher education in Colombia.

of variation in SPP eligibility rates, and when using women aged 15-19 with completed education less than eighth grade (and thus likely a school dropout) as a comparison group instead of women aged 25-29. We also rule out that possible confounding events drive our results, including Colombia's peace agreement with the Revolutionary Armed Forces of Colombia, the Zika virus epidemic, and the partial introduction of extended school days in some municipalities.

We add to the literature in three important ways. Our primary contribution is documenting teen fertility responses to a large change in post-secondary schooling *opportunities*, which suggest that improving the future economic prospects of young women through college opportunities can reduce teen pregnancy and early childbearing. An existing literature studies the effects of education on teen pregnancy using exogenous variation from school entry policies and mandatory schooling laws (e.g., [Black, Devereux and Salvanes \(2008\)](#); [McCrary and Royer \(2011\)](#); [Alzúa and Velázquez \(2017\)](#)) and from the duration of school days ([Berthelon and Kruger, 2011](#)). Since these policies require additional time to be spent in school, evidence of declines in adolescent fertility in these settings may be due to either an incapacitation effect or a true "human capital" effect of the extra years of contemporaneous education, but not to expanded *future, non-contemporaneous* educational opportunities ([Doleac and Gibbs, 2016](#); [Alzúa and Velázquez, 2017](#)).

Our analysis represents an empirical test of theoretical predictions that increases in economic opportunities (and increases in opportunity costs) influence the fertility decisions of young women ([Becker, 1960](#); [Willis, 1973](#); [Kearney and Levine, 2014](#)). Little is known about how increasing opportunities for schooling affects fertility decisions, where youth still have agency in their schooling choices or where schooling cannot be made compulsory, such as with college attendance. Closest to our work are [Cowan \(2011\)](#) and [Koohi \(2017\)](#) who show that tuition costs at colleges in the United States are positively associated with various risky behaviors of youth, such as the number of sexual partners within the past year ([Cowan, 2011](#)) and the prevalence of adolescent childbearing among undocumented Mexican immigrants ([Koohi, 2017](#)). We advance this work by exploiting a large-scale program that provides a cleaner shock to post-secondary educational opportunities in a context with more certainty around the labor market returns to investments in higher education and where imperfect credit markets and limited financial aid make it more difficult for low-income students to attend college.

Second, our paper is related to a literature that studies the teen fertility impacts

of interventions in developing countries that aim to improve economic opportunities and empowerment for adolescent women (Jensen, 2012; Duflo, Dupas and Kremer, 2015, 2021; Muralidharan and Prakash, 2017; Bandiera, Buehren, Burgess, Goldstein, Gulesci, Rasul and Sulaiman, 2020). We extend this body of work by providing evidence on how opportunities for college attendance, rather than primary or secondary schooling, relates to adolescent fertility decisions. This evidence is particularly relevant for countries where, like Colombia, secondary schooling is relatively accessible and where attending college is increasingly important for economic mobility. Third, by examining understudied non-educational outcomes (Cowan, 2011; Doleac and Gibbs, 2016; Koohi, 2017), we build on the literature of the effects of the *Ser Pilo Paga* program (Londoño-Vélez et al., 2020; Bernal and Penney, 2019; Laajaj et al., 2022) and the effects of college financial aid programs more broadly on the decisions of high school students (Cáceres-Delpiano, Giolito and Castillo, 2018).

The remainder of this paper is organized as follows: In the next section, we discuss the Colombian context and describe the details of the *Ser Pilo Paga* program. The third section describes the data sources we use, discusses the key variables used in our analyses, and presents trends in fertility rates in Colombia. The fourth section discusses our identification strategies and estimation approaches. The fifth section presents our core empirical results and section six tests for the sensitivity and robustness of those results. Finally, section seven concludes.

## 2 Background

### 2.1 Teen fertility in Colombia

Similar to many Latin American and Caribbean countries, teen fertility is high in Colombia. Estimated at 70.7 births per 1,000 women aged 15-19 years in 2014 (when SPP was announced), the adolescent fertility rate in Colombia was slightly higher than the Latin American average, more than twice that of other countries with similar income levels and nearly three times higher than in the United States.<sup>2</sup> These “higher-than-expected” adolescent fertility rates observed in Latin American

---

<sup>2</sup>As a region, Latin America and the Caribbean has the second highest fertility rate for teenagers globally, second only to Sub-Saharan Africa. A general discussion about this phenomenon can be found in Azevedo et al. (2012). A cross-country comparison, highlighting Colombia, is presented in Figure A1. Data are from the World Bank’s World Development Indicators.



countries are likely associated with the high levels of inequality of income (and opportunities) observed in the region (Azevedo et al., 2012).

In contrast, in 2014, Colombia had a lower *total* fertility rate than the average Latin American country, similar to the overall fertility rates in other upper middle-income countries and the United States. About 22% of the overall number of births in the country were from mothers aged 19 or younger that year. As a result, early childbearing is a worrisome phenomenon and a policy concern in Colombia, given its association with worse prospects for the adolescent mothers and their children in terms of health, education, and labor market outcomes (Gaviria, 2010; Azevedo et al., 2012; Urdinola and Ospino, 2015).

Early parenting in Colombia is primarily a female phenomenon. Data from the most recent Demographic and Health Survey (2015) show that adolescent women are 6.4 times more likely to have at least one child than adolescent men—13.6 percent versus 2.1 percent (Flórez and Soto, 2019). Furthermore, birth records data indicate that only 22 percent of births to adolescent women between 2008 and 2014 had a teenage father.<sup>3</sup> While teenage pregnancy affects all income groups, it is particularly worrying among low-income women. Low-income Colombian teenagers are five times more likely to have ever been pregnant than their high-income peers (Flórez and Soto, 2019).

In the last decade and a half, Colombia has implemented several programs and policies directly aimed at reducing teenage pregnancies.<sup>4</sup> Among the most relevant initiatives is the implementation of the Youth Friendly Health Services Model (SSAAJ, from the Spanish acronym) and the Program of Education in Sexuality and Construction of Citizenship (PESCC), both launched in 2007-2008 and scaled up nationally in subsequent years.<sup>5</sup> In 2012, the national government additionally launched a strategic framework to address the issue comprehensively, articulating different actors within the public sector.<sup>6</sup> On top of others not directly targeted at reducing fertility like *Familias en Acción*, the conditional cash transfer program in Colombia, these initiatives likely contributed to the downward trend in teenage fertility observed in the country since the mid-2000s after a concerning period of increase during the 1990s (Flórez and Soto, 2019; Attanasio, Sosa, Medina, Meghir

---

<sup>3</sup>The age of consent in Colombia is 14 years old.

<sup>4</sup>See part three of Vargas Trujillo, Flórez, Cortés and Ibarra, eds (2019) for a recent review.

<sup>5</sup>*Modelo de Servicios de Salud Amigables para Adolescentes y Jóvenes* (SSAAJ) and *Programa de Educación para la Sexualidad y Construcción de Ciudadanía* (PESCC) in Spanish.

<sup>6</sup>National Department of Planning (DNP). *Documento CONPES Social* No. 147.

and Posso-Suárez, 2021).<sup>7</sup>

## 2.2 *Ser Pilo Paga* and higher education in Colombia

*Ser Pilo Paga* was announced by surprise on October 1st of 2014 by President Santos's administration. The program was publicly funded and covered recipients' full tuition cost of attending an undergraduate program at any university in Colombia with a High Quality Accreditation. The aid came in the form of a loan that is forgiven upon graduation, although only about 1.9 percent of SPP beneficiaries from the first three cohorts had dropped out of the program (Londoño-Vélez et al., 2020).<sup>8</sup> Additionally, SPP recipients would receive a biannual stipend of at least the national minimum wage to help cover students' living expenses.

Eligibility for SPP was based on both need and merit. First, students must score above a cutoff on the SABER 11, which is similar to the SAT in the United States. The SABER 11 exam is taken by nearly all high school seniors regardless of their plans to attend an institution of higher education. SABER 11 scores play a significant role in college admissions, with about four-fifths of institutions using them in admissions considerations (OECD and World Bank, 2012). The SABER 11 cutoff score is placed at approximately the 91st percentile each year.

Second, students must be below a cutoff on the SISBEN, Colombia's wealth index used to target social welfare programs. The SISBEN cutoff varies by geographic location. The cutoff is 57.21 (over 100) in the 14 main metropolitan areas, 56.32 in other urban areas, and 40.75 in rural areas. Between 2015 and 2018, there were about 10,000 SPP beneficiaries per year (43% of them women), which represents about one-third of students attending an institution with High Quality Accreditation.

In the first year of the program, students had already taken the SABER 11 exam before SPP was announced. Moreover, there was insufficient time to request a reevaluation of their household wealth index before determining eligibility for SPP. Thus, students in this first cohort had no opportunities to influence their test scores or wealth index scores in response to the SPP eligibility cutoffs.

Tuition at the high-quality private universities is very expensive, both compared to private universities in other countries and to the public universities in Colombia

---

<sup>7</sup>Since all these policies were implemented years before SPP was introduced, we do not view them as threats to our identification strategy, but rather as possible factors explaining the decline in adolescent fertility observed before SPP.

<sup>8</sup>The SPP program considers students to have dropped out if they have not attended a high-quality institution for three or more consecutive semesters.



([OECD and World Bank, 2012](#)). Since the tuition at the high-quality public universities is relatively low, these institutions are historically oversubscribed, leading to highly selective admissions. Prior to SPP, there were very few financial aid opportunities for high-achieving, low-income students. Only 11 percent of first-year undergraduate students had a student loan before SPP ([Ferreyra, Avitabile, Botero Álvarez, Haimovich Paz and Urzúa, 2017](#)).

### 3 Data and key variables

This section describes our data sources and key variables. We gather data from publicly available sources on births and population counts in Colombia in order to calculate age-specific fertility rates. To compute a measure that indicates which municipalities were more or less affected by SPP, we collect SABER 11 test score data to calculate SPP eligibility rates.

#### 3.1 Data sources

We use the universe of birth records and annual population estimates from the Colombian National Department of Statistics (DANE, from the Spanish acronym) from 2008 to 2020. Individual birth records contain information about the mother's age in 5-year intervals (i.e., 15-19, 20-24, 25-29, etc.) and about her municipality of *residence* (in addition to where the birth took place). The records also contain the year and month of occurrence of each birth. We use this data to create a municipality by age group and year panel dataset of age-specific fertility rates, which is our primary outcome.

We also use administrative data from the Colombian Institute for the Assessment of Education (ICFES) containing student-level information of the national standardized high school exit exam, SABER 11, including test scores and socio-demographic characteristics. Importantly, this data includes information about SISBEN eligibility and the municipality of residence of the student.

Finally, we complement our data with pre-SPP municipality characteristics which we obtain from the Center for the Study of Economic Development (CEDE) from Universidad de los Andes, the Ministry of Education, and DANE.

### 3.2 Construction of analysis measures

We use the birth records and population estimates to create a municipality-of-residence by age group panel dataset of age-specific fertility rates, our primary outcome. Our main estimates use the natural log of these fertility rates. Throughout the descriptive and econometric analyses that follow, we account for the lag between conception and birth by using the year-month of birth and the gestational age at delivery to approximate the year-month of conception of each newborn. For 80 percent of births in our sample, this is equivalent to assuming that conception occurred nine months before the reported date of birth.

We define our municipality-level treatment intensity measure as the rate of female SABER 11 test takers in 2014 who are eligible for SPP in each municipality.<sup>9</sup> We then separate the sample at the median, the top half representing the *treatment municipalities* and the bottom half representing the *comparison municipalities*. We do not observe the exact SISBEN score of the students in the SABER 11 data and, therefore, their precise eligibility on the SISBEN margin. However, students report if they are categorized as SISBEN level 1 or 2. A SISBEN level of 1 or 2 is roughly equivalent to being eligible for SPP on the SISBEN margin, whereas students with higher SISBEN levels or not categorized are ineligible. On the SABER 11 margin, we determine students' eligibility using their test scores and the SPP threshold established by the government for 2014. For these students, the SPP program was announced after they had taken the SABER 11 exam. Thus, our eligibility rates avoid possible endogenous responses to the announcement of the program or its eligibility thresholds.

We attempt to assess the validity of our treatment intensity measure by estimating whether it is associated with an increase in SABER 11 test scores after SPP is introduced. This is essentially testing whether we can replicate the results from [Bernal and Penney \(2019\)](#) and [Laajaj et al. \(2022\)](#) using our treatment measure. We use individual-level data on female SABER 11 test takers between 2010 and 2016 and estimate a triple difference model that compares standardized test scores of SISBEN-eligible students between treatment and comparison municipalities. See

---

<sup>9</sup>ICFES administers the SABER 11 exam in both the spring and fall semesters each year, with the vast majority of students taking the exam in the fall semester. SPP eligibility on the SABER 11 margin was based on exams taken in the fall semester. Typically, only students in a limited set of private schools whose academic calendar is synchronized with the United States take the SABER 11 exam during the first (spring) semester of the year. For example, in 2014 (the year when SPP was introduced), 95.6 percent of the test takers took the test in the second (fall) semester.

[Appendix B](#) for a full description of this analysis. Consistent with the existing evidence, we find that, after the introduction of SPP, SABER 11 test scores increased in treatment municipalities for SISBEN-eligible students by about 0.03 standard deviations, relative to comparison municipalities. These findings support the notion that our treatment intensity measure is adequately capturing the mechanisms underlying the introduction of SPP.

### 3.3 Analytic sample and summary statistics

We restrict our sample to municipalities (i) with at least one observed birth from each age group we use in our empirical analysis for all years between 2008 to 2020 and (ii) with SABER 11 information in 2014. By doing this, we drop extremely small municipalities from our sample. Our results are not sensitive to the exclusion of these municipalities. Our final analysis sample consists of a balanced panel of 1,067 municipalities for conception years 2008-2019 (out of 1,122 in the country and 1,105 with SABER 11 information in 2014).

[Table 1](#) displays means and standard deviations of SPP eligibility rates in 2014, weighted by the number of female students in each municipality, for both treatment and comparison municipalities. Comparison municipalities had about 20 fewer SPP eligible students per 1,000 female students in 2014. [Figure A3](#) plots the full distribution of SPP eligibility rates for the municipalities in our sample. About 36 percent of municipalities had zero SPP eligible female students in 2014. We do not interpret these municipalities as completely “untreated.” Since, as [Londoño-Vélez et al. \(2020\)](#) and [Laajaj et al. \(2022\)](#) document, SPP had effects on students throughout the distribution of students’ achievement, our view is that students did not need to be eligible for SPP to be affected by the introduction of the program. We use eligibility rates (in 2014) to characterize municipalities as more or less affected by SPP.

[Figure A4](#) visualizes the municipality-level variation in the discrete version of SPP eligibility rates in 2014. While there are some clusters of treatment municipalities at a local level, there are treatment and comparison municipalities in every region of Colombia. [Table A1](#) suggests that these two groups of municipalities were different, on average, in terms of pre-SPP characteristics. For example, treatment municipalities have larger populations, lower poverty levels, and higher secondary school enrollment rates. Importantly, since our identification strategy relies on an assumption of parallel fertility rate trends in absence of SPP, these differences do

**Table 1.** Summary statistics of key variables.

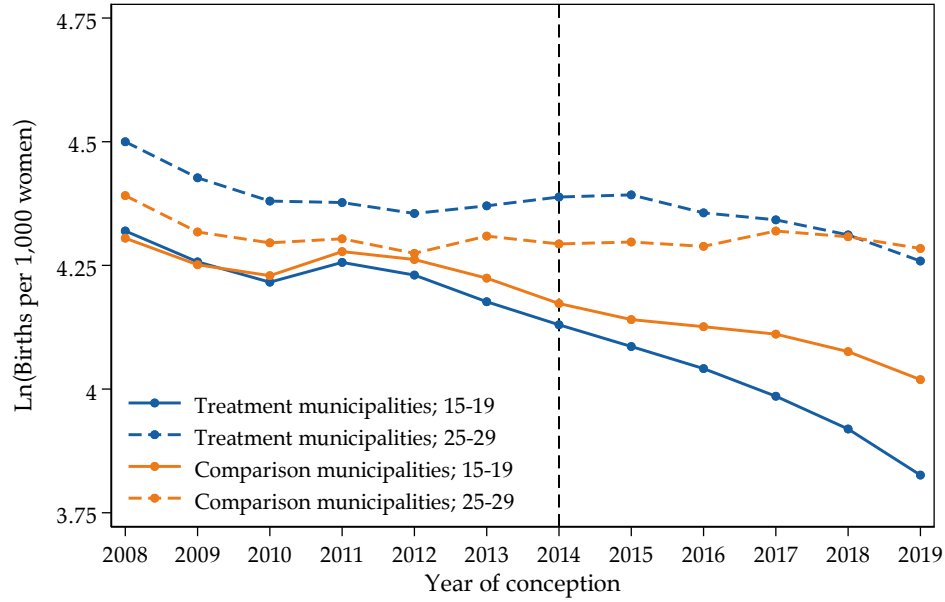
	Treatment municipalities	Comparison municipalities	All
<i>Pre-SPP births per 1,000 women (2008-2013)</i>			
Age 15-19	72.1 (20.2)	77.1 (29.6)	73.6 (23.5)
Age 25-29	84.0 (20.8)	80.2 (27.7)	83.0 (23.0)
<i>SPP eligibility rates (2014)</i>			
Per 1,000 female students	26.8 (15.5)	6.6 ( 4.8)	21.6 (16.2)
Number of municipalities	541	526	1,067

*Notes:* This table shows means and standard deviations (in parentheses) for age-specific municipality-level birth rates and SPP eligibility rates between treatment and comparison municipalities. Birth rates are averages from 2008 to 2013 and are weighted using each municipality's annual age-specific female population. SPP eligibility rates are from 2014, the first cohort of students eligible for SPP, and are averaged using the number of female students in each municipality as weights. Treatment municipalities are above the median in female eligibility rates for SPP in 2014, while comparison municipalities are below the median.

not invalidate our empirical strategy. [Figure A4](#) also shows the municipalities that had at least one SPP-eligible higher education institution. There were 15 municipalities with an SPP-eligible institution in 2014. This increased to 20 in 2016 and 21 in 2017.

[Table 1](#) also displays means of municipality-level fertility rates during the pre-SPP period for both 15-19 year olds and 25-29 year olds, weighted by the annual age-specific female population. These are the two age groups we use in our triple difference empirical strategy, which we describe in detail in the next section. Compared to the comparison municipalities, treatment municipalities have slightly lower fertility rates for women aged 15-19, but slightly higher fertility rates for women aged 25-29. [Figure A5](#) plots the complete distribution of adolescent fertility rates, which shows a substantial amount of overlap between the distributions of treatment and comparison municipalities. Finally, [Figure 1](#) shows the raw trends in average log fertility rates between age groups for both treatment and comparison municipalities. This figure mimics our empirical approaches discussed in the next section.

**Figure 1.** Trends in fertility rates for 15-19 year olds and 25-29 year olds by treatment and comparison municipalities.



*Notes:* This figure plots trends in average log fertility rates between age groups for both treatment and comparison municipalities. Treatment municipalities are above the median in female eligibility rates for SPP in 2014, while comparison municipalities are below the median. The averages weight municipalities by the annual population of women in each municipality and age group.

## 4 Empirical analysis

To estimate the effect of SPP on teen fertility, we follow difference-in-differences approaches. Our designs exploit variation in the share of female students eligible for the program across municipalities. Our preferred triple difference approach also leverages the fact that the introduction of SPP did not affect the education and fertility decisions of older women not targeted by the program. This section first describes the identifying assumptions behind our research designs, and then presents the estimation procedures.

### 4.1 Identification strategies

We use two approaches to estimate the impact of SPP on the adolescent fertility rate. First, we estimate a simple difference-in-differences model that compares the fertility rate of 15-19 year olds before and after the introduction of SPP between municipalities with different eligibility rates for the program. The identifying

assumption underlying this approach is that the trends in birth rates observed in comparison municipalities provides a good counterfactual for the trends in treatment municipalities in the absence of SPP.

Our second and preferred approach is a triple difference model that additionally uses women aged 25-29 as a within-municipality comparison group. We choose 25-29 year olds as our within-municipality comparison group because it is the group closest in age to the 15-19 year olds that is likely not affected by the introduction of SPP. The 25-29 year old group cannot be SPP beneficiaries and most are likely past their college-going years. The 20-24 group is partially affected by the introduction of SPP during our sample period given the nature of our birth records data and is also more likely to be affected by the general equilibrium effects of SPP on the higher education market.

The identifying assumption for this triple difference design is that in the absence of the policy, the differentials in fertility outcomes between 15-19 and 25-29 years old in municipalities with higher SPP eligibility rates (treatment municipalities) would have evolved similarly to these differentials in municipalities with lower SPP eligibility rates (comparison municipalities). This is the usual parallel trends assumption underlying double difference-in-differences designs applied to the triple difference case. We provide evidence in support of this assumption when we discuss our main results in [section 5](#).

We prefer the triple difference approach because it better mitigates potential bias coming from unobserved, *time-varying* heterogeneity across municipalities. By including a within-municipality comparison group, the triple difference approach accounts for municipality-specific factors that might coincide with the introduction of SPP that the simple difference-in-differences approach cannot account for. While we provide empirical support that the parallel trends assumptions are satisfied in both the difference-in-differences and triple difference approaches, we view the identifying assumptions in the triple difference approach as more theoretically plausible.

The main identification threat to the preferred triple difference strategy is the existence of other confounding events or policies that could have differentially affected the fertility rate of women in different age groups *and* are also correlated with our municipality-level treatment variable. We provide evidence that other majors events that occurred in the country around 2014 cannot explain our results in [section 6](#).



## 4.2 Estimation approaches

We implement the difference-in-differences design with an event-study specification that we estimate by Ordinary Least Squares (OLS) using the sample of 15-19 year olds. More specifically, we use the following specification:

$$Y_{mt} = SPP_m^* \times \sum_{\substack{\tau=-7 \\ \tau \neq -1}}^4 \alpha_\tau \mathbb{1}[t = \tau] + \theta_m + \theta_{d(m)t} + \epsilon_{mt}, \quad (1)$$

where  $Y_{mt}$  is the log of the teen fertility rate in municipality  $m$  and year  $t \in [-7, 4]$ , which is measured in years relative to October 1, 2014, the date in which SPP was first announced. In Equation 1,  $SPP_m^*$  denotes treatment and comparison municipalities and is defined as  $SPP_m^* = \mathbb{1}[SPP_m > \text{median}(SPP_m)]$  with  $SPP_m$  being the rate of female students eligible for the program in a given municipality in 2014.<sup>10</sup>  $\theta_m$  and  $\theta_{d(m)t}$  are municipality fixed effects and year fixed effects that we allow to be department-specific, respectively.<sup>11</sup> Finally,  $\epsilon_{mt}$  is an error term.

In a similar fashion, we implement our preferred triple difference design with the following specification that we also estimate by OLS but using the sample of 15-19 and 25-29 year olds:

$$Y_{amt} = Teen_a \times SPP_m^* \times \sum_{\substack{\tau=-7 \\ \tau \neq -1}}^4 \beta_\tau \mathbb{1}[t = \tau] + \gamma_{am} + \gamma_{mt} + \gamma_{ad(m)t} + \epsilon_{amt}, \quad (2)$$

where  $Y_{amt}$  is the log fertility rate of age group  $a \in \{15-19, 25-29\}$  in municipality  $m$  and year  $t \in [-7, 4]$ , again measured in years relative to the announcement of SPP.  $Teen_a$  is an indicator for age group defined as  $Teen_a = \mathbb{1}[a = 15-19]$ .  $\gamma_{am}$  and  $\gamma_{mt}$  are age group by municipality and municipality by year fixed effects, respectively.  $\gamma_{ad(m)t}$  are age group by year fixed effects that we also allow to be department-specific. Finally,  $\epsilon_{amt}$  is an error term.

In Equation 1 and Equation 2,  $\alpha_{\tau, t \geq 0}$  and  $\beta_{\tau, t \geq 0}$  represent the average treatment

<sup>10</sup>We show in section 6 that our results are robust to different estimation approaches, including using a continuous version of our treatment intensity measure, excluding capital cities from the sample, using all available municipalities with the inverse hyperbolic sine transformation of fertility rates as the outcome, and using fertility rates in levels as the outcome estimated by either OLS or Poisson models.

<sup>11</sup>Departments in Colombia are similar to states in the United States. A group of municipalities forms each department.

effect of SPP on teen fertility at time  $t = \tau$  after the introduction of the program. For the estimation, we use the year before the announcement of SPP ( $t = -1$ ) as our reference period. For Equation 2, and in the standard way for triple difference specifications, we include three two-way interactions between age groups, municipalities, and years. The age group by municipality fixed effects ( $\gamma_{am}$ ) control for time-invariant, municipality-specific factors (both observed and unobserved) that affect fertility rates and that are potentially different by age groups. The municipality by year fixed effects ( $\gamma_{mt}$ ) control for municipality-specific trends in fertility rates common to all age groups. Finally, the age group by department year effects ( $\gamma_{ad(m)t}$ ) account for age-specific trends in fertility and arbitrary shocks to fertility that are common to all municipalities in a given region. As mentioned earlier, the remaining and identifying source of variation we leverage is the *differential* effect that SPP had on the adolescent fertility rate in the treatment municipalities (relative to the comparison municipalities in the same department).

To summarize the event-study estimates of SPP's effects in a single estimate, we also estimate a version of Equation 1 and Equation 2 that replaces the year indicators with a single post-SPP indicator variable. Specifically, we estimate the following two equations by OLS:

$$Y_{mt} = \alpha (SPP_m^* \times Post_t) + \theta_m + \theta_{d(m)t} + \xi_{mt}, \text{ and} \quad (3)$$

$$Y_{amt} = \beta (Teen_a \times SPP_m^* \times Post_t) + \gamma_{am} + \gamma_{mt} + \gamma_{ad(m)t} + u_{amt}, \quad (4)$$

where  $Post_t = \mathbb{1}[t \geq 0]$  and everything else is defined as in Equation 1 and Equation 2. In Equation 3 and Equation 4,  $\alpha$  and  $\beta$  are the summary difference-in-differences and triple difference parameters across all post-SPP years, respectively.

In all our regressions, we cluster the standard errors at the municipality level and weight each cell by the population of women in each municipality and age group. Excluding weights would give each municipality equal weight in the regression by default, which we argue is inappropriate in our context, particularly due to the large population variation across municipalities in Colombia.<sup>12</sup>

---

<sup>12</sup>If we were able to run an individual-level regression, there would naturally be many more observations from municipalities with larger populations. Our use of population weights essentially approximates the use of individual-level data.

## 5 Results

This section reports and discusses our results. We begin by presenting our difference-in-differences estimates of the teen fertility impacts of *Ser Pilo Paga*. We then present our preferred estimates which use a triple difference approach. Finally, we present supplemental analyses to explore the mechanisms and heterogeneity of the main results.

### 5.1 Difference-in-differences estimates

Figure 2 shows the event study estimates of  $\alpha_\tau$  from the simple difference-in-differences specification in Equation 1 that compares adolescent fertility rates across municipalities with higher and lower initial SPP eligibility rates, before and after the introduction of the program. None of the coefficients in the pre-period are statistically significant at the 5 percent level. Following Borusyak, Jaravel and Spiess (2021), we conduct a more formal test of pre-period trends by estimating the difference-in-differences specification using only the set of untreated observations and running a joint F-test of these coefficients. With a  $p$ -value of 0.711, this test cannot reject the null hypothesis that the pre-period coefficients are jointly equal to zero.

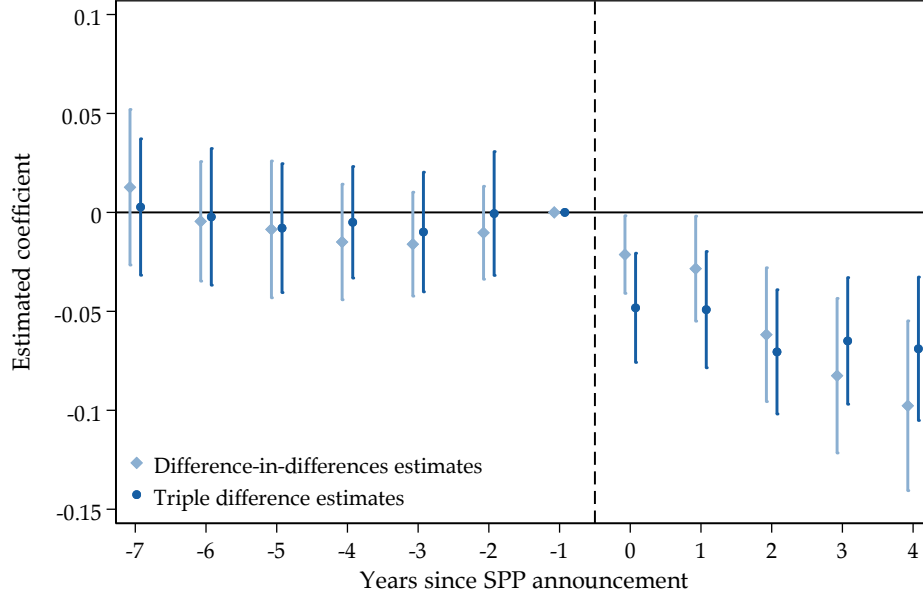
After the introduction of SPP, fertility rate trends between treatment and comparison municipalities change significantly, with teen fertility rates decreasing more in treatment municipalities. All post-period coefficients are negative and statistically significant at the 5 percent level. Column 1 of Table 2 presents the summary estimate from the difference-in-differences design which implies that after SPP was introduced fertility rates of women aged 15-19 in treatment municipalities decreased by about 5.2 percent relative to comparison municipalities.<sup>13</sup>

In support of our use of initial SPP eligibility rates to characterize municipalities as more or less affected by SPP, we show that the teen fertility impacts of SPP are larger in municipalities with higher eligibility rates. To do this, we replace the indicator for being above the median in SPP eligibility with indicators for the *quartile* of SPP eligibility rates. Municipalities in the first (lowest) quartile of eligibility become the reference group in the regression, which we note all had zero female students eligible for SPP in 2014. Column 3 of Table 2 reports the summary results, but we also show the event-study equivalent of this specification

---

<sup>13</sup>The exact percentage changes implied by our estimates are given by  $100 \times [\exp(\hat{\alpha}) - 1]$ .

**Figure 2.** Event study estimates.



*Notes:* This figure plots the difference-in-differences estimates of  $\alpha_\tau$  from Equation 1 and the triple difference event study estimates of  $\beta_\tau$  from Equation 2. The dots and diamonds represent the estimated coefficients and the vertical lines represent 95 percent confidence intervals. All estimates are weighted by the annual population of women in each municipality and age group. Standard errors are clustered at the municipality level.

in Figure A6, Panel (a).

Relative to municipalities in the 1st quartile of SPP eligibility, we estimate that 2nd quartile municipalities experienced a 3.1 percent decrease (not statistically significant) in adolescent fertility. Meanwhile, we estimate that the 3rd and 4th quartile municipalities experienced a 6.9 and 7.9 percent decrease, respectively. The estimates for the 3rd and 4th quartile municipalities are both statistically different from the estimate for the 2nd quartile at the 10 percent significance level. These results highlight that our estimates are not dependent upon how we split municipalities into treatment and comparison groups.

## 5.2 Triple difference estimates

Figure 2 also displays our preferred estimates of the fertility impacts of SPP ( $\beta_\tau$ ) using the triple difference event-study specification in Equation 2. Similar to the difference-in-differences results, the pre-period coefficients are again close to zero and not statistically significant. An F-test of the null hypothesis that the pre-period

**Table 2.** Summary difference-in-differences and triple difference estimates.

	Log fertility rate			
	(1)	(2)	(3)	(4)
$SPP \times Post$	-0.052*** (0.016)			
$Teen \times SPP \times Post$		-0.057*** (0.011)		
$SPP_{Quartile} \times Post$				
1st quartile			[Reference]	
2nd quartile			-0.031 (0.027)	
3rd quartile			-0.069*** (0.018)	
4th quartile			-0.079*** (0.021)	
$Teen \times SPP_{Quartile} \times Post$				
1st quartile				[Reference]
2nd quartile				-0.040** (0.019)
3rd quartile				-0.079*** (0.015)
4th quartile				-0.091*** (0.018)
Observations	12,804	25,608	12,804	25,608
Treatment municipalities	541	541	–	–
Comparison municipalities	526	526	–	–
Pre-trends test $p$ -value	0.711	0.985	–	–

*Notes:* This table presents summary difference-in-differences and triple difference estimates using Equation 3 and Equation 4, respectively. Column 1 presents the difference-in-differences estimates. Column 2 presents the main triple difference estimates. Columns 3 and 4 use indicators for the quartile of SPP eligibility rates instead of an indicator for being above the median in SPP eligibility. The reference group here are municipalities in the first (lowest) quartile of SPP eligibility. All estimates are weighted by the annual population of women in each municipality and age group. Standard errors are clustered at the municipality level and presented in parentheses (\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).

coefficients are jointly zero using only the set of untreated observations produces a  $p$ -value of 0.985 (Borusyak et al., 2021). This provides empirical support in favor of the parallel trends assumption of the triple difference research design.

Starting in the first year after the introduction of SPP, there is a distinct decrease in the fertility rate of women aged 15-19 in treatment municipalities relative to comparison municipalities. All post-SPP coefficients are negative and statistically significant at the 5 percent level. We report the summary estimates using [Equation 4](#) in column 2 of [Table 2](#). The triple difference estimate in this specification indicates that SPP reduced fertility rates of women aged 15-19 in treatment municipalities by 5.7 percent relative to comparison municipalities.<sup>14</sup> This effect accounts for approximately one-fourth of the overall decrease in adolescent fertility observed in Colombia in the years following the announcement of SPP.

To illustrate even more clearly that the timing of the changes in fertility rate trends aligns with the timing of the introduction of SPP, we estimate [Equation 2](#) using a quarterly-level data set.<sup>15</sup> SPP was announced on October 1st in 2014. Thus, for SPP to be responsible for the relative decrease in adolescent fertility that we observe, we would expect quarterly-level effects of SPP to be apparent starting exactly in the fourth quarter of 2014. Indeed, this is what we observe from the quarterly-level estimates in [Figure 3](#). Moreover, the summary estimate at the quarterly level is identical to our main estimate at the annual level.

Next, we show that the estimated effect of SPP increases (in magnitude) as a municipality's initial SPP eligibility rate increases. To show this, we again use indicators for the *quartile* of SPP eligibility rates instead of an indicator for being above the median in SPP eligibility. The reference group now becomes municipalities in the first (lowest) quartile of SPP eligibility. The summary results are reported in [Figure 4](#) and in [Table 2](#). The estimates from the event-study equivalent of this specification are shown in [Figure A6](#), Panel (b). Relative to municipalities in the 1st quartile of eligibility, 2nd quartile municipalities experienced a 4 percent decrease in fertility rates, while the 3rd and 4th quartile municipalities experienced a 7.9 and 9.1 percent decrease, respectively. The estimates for each quartile are statistically different from the 1st quartile at the 5 percent level (1 percent for the 3rd and 4th quartiles), and the estimates for the 3rd and 4th quartiles are statistically different from the estimate for the 2nd quartile at the 1 percent level.

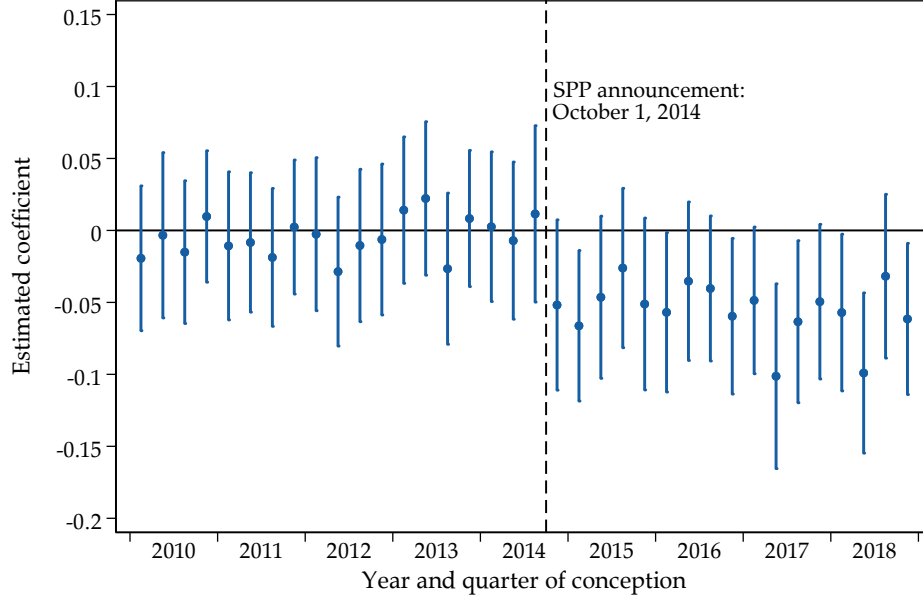
These results highlight two important points. First, our main estimates are

<sup>14</sup>The exact percentage changes implied by our estimates are, again, given by  $100 \times [\exp(\hat{\beta}) - 1]$ .

<sup>15</sup>Since there are some municipalities that have zero births in some cells at the quarterly level, we use the inverse hyperbolic sine transformation of the birth rate as the outcome instead of the log transformation. The interpretation of the estimated coefficients from this model is similar to our main log model. See [Bellemare and Wichman \(2020\)](#) for a discussion and formal derivation. In [Figure A7](#), we show our results hold using a Poisson model.



**Figure 3.** Triple difference event study estimates using quarterly data.



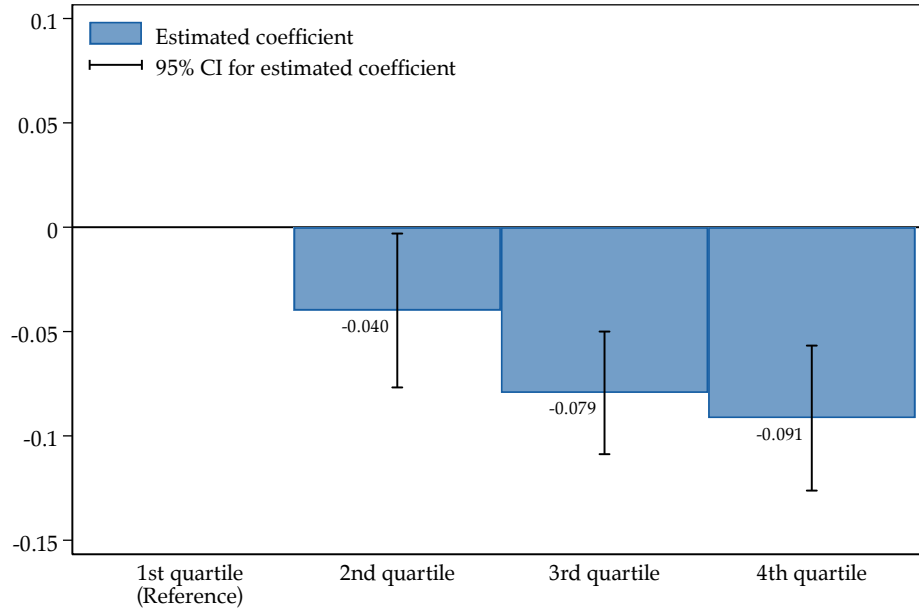
*Notes:* This figure plots the triple difference event study estimates of  $\beta_\tau$  from Equation 2 using quarterly data instead of annual data. In contrast to the annual specification we use the inverse hyperbolic sine transformation of the birth rate as the outcome instead of the log transformation, because at the quarterly level, some municipalities have zero births in some of the cells. The quarters in 2008 are the reference period. Only estimates for four years around 2014 are plotted. The dots represent the estimated coefficients and the vertical lines represent 95 percent confidence intervals. All estimates are weighted by the annual population of women in each municipality and age group. Standard errors are clustered at the municipality level.

not purely the product of a fortuitous split of municipalities at the median of SPP eligibility. Second, our estimates represent relative effects, not total effects. We use the initial SPP eligibility rate to compare municipalities that are plausibly more and less affected by this nationally implemented program. That the estimated effects for municipalities in the 3rd and 4th quartile of eligibility are larger in magnitude than the main estimate suggests that the *total* effects of SPP on adolescent fertility rates in Colombia are likely also larger than what our main estimates indicate.

### 5.3 Mechanisms

In this section, we aim to assess the extent to which our estimates are explained by direct effects—where incapacitation or income effects drive the results through receipt of the SPP scholarship—or by indirect effects—where results are driven by behavioral responses before being able to receive the scholarship, such as motiva-

**Figure 4.** Summary triple difference estimates by quartile of initial SPP eligibility.



*Notes:* This figure plots summary triple difference estimates of  $\beta$  from Equation 4 using indicators for the quartile of SPP eligibility rates instead of an indicator for being above the median in SPP eligibility. All estimates are weighted by the population of women from each age group in each municipality. Standard errors are clustered at the municipality level.

tional effects, changing opportunity costs, and peer effects. Since our data limits us to estimate effects on women within an age range of 15 to 19 years old, our main estimates may only reflect the direct effects of students receiving SPP, going to college, and reducing (or delaying) childbearing that would have occurred during their teen college-age years (i.e., 18 or 19 years old).

To analyze these mechanisms, we first assess the extent to which the direct effects explain our results. To do this, we begin by applying our estimates to the annual post-SPP birth rates and population counts to calculate the number of fewer births implied by our results. We then compare this to the actual number of female SPP recipients from 2015 to 2018. Our estimates imply that there were 23,878 fewer births to teenage mothers as a result of the introduction of SPP. This is larger than the 17,149 female SPP scholarship recipients during the same period. These calculations suggest that incapacitation or income effects from receiving the SPP scholarship cannot fully explain the effects we observe.

Next, returning to Figure 3, the timing of the effects we observe is also informative for the mechanisms driving the results. If incapacitation effects or income

effects primarily explain our results, we would expect to begin to see effects a few quarters after SPP's announcement when college enrollment began and scholarship funds were disbursed for the first cohort of SPP beneficiaries. However, we observe effects in the first quarter after SPP's announcement, which suggests that incapacitation or income effects are likely not the primary source of our estimated effects.

Finally, we use the more granular information in our data about the age of the father to learn more about what ages the fertility effects of SPP are concentrated. While our data only include mother's age in a range of years, the data does include father's age in integer years. Since the onset of SPP can have similar effects on the college opportunities for young men, we assess SPP's effect on teen fatherhood rates in [Table 3](#), using the difference-in-differences specification from [Equation 3](#) except with the age-specific fatherhood rates per 1,000 young men as the outcome.<sup>16</sup> The share of all fathers who are teens is lower than the share of mothers who are teens, and SPP's effects may not be the same between adolescent men and women. Nevertheless, the more disaggregated age-specific effects for adolescent men may still be informative for assessing which ages the fertility effects of SPP are concentrated and thus the class of mechanisms involved. For instance, larger effects for younger teenagers would suggest an important role for the indirect effects we previously described.

For all teenage men aged 15-19, we estimate that teenage fatherhood rates decreased by 6.2 percent in treatment municipalities relative to comparison municipalities after SPP was introduced, suggesting SPP had similar effects on parenthood between teenage men and women. Estimating effects separately by age, we find the smallest effects among 19 years olds, even though this group accounts for the largest share of births to teen fathers. Importantly, we find larger effects on fatherhood rates for teenagers aged 15-17, who are likely to have not yet received an SPP scholarship. Meanwhile, we find the largest effects among 18 year olds. 18 years olds can be high school graduates (and thus potential SPP scholarship recipients), but nearly 17 percent of SABER 11 test takers are 18 years old so many likely have not yet graduated high school.

Together, the results and observations above suggest that our main estimates

---

<sup>16</sup>Since there are some municipalities that have zero births to teen fathers in some cells (particularly at young ages), we use the inverse hyperbolic sine transformation of the fatherhood rate as the outcome instead of the log transformation. In [Table A2](#), we show our results are qualitatively similar using a Poisson model.

**Table 3.** Summary difference-in-differences estimates for teen fatherhood.

Age group:	IHS births per 1,000 men			
	All teens (15-19)	15-17	18	19
	(1)	(2)	(3)	(4)
$SPP \times Post$	-0.062*** (0.022)	-0.063** (0.032)	-0.090*** (0.027)	-0.044* (0.024)
Observations	12,804	12,804	12,804	12,804
Treatment municipalities	541	541	541	541
Comparison municipalities	526	526	526	526
Pre-trends testing $p$ -value	0.492	0.181	0.335	0.504
Pre-SPP share of teen fathers	100	25.7	33.6	40.7

*Notes:* This table presents summary difference-in-differences estimates using [Equation 3](#) with the inverse hyperbolic sine (IHS) transformation of the number of births per 1,000 men in each age group as the outcome, because some municipalities in our main sample have zero births for certain male age groups. Column 1 presents the estimate for all male teens (15-19 years old). Column 2 shows the results for male teens 15-17 years of age. Columns 3 and 4 present the estimates for men 18 and 19 years of age, respectively. All estimates are weighted by the annual population of men in each municipality and age group. Standard errors are clustered at the municipality level and presented in parentheses (\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).

largely consist of indirect effects, where new college-going opportunities created by SPP influenced teen fertility decisions for students before they were even able to benefit directly from the program. This is consistent with [Laajaj et al. \(2022\)](#) who find that SPP caused motivational effects on low-income students resulting in increased 9th grade test scores, years before eligibility for SPP is determined, throughout most of the test score distribution. These effects on 9th grade test scores materialized within the first year after SPP was announced, which supports the plausibility of the immediacy of our observed effects.

#### 5.4 Heterogeneity analysis

In this section, we explore how the fertility effects of SPP vary by pre-policy municipality characteristics. Specifically, in line with the idea that SPP represented a shock to post-secondary educational opportunities and opportunities for social mobility more broadly, we examine whether the program had different impacts based on the municipality's level of income inequality and students' college-going expectations before the policy's implementation.

We use a Gini coefficient as a proxy of income inequality, measured in 2005 at the municipality level. We then separately estimate [Equation 4](#) for those below the

median (lower inequality) and above the median (higher inequality). Following the discussion in [Kearney and Levine \(2014\)](#), we expect the fertility effects of SPP to be larger in magnitude in municipalities that ex-ante had higher levels of inequality and arguably more limited chances of economic mobility.

We present the results of this analysis in Panel A of [Table 4](#). Using subsets of municipalities necessarily changes the composition of treatment and comparison units used in the estimation. Thus, ensuring that the parallel trends assumption is still reasonably satisfied is important. Using the pre-trends test described in [section 5](#), we cannot reject the null hypothesis that all pre-period coefficients for both subsets of municipalities are zero. Column 1 reproduces our core results for the overall sample of municipalities for which we have a measure of inequality (1,009 out of 1,067 in the main estimation sample). The estimated reduction in teen fertility in municipalities with higher levels of income inequality is 2.5 percentage points greater than in those with lower inequality (column 3 versus column 2). In fact, although negative, the estimated change in fertility in lower inequality municipalities is not statistically significant. However, we cannot reject the null hypothesis that these two coefficients are the same using a two-sided F-test ( $p$ -value = 0.295).

To assess the heterogeneous impacts of SPP by expectations of attending college, we utilize survey responses on the higher education expectations of a 10 percent random sample of SABER 11 test takers in 2013 and 2014 (pre-SPP). We use low-income (SISBEN 1 and 2) female students' responses to a question that asks about how likely they are to enroll in a higher education program immediately after finishing high school. At the municipality level, we calculate the share of respondents who indicate they are likely or highly likely to attend college. To guarantee that we have a reasonable number of students in each municipality, we only include municipalities for which we observe at least 5 percent of the test takers and at least ten students in this analysis. This limits the estimation sample to 658 municipalities.<sup>17</sup> Finally, we group municipalities by whether they are above or below the median share of students who meet this criteria and estimate [Equation 4](#) separately for these two subsets of municipalities.

---

<sup>17</sup>Results are qualitatively similar when using the slightly more stringent requirement of keeping municipalities where we observe at least 10 percent of the students.

**Table 4.** Triple difference estimates by group of municipalities.

Log fertility rate			
<b>Panel A. Baseline income inequality</b>			
Municipalities:	All	Below median	Above median
	(1)	(2)	(3)
<i>Teen × SPP × Post</i>	-0.054*** (0.011)	-0.029 (0.019)	-0.054*** (0.015)
Observations	24,216	11,928	13,680
Treatment municipalities	511	293	248
Comparison municipalities	498	204	322
Pre-trends test <i>p</i> -value	0.966	0.486	0.833
<b>Panel B. Baseline college-going expectations</b>			
Municipalities:	All	Below median	Above median
	(1)	(2)	(3)
<i>Teen × SPP × Post</i>	-0.053*** (0.013)	-0.068*** (0.023)	-0.039*** (0.014)
Observations	15,792	7,896	7,896
Treatment municipalities	389	179	210
Comparison municipalities	269	150	119
Pre-trends test <i>p</i> -value	0.949	0.276	0.350
<b>Panel C. Baseline share of teen births from adolescent fathers</b>			
Municipalities:	All	Below median	Above median
	(1)	(2)	(3)
<i>Teen × SPP × Post</i>	-0.057*** (0.011)	-0.002 (0.016)	-0.062*** (0.016)
Observations	25,608	12,744	12,864
Treatment municipalities	541	234	307
Comparison municipalities	526	297	229
Pre-trends test <i>p</i> -value	0.985	0.532	0.915

*Notes:* This table presents summary triple difference estimates using Equation 4 for a subgroup of municipalities indicated in each column. In Panel A, income inequality is measured by the 2005 municipal Gini obtained from CEDE's municipal panel. The Gini is not available for some of the municipalities. In Panel B, we classify municipalities according to the pre-SPP measure of college-going expectations described in the text. This measure is not available for some of the municipalities. In Panel C, the share of teen births from adolescent fathers corresponds to the average of 2008-2013. Two-sided *p*-value from an *F*-test of equality,  $H_0(2) = (3)$ : Panel A *p*-value = 0.295; Panel B *p*-value = 0.293; Panel C *p*-value = 0.009. All estimates are weighted by the annual population of women in each municipality and age group. Standard errors are clustered at the municipality level and presented in parentheses. (\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).



The triple difference coefficients are presented in Panel B of [Table 4](#). First, we cannot reject the null that all pre-period coefficients for both subsets of municipalities are zero. Second, the overall coefficient in column 1 is similar to the one in our main analysis. We then estimate a reduction in fertility almost 3 percentage points larger (in magnitude) in municipalities where, before SPP, the expectations of immediate enrollment in higher education after high school graduation were smaller (column 2 vs column 3). This time we also cannot reject the null hypothesis that the two coefficients are the same using a two-sided F-test ( $p$ -value = 0.293).

We explore an additional source of heterogeneity. In Panel C of [Table 4](#), we present results by splitting up municipalities according to the pre-SPP share of teen births to teen fathers. We find that the relative reduction in teen fertility is driven by municipalities where female teenagers tend to have children with other teenagers. In conjunction with the results presented in [Table 3](#), this suggests a reinforcement of incentives to avoid parenthood when potential fathers also face a positive shock to college-going opportunities, or alternatively, an increase in bargaining power for female teenagers in relationships with peers of similar age.

Given the previous results, we complement our heterogeneity analysis by presenting in [Figure A8](#) a slightly more detailed version of the results in [Table 4](#). There, we divide municipalities in deciles according to the level of each characteristic discussed before. After recasting the triple difference as a simple difference-in-differences by calculating the within-municipality difference between the (log) fertility rate of teens and non-teens, we use [Equation 3](#) to estimate the triple difference reduction in fertility for each decile. Following [Muralidharan and Prakash \(2017\)](#), we then plot these estimates using a lowess regression smoothing of the triple difference coefficients on the average level of the characteristic for the deciles. By thoroughly exploiting the variation in the measures of income inequality and college-going expectations, this exercise helps us confirm the takeaways from [Table 4](#). Together, this set of results support the idea that SPP represented a bigger shock to economic opportunity in municipalities with greater ex-ante income inequality and with greater perceived limited opportunities for tertiary education.

## 6 Robustness

This section overviews a series of analyses that assess the robustness of our preferred triple difference results to alternative definitions of treatment and compari-

son units, possibly confounding events, and other sensitivity checks.

### 6.1 Alternative definitions of treatment and comparison units

We showed in [section 5](#) that our main results are not dependent upon a convenient splitting of the sample into treatment and comparison municipalities at the median of SPP eligibility. [Table A3](#) also shows that these results hold when using a continuous version of our treatment intensity measure, excluding capital cities from the sample, using the inverse hyperbolic sine (IHS) transformation of fertility rates as the outcome, and using fertility rates in levels as the outcome (estimated with either OLS or Poisson models). Both the IHS transformation and using the rates in levels allow us to keep the few municipalities with zero birth counts in our estimation. We build on this here by exploring additional definitions of treatment and control units using alternative sources of variation. Combined, these analyses show that our main results are not solely dependent upon our use of initial municipality-level SPP eligibility rates, or using women aged 25-29 as a comparison group.

First, we estimate a triple difference model similar to [Equation 2](#) which replaces  $SPP_m^*$  with an indicator for whether a municipality is below the median distance (closer) to the nearest SPP-eligible institution.<sup>18</sup> Thus, this empirical approach does not rely at all upon a municipality's initial SPP eligibility rate. We plot the event study estimates from this analysis in [Figure A9](#). The results show that the age group differentials in log birth rates trend similarly, with no clear pattern, between municipalities that are closer and farther from SPP-eligible institutions from 2008 through 2014. After the introduction of SPP, however, these trends begin to diverge, with birth rates declining in municipalities closer to SPP-eligible institutions relative to those farther away. The summary estimate, shown in column 2 of [Table A4](#), implies that municipalities closer to SPP-eligible institutions experienced a 9.6 percent decrease in teen fertility rates after SPP was introduced relative to municipalities that are farther away.

Second, we estimate a triple difference model that does not rely on using women aged 25-29 as the within-municipality control group. Here, we instead use young women whose birth record indicates their highest level of education completed being seventh grade or less as a comparison group. We replace  $Teen_a$  in [Equation 2](#)

---

<sup>18</sup>See [Figure A4](#) for a map of where SPP-eligible institutions are located in Colombia. We calculate distance to institutions using only the initial set of institutions that were SPP-eligible at the start of the program.

with an indicator for having completed eighth grade or more using only data on births to women aged 15-19.<sup>19</sup> The intuition is that any woman aged 15-19 whose highest grade completed is seventh grade or less is likely to have already dropped out of school. As school dropouts, these women are likely to be less affected by the introduction of SPP than women aged 15-19 who have completed eighth grade or higher.

The results of this approach are reported in [Figure A10](#). Panel (a) shows, separately for treatment and comparison municipalities, the trends in the differential in log number of births between women age 15-19 who have completed eighth grade or higher and those who have completed only seventh grade or lower.<sup>20</sup> The plot shows that births to women with eighth grade or higher education is increasing over time relative to women with less than eighth grade education. But, these differentials trend very similarly between treatment and comparison municipalities from 2008 to 2014. Starting in 2015, however, the trends begin to diverge between these groups of municipalities, where the trends in grade level differentials flatten in treatment municipalities but continue to increase in comparison municipalities.

Panel (b) plots the triple difference estimates that compares the changes in the number of births to women aged 15-19 over time between those completing eighth grade or more versus seventh grade or less, and between treatment and comparison municipalities. The results mirror the trends in Panel (a): grade level differences in births between treatment and comparison municipalities trend similarly through 2014, but begin to diverge significantly after SPP is introduced in 2015. The summary estimate, presented in column 3 of [Table A4](#), implies that births to adolescent women who have completed at least eighth grade decreased by 10.7 percent in treatment municipalities relative to comparison municipalities. This analysis highlights that our results are not dependent upon using older women aged 25-29 as a comparison group.

---

<sup>19</sup>We also restructure the analysis data set to have one observation per year per municipality per grade-level (i.e., eighth grade or above, or seventh grade or less).

<sup>20</sup>We use the number of births instead of a birth rate because we cannot reliably calculate the number of young women in each municipality with above or below an eighth grade level of education. Also, we use an inverse hyperbolic sine transformation instead of a log transformation since some municipalities have zero recorded births in these year by grade-level cells.

## 6.2 Possible confounding events

Since our setting involves a single treatment time period, we are potentially vulnerable to events that happened simultaneously (or around the same time) as the introduction of SPP. Although, we note that to truly be a threat to identification, these simultaneous events would have to differentially affect women of the different age groups and be correlated with SPP eligibility rates. Nevertheless, we assess whether three events that occurred at a similar time might be driving our results: 1) the unilateral permanent ceasefire by the Revolutionary Armed Forces of Colombia (FARC, from the Spanish acronym) in December 2014 as part of the by then ongoing peace process between the guerrilla group and the Colombian government, 2) the Zika virus epidemic, which occurred from October 2015 to July 2016, and 3) the *Jornada Única* initiative, which gradually transitioned some public secondary schools that were operating half-day shifts into full school days beginning in 2015.

For each of these possibly confounding events, we re-estimate our main specification using only a subset of municipalities that were likely unaffected by the relevant event. If these events are not driving our results, we would expect to see estimates based on these subsets of municipalities that are similar to our main estimates. We provide a full description of these analyses in [Appendix C](#) and report these estimates in [Table C1](#). Indeed, we consistently estimate large and statistically significant effects of SPP in each of these subsample analyses. We conclude that these three events cannot explain the effects we observe.

## 6.3 Other sensitivity checks

Although our empirical strategy does not rely on a staggered rollout design, the recent developments in the difference-in-differences literature have documented issues with estimating difference-in-differences designs with linear regressions and fixed effects specifications even with non-staggered binary treatments, which is the case of our specifications in [subsection 4.2](#) ([Borusyak et al., 2021](#); [de Chaisemartin and D’Haultfœuille, 2022a,b](#)). We implement the imputation estimator developed by [Borusyak et al. \(2021\)](#) and the difference-in-differences (DID) estimator proposed by [de Chaisemartin and D’Haultfœuille \(2022a\)](#) that are robust to treatment effect heterogeneity to estimate the models in [Equation 2](#) and [Equation 4](#). Results using these alternative estimators are reported in [Figure A11](#) and [Table A5](#). The triple difference estimate using the imputation estimator is 8.3 percent and for the DID

estimator is 7.4 percent.

To assess whether the decline in teen fertility we observe in treatment municipalities is the result of pure chance, we perform a permutation test that randomly assigns municipalities to be treatment or comparison municipalities. We then compare our main estimate to a distribution of estimates across 5,000 randomly assigned groups of treatment municipalities. To do this, we use the randomization inference routine developed by [Heß \(2017\)](#) and the specification in [Equation 4](#). We report the results in [Figure A13](#). Reassuringly, we see that our main estimate is in the far left tail of the distribution of estimated triple difference coefficients. Also, to be sure our results are not driven by a small number of municipalities, we re-estimate our main specification while each time excluding municipalities in a single department (reported in [Figure A12](#)). We also estimate our main specification by excluding all municipalities that include a department’s capital city (reported in column 2 of [Table A3](#)). The results from each of these regressions produce estimates that are very similar to our main estimates.

Finally, we perform a placebo-in-time strategy to further support the validity of the parallel trends assumption required for our estimates to have a causal interpretation. In [Table A6](#), we use 2008-2014 data and estimate the same specification in [Equation 4](#) pretending that SPP was introduced in years 2008-2013. The overall estimated effects after each of the placebo treatment years are always statistically insignificant and close to zero. The parallel trends assumption is an assumption about counterfactuals and, therefore, untestable. We have shown robust evidence that the differentials in fertility rates between younger and older women were not trending differently among treatment and comparison municipalities before SPP, which adds support that this would have been the case during the post-period had the program not been introduced.

## 7 Conclusion and Discussion

In this paper, we study the teen fertility impacts of *Ser Pilo Paga*, Colombia’s generous college financial aid program for high-achieving, low-income students. After the 2014 introduction of the program, we find that teen fertility rates decreased by about 6 percent in municipalities more affected by SPP relative to less affected municipalities. Due to SPP being a nationally implemented policy, these estimates are necessarily relative effects. The *total* effects of SPP on teen fertility rates nationwide

are likely to be larger than the relative estimates indicate.

While our data limits us from precisely identifying the mechanisms driving our results, our analyses point to effects largely coming from behavioral responses prior to students going to college, potentially including channels such as motivational effects, increased opportunity costs, and/or peer effects. We also find larger effects of SPP in areas where the pre-SPP levels of income inequality were greater. This is consistent with [Kearney and Levine \(2014\)](#) who present empirical and theoretical evidence that suggests inequality—and the “economic hopelessness” that inequality cultivates—explains a large share of the variation in teen childbearing rates. Placing our estimates in the context of existing research on the determinants of teen fertility in similar settings, we find that SPP’s effects on teen fertility is smaller than *Familias en Acción*—Colombia’s conditional cash transfer program—([Attanasio et al., 2021](#)), but larger than a Chilean reform that lengthened the school day from half to full-day shifts ([Berthelon and Kruger, 2011](#)).

Our results suggest that increasing future economic opportunities for young women can lead to meaningful reductions in teen fertility, consistent with some of the policy considerations discussed by [Kearney and Levine \(2012\)](#) in the context of the United States. Prior to SPP, Colombia was characterized by large socio-economic gaps in college enrollment due to severe financial constraints, low access to credit, and high college tuition costs. We posit that, in countries with high inequality, college financial aid programs like SPP that decrease inequality of opportunity can have behavioral effects on teen childbearing and perhaps other outcomes. The characteristics of SPP—namely its generosity, salience, and simplicity—would seem to be important in accounting for the far-reaching impacts of the program, which is consistent with the college financial aid literature more broadly (e.g., [Bettinger, Long, Oreopoulos and Sanbonmatsu \(2012\)](#); [Dynarski, Libassi, Micheltmore and Owen \(2021\)](#)).



## References

- Aizer, Anna, Paul Devereux, and Kjell Salvanes**, “Grandparents, Moms, or Dads? Why Children of Teen Mothers Do Worse in Life,” *Journal of Human Resources*, 2022, 57 (6), 2012–2047.
- Alzúa, María Laura and Cecilia Velázquez**, “The effect of education on teenage fertility: causal evidence for Argentina,” *IZA Journal of Development and Migration*, 2017, 7 (1), 1–23.
- Attanasio, Orazio, Lina Cardona Sosa, Carlos Medina, Costas Meghir, and Christian Manuel Posso-Suárez**, “Long Term Effects of Cash Transfer Programs in Colombia,” Working Paper 29056, National Bureau of Economic Research July 2021.
- Azevedo, Joao Pedro, Marta Favara, Sarah E. Haddock, Luis F. López-Calva, Miriam Muller, and Elizaveta Perova**, *Teenage pregnancy and opportunities in Latin America and the Caribbean: On teenage fertility decisions, poverty and economic achievement*, World Bank, Washington, DC, 2012.
- Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman**, “Women’s Empowerment in Action: Evidence from a Randomized Control Trial in Africa,” *American Economic Journal: Applied Economics*, 2020, 12 (1), 210–259.
- Becker, Gary S.**, “An Economic Analysis of Fertility,” in Gary S. Becker, ed., *Demographic and Economic Change in Developed Countries*, Princeton, NJ: Princeton University Press, 1960, pp. 209–231.
- Bellemare, Marc F. and Casey J. Wichman**, “Elasticities and the Inverse Hyperbolic Sine Transformation,” *Oxford Bulletin of Economics and Statistics*, 2020, 82 (1), 50–61.
- Bernal, Gloria L. and Jeffrey Penney**, “Scholarships and Student Effort: Evidence from Colombia’s Ser Pilo Paga Program,” *Economics of Education Review*, 2019, 72, 121–130.
- Berthelon, Matias E. and Diana I. Kruger**, “Risky behavior among youth: Incapacitation effects of school on adolescent motherhood and crime in Chile,” *Journal of Public Economics*, 2011, 95 (1), 41–53.
- Bettinger, Eric P., Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu**, “The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment,” *Quarterly Journal of Economics*, 2012, 127 (3), 1205–1242.

- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes,** “Staying in the Classroom and Out of the Maternity Ward? The Effect of Compulsory Schooling Laws on Teenage Births,” *The Economic Journal*, 2008, 118 (530), 1025–1054.
- Borusyak, Kirill, Xavier Jaravel, and Jann Spiess,** “Revisiting Event Study Designs: Robust and Efficient Estimation,” 2021.
- Correia, Sergio, Paulo Guimarães, and Tom Zylkin,** “Fast Poisson estimation with high-dimensional fixed effects,” *The Stata Journal*, 2020, 20 (1), 95–115.
- Cowan, Benjamin W.,** “Forward-thinking teens: The effects of college costs on adolescent risky behavior,” *Economics of Education Review*, 2011, 30 (5), 813–825. Special Issue on Education and Health.
- Cáceres-Delpiano, Julio, Eugenio Giolito, and Sebastián Castillo,** “Early impacts of college aid,” *Economics of Education Review*, 2018, 63, 154–166.
- de Chaisemartin, Clément and Xavier D’Haultfœuille,** “Difference-in-Differences Estimators of Intertemporal Treatment Effects,” Working Paper 29873, National Bureau of Economic Research March 2022.
- and —, “Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey,” *The Econometrics Journal*, 06 2022.
- Doleac, Jennifer L. and Chloe R. Gibbs,** “A Promising Alternative: How Making College Free Affects Teens’ Risky Behaviors,” *Unpublished manuscript*, 2016.
- Duflo, Esther, Pascaline Dupas, and Michael Kremer,** “Education, HIV, and Early Fertility: Experimental Evidence from Kenya,” *American Economic Review*, 2015, 105 (9), 2757–2797.
- , —, and —, “The Impact of Free Secondary Education: Experimental Evidence from Ghana,” *NBER Working Paper*, 2021, (28937).
- Dynarski, Susan, CJ Libassi, Katherine Micheltore, and Stephanie Owen,** “Closing the Gap: The Effect of Reducing Complexity and Uncertainty in College Pricing on the Choices of Low-Income Students,” *American Economic Review*, 2021, 111 (6), 1721–1756.
- Ferreira, Maria Marta, Ciro Avitabile, Javier Botero Álvarez, Francisco Haimovich Paz, and Sergio Urzúa,** “At a Crossroads—Higher Education in Latin America and the Caribbean,” Report, Washington, DC: World Bank Group 2017.
- Flórez, Carmen Elisa and Victoria Eugenia Soto,** “Tendencias del embarazo en la adolescencia y sus determinantes próximos en Colombia.” In [Vargas Trujillo et al., eds \(2019\)](#) pp. 59–98.

- Gamboa, Luis Fernando and Paul Rodríguez-Lesmes**, “The Fertility-Inhibiting Effect of Mosquitoes: Socio-economic Differences in Response to the Zika Crisis in Colombia,” *Economics and Human Biology*, 2019, 35, 63–72.
- Gaviria, Alejandro**, “Cambio social en Colombia durante la segunda mitad del siglo XX,” Documentos CEDE 30, Universidad de los Andes, Facultad de Economía, CEDE October 2010.
- Guerra-Cújar, María Elvira, Mounu Prem, Paul Rodríguez-Lesmes, and Juan F. Vargas**, “The Peace Baby Boom: Evidence from Colombia’s Peace Agreement with the FARC,” *LACEA Working Paper Series*, October 2020, (0052).
- Heß, Simon**, “Randomization inference with Stata: A guide and software,” *Stata Journal*, 2017, 17 (3), 630–651.
- Hincapie, Diana**, “Do Longer School Days Improve Student Achievement? Evidence from Colombia,” IDB Working Paper Series IDB-WP-679, Washington, DC 2016.
- Jensen, Robert**, “Do Labor Market Opportunities Affect Young Women’s Work and Family Decisions? Experimental Evidence from India,” *Quarterly Journal of Economics*, 2012, 127 (2), 753–792.
- Kearney, Melissa S. and Phillip B. Levine**, “Why is the Teen Birth Rate in the United States So High and Why Does It Matter?,” *Journal of Economic Perspectives*, 2012, 26 (2), 141–166.
- and —, “Income Inequality and Early Nonmarital Childbearing,” *Journal of Human Resources*, 2014, 49 (1), 1–31.
- Koohi, Shiva**, “College prospects and risky behavior among Mexican immigrant youth: The effects of in-state tuition policies on schooling and childbearing,” *Economics of Education Review*, 2017, 58, 162–174.
- Laajaj, Rachid, Andrés Moya, and Fabio Sánchez**, “Equality of opportunity and human capital accumulation: Motivational effect of a nationwide scholarship in Colombia,” *Journal of Development Economics*, 2022, 154, 102754.
- Londoño-Vélez, Juliana, Catherine Rodríguez, and Fabio Sánchez**, “Upstream and Downstream Impacts of College Merit-Based Financial Aid for Low-Income Students: Ser Pilo Paga in Colombia,” *American Economic Journal: Economic Policy*, May 2020, 12 (2), 193–227.
- McCrary, Justin and Heather Royer**, “The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth,” *American Economic Review*, 2011, 101, 158–195.

- Muralidharan, Karthik and Nishith Prakash**, “Cycling to School: Increasing Secondary School Enrollment for Girls in India,” *American Economic Journal: Applied Economics*, 2017, 9 (3), 321–350.
- OECD and World Bank**, “Reviews of National Policies for Education—Tertiary Education in Colombia,” Report, Paris: OECD Publishing 2012.
- Prem, Mounu, Juan F. Vargas, and Olga Namen**, “The Human Capital Peace Dividend,” *Journal of Human Resources*, 2021, pp. 0320–10805R2.
- , **Santiago Saavedra, and Juan F. Vargas**, “End-of-conflict deforestation: Evidence from Colombia’s peace agreement,” *World Development*, 2020, 129, 104852.
- Urdinola, B. Piedad and Carlos Ospino**, “Long-term consequences of adolescent fertility: The Colombian case,” *Demographic Research*, 2015, 32 (55), 1487–1518.
- Vargas Trujillo, Elvia, Carmen Elisa Flórez, Darwin Cortés, and Marta Carolina Ibarra, eds**, *Embarazo temprano: Evidencias de la investigación en Colombia*, Universidad de los Andes, Ediciones Uniandes-Universidad del Rosario, 2019.
- Willis, Robert J.**, “A New Approach to the Economic Theory of Fertility Behavior,” *Journal of Political Economy*, 1973, 81 (2), S14–S64.

## A Appendix Tables and Figures

**Table A1.** Municipality characteristics.

	Comparison municipality mean (1)	Treatment municipality difference (2)
Ln(population) (2008)	9.379 (0.040)	0.402*** (0.064)
Rural share of population (2008)	0.617 (0.010)	-0.085*** (0.014)
Distance to department's capital (km)	87.335 (2.666)	-19.266*** (3.419)
Distance to nearest SPP eligible institution (km)	111.270 (4.707)	-25.403*** (6.018)
Poverty incidence (2005)	0.548 (0.005)	-0.066*** (0.006)
Gini coefficient (0-1) (2005)	0.462 (0.002)	-0.016*** (0.002)
Public expenditure per capita (2008)	1,170.896 (179.473)	-272.104*** (180.781)
Tax revenue per capita (2008)	94.599 (4.230)	48.854*** (8.575)
Public investment in education per capita (2008)	329.604 (159.372)	-109.323*** (195.344)
Gross enrollment rate 6th-9th grade (2011)	0.975 (0.012)	0.091*** (0.016)
Gross enrollment rate 10th-11th grade (2011)	0.646 (0.011)	0.140*** (0.015)
Dropout rate 6th-9th grade, public schools (2011)	0.050 (0.001)	0.002*** (0.002)
Dropout rate 10th-11th grade, public schools (2011)	0.040 (0.001)	-0.001*** (0.002)
Exposed to FARC (0/1) (2011-2014)	0.142 (0.015)	-0.042** (0.020)
Number of municipalities	526	541

*Notes:* This table compares pre-SPP characteristics between treatment and comparison municipalities. Columns 1 and 2 present results of a regression of a municipality characteristic on an indicator for being a treatment municipality. Column 1 shows the coefficients on the intercept term and represents the mean of comparison municipalities. Column 2 shows coefficients on the treatment indicator and represents the mean difference between treatment and comparison municipalities. Money variables are measured in 2019 thousand Colombian pesos. Robust standard errors are presented in parentheses (\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ). Significance stars are suppressed for coefficients on the intercept term.

**Table A2.** Summary difference-in-differences estimates for teen fatherhood (Poisson model).

Age group:	Births per 1,000 men			
	All teens (15-19)	15-17	18	19
	(1)	(2)	(3)	(4)
<i>SPP × Post</i>	-0.081*** (0.021)	-0.076** (0.031)	-0.112*** (0.025)	-0.066*** (0.021)
Observations	12,804	12,744	12,792	12,804
Treatment municipalities	541	541	541	541
Comparison municipalities	526	526	526	526
Pre-trends testing <i>p</i> -value	0.398	0.259	0.554	0.200
Pre-SPP share of teen fathers	100	25.7	33.6	40.7

*Notes:* This table presents summary difference-in-differences estimates following Equation 3 and a Poisson model with the number of births per 1,000 men in each age group as the outcome. For estimation, we use Correia, Guimarães and Zylkin (2020)'s Stata command. Column 1 presents the estimate for all male teens (15-19 years old). Column 2 shows the results for male teens 15-17 years of age. Columns 3 and 4 present the estimates for men 18 and 19 years of age, respectively. All estimates are weighted by the annual population of men in each municipality and age group. Standard errors are clustered at the municipality level and presented in parentheses (\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).

**Table A3.** Robustness to alternative specifications and sample of municipalities.

	Log fertility rate			IHS fertility rate	Fertility rate	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Teen</i> × <i>SPP</i> × <i>Post</i>	-0.057*** (0.011)	-0.057*** (0.014)	-0.035*** (0.008)	-0.059*** (0.011)	-0.060*** (0.011)	-3.204*** (0.701) [-4.5%]
Treatment measure	Discrete	Discrete	Continuous	Discrete	Discrete	Discrete
Exclude capital cities	No	Yes	No	No	No	No
Model	Linear	Linear	Linear	Linear	Poisson	Linear
Observations	25,608	24,840	25,608	26,520	26,514	26,520
Treatment municipalities	541	514	–	552	552	552
Comparison municipalities	526	521	–	553	553	553
Pre-trends test <i>p</i> -value	0.985	0.951	0.684	0.946	0.984	0.972

*Notes:* This table reports results from variations of our main specification and sample of municipalities. Column 1 replicates our main results. In column 2, we exclude the group of 32 municipalities corresponding to capitals of departments. In column 3, we replace  $SPP_m^*$  in Equation 4 by the standardized rate of female SABER 11 test takers in 2014 who are eligible for SPP in each municipality ( $StdSPP_m$ ). Our pre-trends test cannot be applied with the continuous version of the treatment in column 3. Following Muralidharan and Prakash (2017), we replace it by a differential linear pre-trends test using pre-SPP data. Specifically, the reported *p*-values in column 3 comes from testing the significance of  $\beta$  in the equation  $Y_{amt} = \alpha + \beta (Teen_a \times StdSPP_m \times t) + \gamma_{am} + \gamma_{mt} + \gamma_{ad(m)t} + \epsilon_{amt}$ . Column 4 uses the inverse hyperbolic sine (IHS) transformation of the birth rate as the outcome instead of the log transformation. It includes the group of municipalities with zero births in any given period for any age group (38 out of 1,105, or 3.4%). Columns 5 and 6 use the fertility rate (in levels) as the outcome variable instead of the log fertility rate and include all municipalities. Column 5 estimates the model using a Poisson pseudomaximum likelihood (PPML) regression using Correia, Guimarães and Zylkin (2020)'s Stata command. For column 6, the implied percentage change with respect to the pre-SPP mean adolescent fertility rate of treatment municipalities is shown in brackets. All estimates are weighted by the annual population of women in each municipality and age group. Standard errors are clustered at the municipality level and are reported in parentheses (\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).



**Table A4.** Robustness to alternative treatment definitions.

	Log fertility rate		IHS births
	(1)	(2)	(3)
Triple difference coefficient	-0.057*** (0.011)	-0.096*** (0.015)	-0.107*** (0.022)
Across-municipality treatment	Above median SPP eligibility rate	Below median distance SPP-eligible HEI	Above median SPP eligibility rate
Within-municipality comparison	25-29	25-29	15-19, < 8th grade
Sample (age groups)	15-19 & 25-29	15-19 & 25-29	15-19
Observations	25,608	25,608	25,608
Treatment municipalities	541	535	541
Comparison municipalities	526	532	526
Pre-trends test $p$ -value	0.985	0.023	0.369

*Notes:* This table reports results from variations of our across-municipality treatment definition and within-municipality comparison group. Column 1 replicates our main triple difference results. In column 2, we use the municipality-level distance to the nearest municipality with an SPP-eligible HEI as our across-municipality treatment definition. Treatment municipalities are those below the median distance (closer), while comparison municipalities are those above the median (farther away). In column 3, we use the across-municipality treatment definition as column 1 but replace the 25-29 years old women as the within-municipality comparison group with teenagers with a level of education less than 8th grade. The outcome in column 3 is the inverse hyperbolic sine (IHS) transformation of the number of births from teenagers in each education-municipality-year cell, because some municipalities in our main sample have zero births for certain education groups. All estimates are weighted by the annual population of women in each municipality and age group. Standard errors are clustered at the municipality level and are reported in parentheses (\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).

**Table A5.** Robustness to alternative estimators.

Estimation:	Log fertility rate		
	OLS	Imputation Estimator	DID
	(1)	(2)	(3)
$Teen \times SPP \times Post$	-0.057*** (0.011)	-0.083*** (0.015)	-0.074*** (0.024)
Observations	25,608	25,608	5,335
Treatment municipalities	541	541	541
Comparison municipalities	526	526	526
Pre-trends test $p$ -value	0.985	0.985	0.875

Notes: This table reports results using [Borusyak et al. \(2021\)](#)'s imputation estimator and [de Chaisemartin and D'Haultfoeuille \(2022a\)](#)'s DID estimator. Column 1 replicates our main triple difference result using OLS to estimate [Equation 4](#). Column 2 reports the treatment effect using the imputation estimator including the same set of fixed effects from [Equation 4](#). See [Borusyak et al. \(2021\)](#) for details. Column 3 reports the average effect from the DID estimator using the same specifications from [Equation 4](#), but we recast the triple difference as a simple difference-in-differences by subtracting the within-municipality log fertility rate of older women from the the log fertility rate of teenagers. The department-specific year trends in column 3 are handled non-parametrically. The pre-trends test in column 3 corresponds to the  $p$ -value of a joint test of significance for four placebos before the policy introduction. See [de Chaisemartin and D'Haultfoeuille \(2022a\)](#) for details. All estimates are weighted by the annual population of women in each municipality and age group. Standard errors clustered at the municipality level are reported in parentheses and for the DID estimator they are based on 5,000 bootstrap replications (\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).

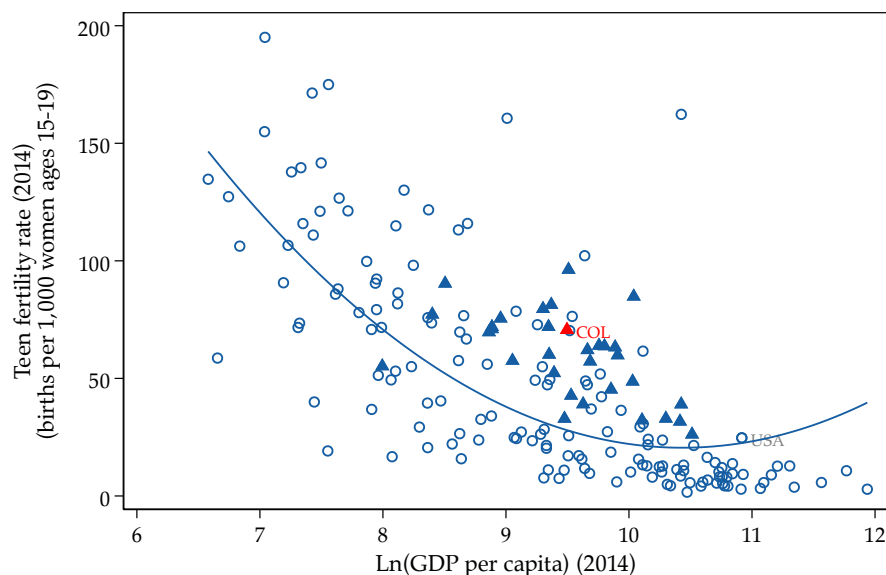
**Table A6.** Robustness to placebo treatment years.

Placebo year ( $h$ ):	Log fertility rate					
	2008	2009	2010	2011	2012	2013
	(1)	(2)	(3)	(4)	(5)	(6)
$Teen \times SPP \times Post$	-0.007 (0.012)	-0.005 (0.011)	-0.002 (0.010)	-0.001 (0.010)	0.004 (0.011)	0.003 (0.013)
Observations	14,938	14,938	14,938	14,938	14,938	14,938
Treatment municipalities	541	541	541	541	541	541
Comparison municipalities	526	526	526	526	526	526

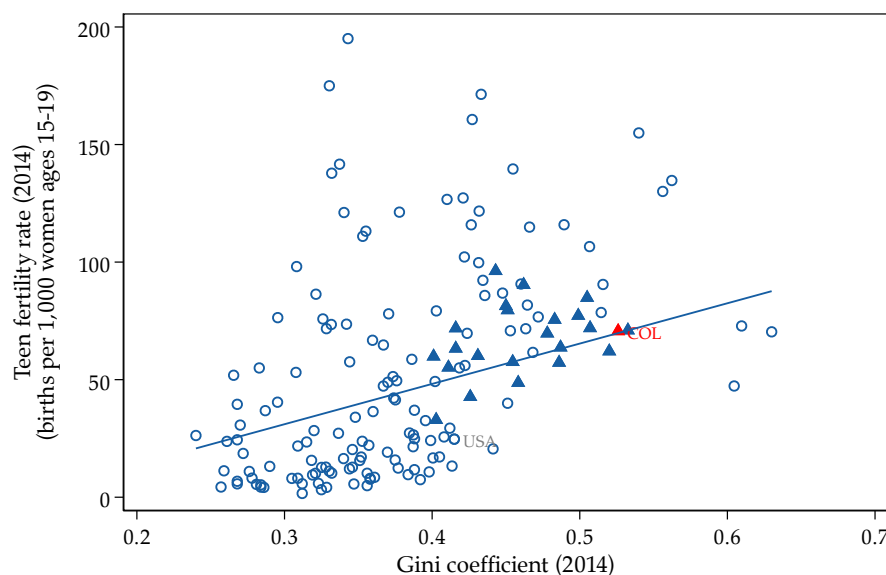
Notes: Each column in this table assumes that SPP was introduced in year  $h$  instead of 2014 and estimates [Equation 4](#) with  $Post_t = \mathbb{1}[t > h]$  and  $t \in [2008, 2014]$ . All estimates are weighted by the annual population of women in each municipality and age group. Standard errors are clustered at the municipality level and are reported in parentheses (\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).

**Figure A1.** Correlates of teen fertility rates across countries.

**(a)** Teen fertility rate and income per capita.



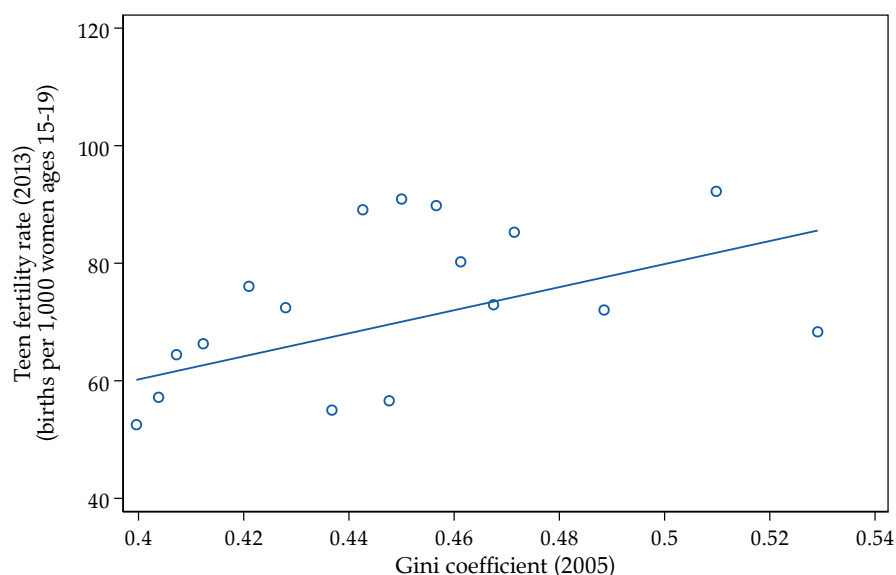
**(b)** Teen fertility rate and income inequality.



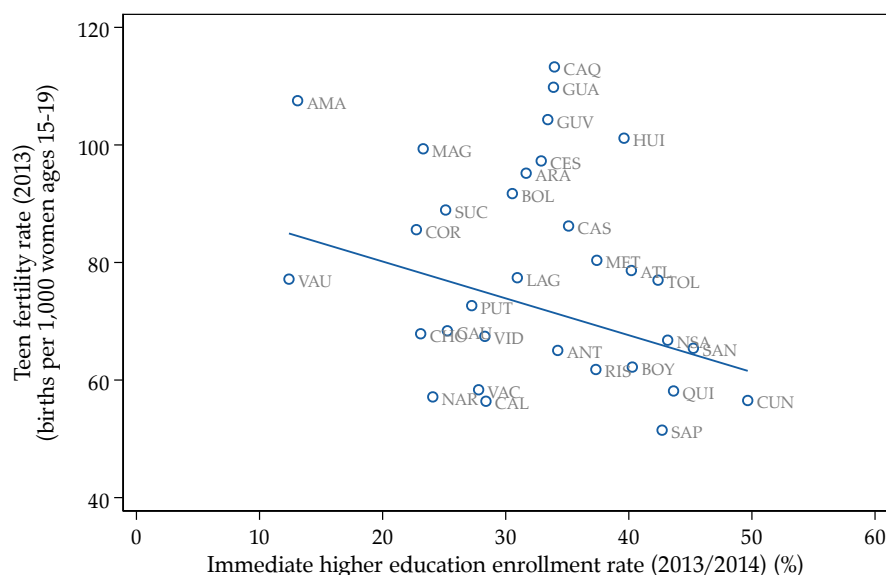
*Notes:* This figure shows some correlates of teen fertility across countries. Panel (a) shows the relationship between teen fertility rates and GDP per capita (PPP) in 2014 for a sample of 183 countries. Panel (b) shows the relationship between teen fertility rates and the Gini coefficient in 2014 for a sample of 161 countries. In Panel (b), for some countries, the Gini corresponds to the closest year before 2014 in case 2014 was unavailable. The data come from the World Bank's World Development Indicators. Countries in Latin America and the Caribbean are shown in solid triangles, while all other countries are shown in hollow circles. Colombia (COL) is highlighted in red, and the US (USA) position is included for reference. In Panel (a), the solid line corresponds to a quadratic fit weighted by the population of women ages 15-19 in each country. The coefficients are -177.032 (s.e. = 63.185) and 8.484 (s.e. = 3.549). In Panel (b), the solid line corresponds to a linear fit weighted by the population of women ages 15-19 in each country. The slope is 171.453 (s.e. = 57.503).

**Figure A2.** Correlates of teen fertility rates in Colombia.

**(a)** Teen fertility rate and income inequality across municipalities.

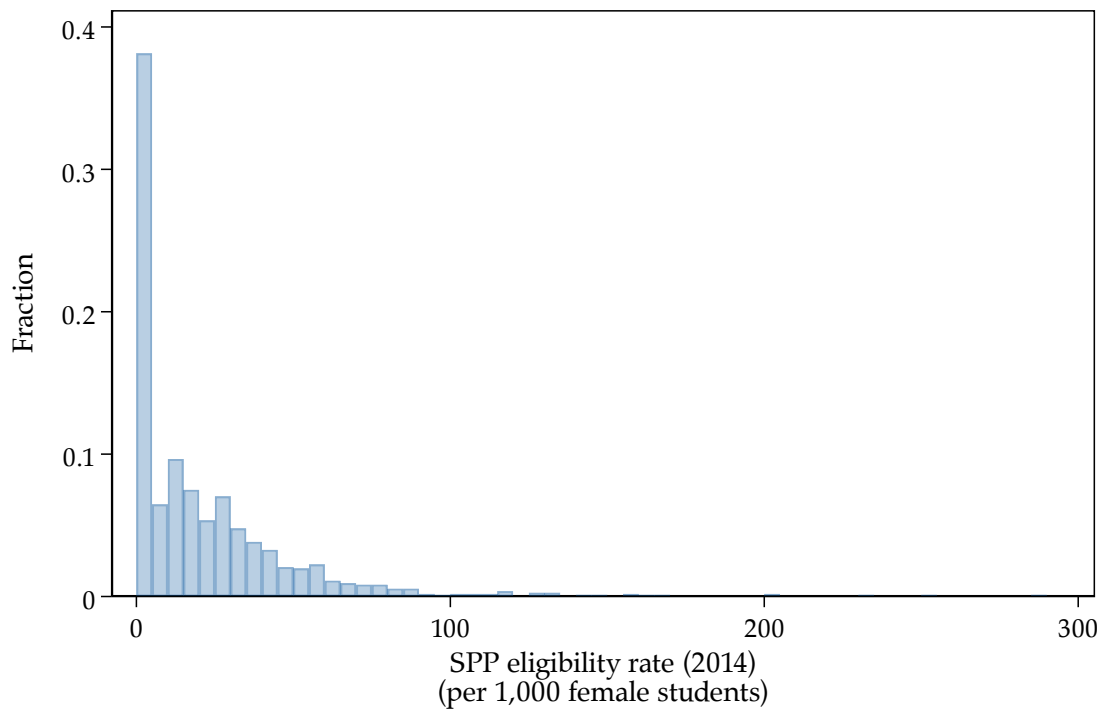


**(b)** Teen fertility rate and college enrollment across departments.



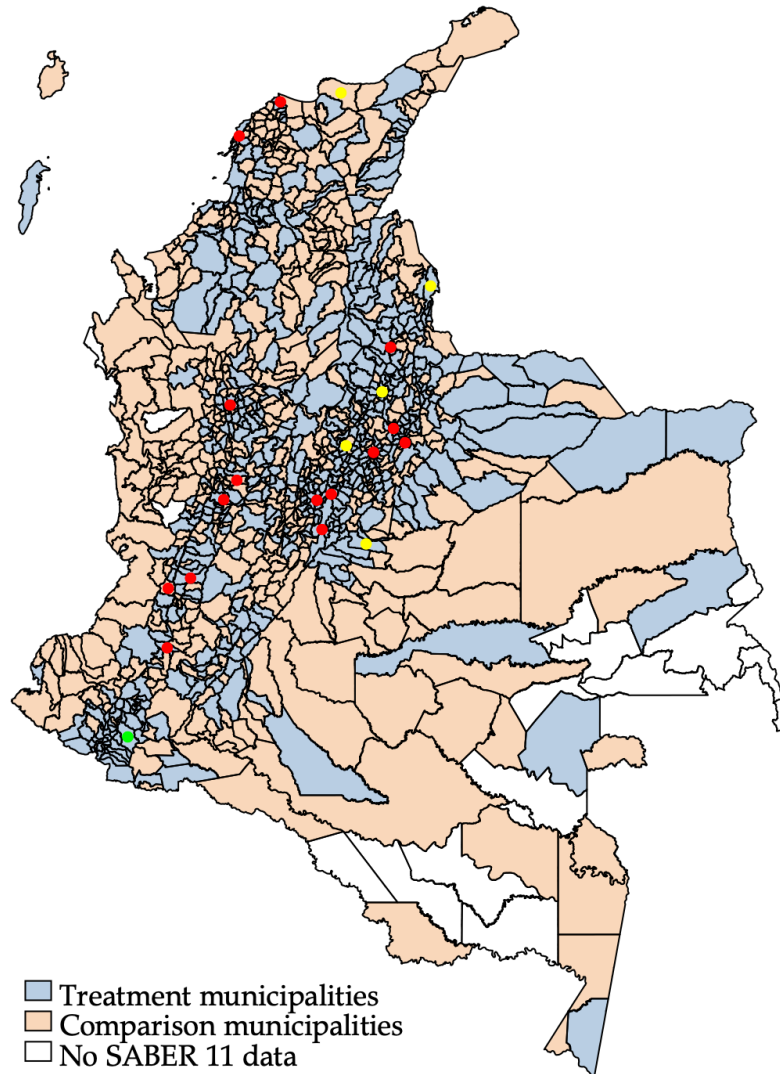
*Notes:* This figure shows some correlates of teen fertility across municipalities and departments in Colombia. Departments in Colombia are similar to states in the United States. A group of municipalities forms each department. Panel (a) shows a binned scatterplot of the relationship between teen fertility rates in 2013 and the Gini coefficient in 2005 for a sample of 1,040 municipalities. The solid line corresponds to a linear fit weighted by the population of women ages 15-19 in each municipality. The slope is 196.207 (s.e. = 44.515). Panel (b) shows the relationship between teen fertility rates in 2013 and the immediate higher education enrollment rate for the 2013 cohort of high school seniors for the 32 departments in Colombia. This rate is measured as the percentage of students in 11th grade in 2013 that enrolled in higher education in 2014. The solid line corresponds to a linear fit weighted by the population of women ages 15-19 in each department. The slope is -0.627 (s.e. = 0.304). Both panels adjust birth rates for the lag between conception and birth. The data come from the official birth records and the Ministry of Education.

**Figure A3.** Distribution of SPP eligibility rates.



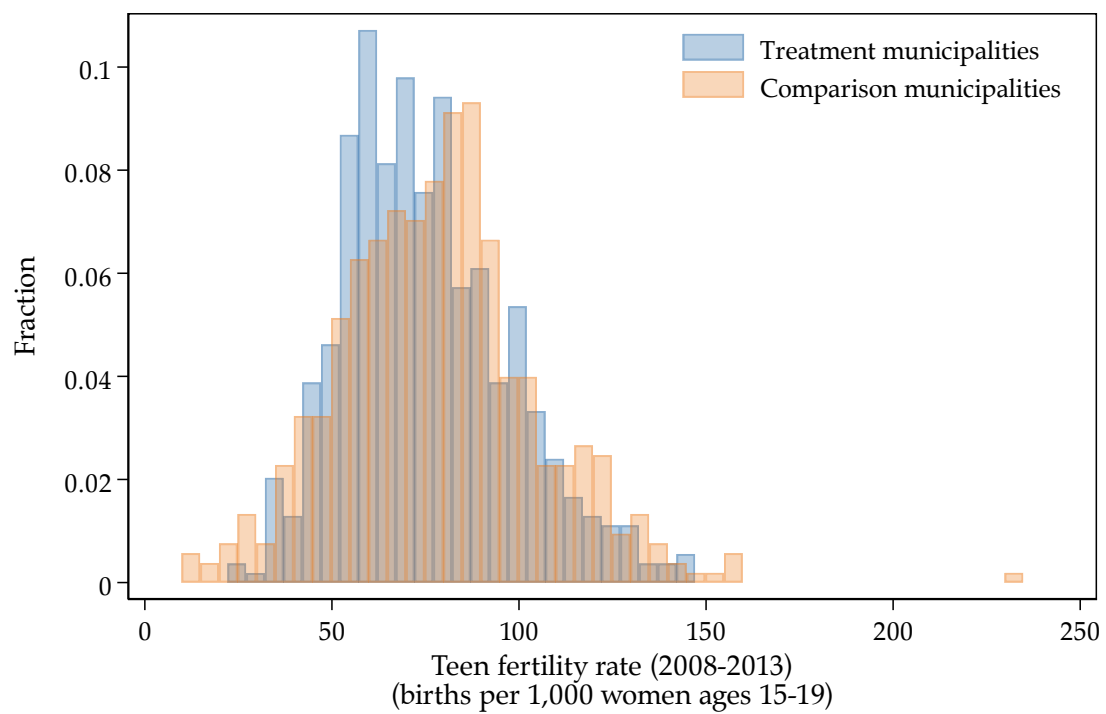
*Notes:* This histogram shows the distribution of SPP eligibility rates in 2014 for the 1,067 municipalities in our main sample. For the overall sample of municipalities, the minimum value is 0 and the maximum value is 285.7 (eligible female students per 1,000 female students). The median eligibility rate is 12.7. The 25th and 75th percentiles are 0 and 30.8, respectively.

**Figure A4.** Municipality-level variation in SPP eligibility.



*Notes:* This map displays the geographic distribution of treatment and comparison municipalities. Treatment municipalities are above the median in female eligibility rates for SPP in 2014, while comparison municipalities are below the median. Dots indicate municipalities with at least one SPP-eligible HEI (red: from SPP 1; yellow: added in SPP 3; green: added in SPP 4).

**Figure A5.** Distribution of teen fertility rates.

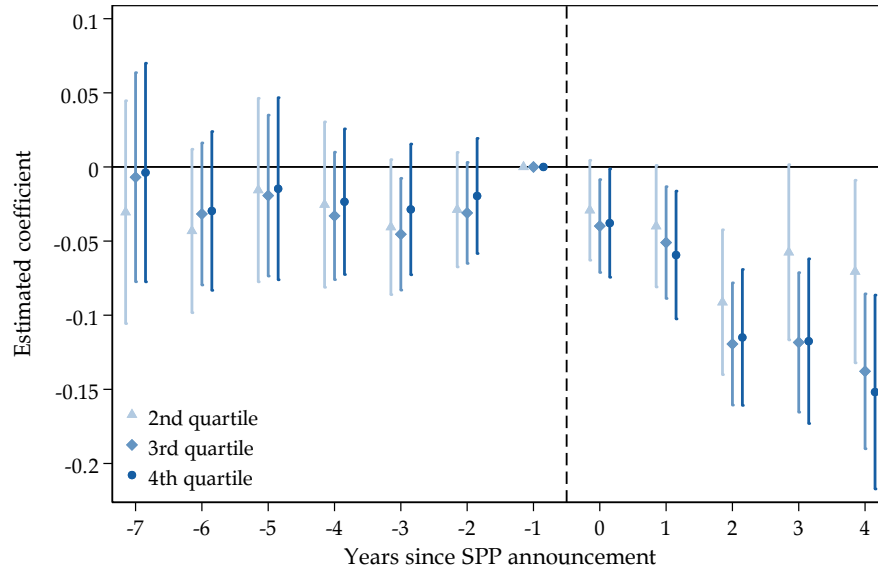


*Notes:* This histogram shows the distribution of teen eligibility rates in 2008-2013 for the 1,067 municipalities in our main sample separately for treatment and comparison municipalities. Treatment municipalities are above the median in female eligibility rates for SPP in 2014, while comparison municipalities are below the median.

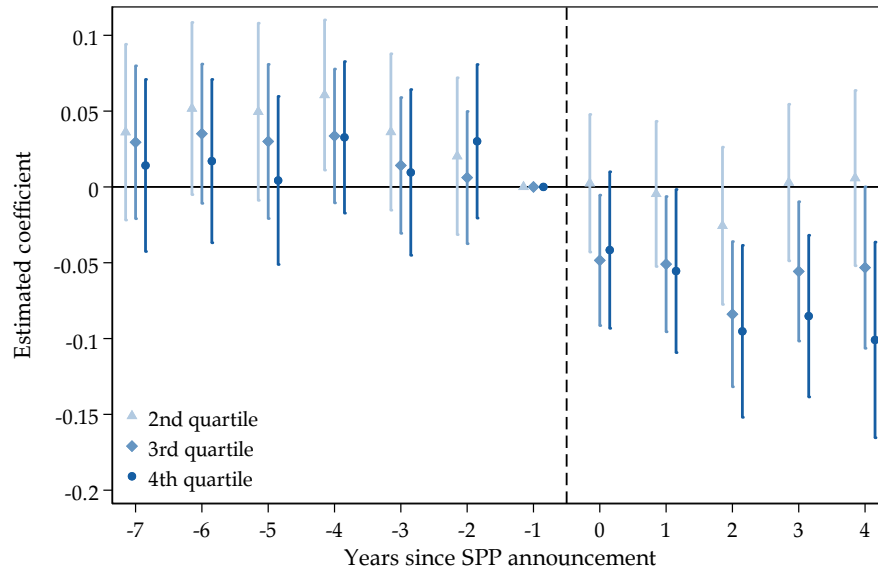


**Figure A6.** Event study estimates by quartile of initial SPP eligibility.

**(a) Difference-in-differences results**

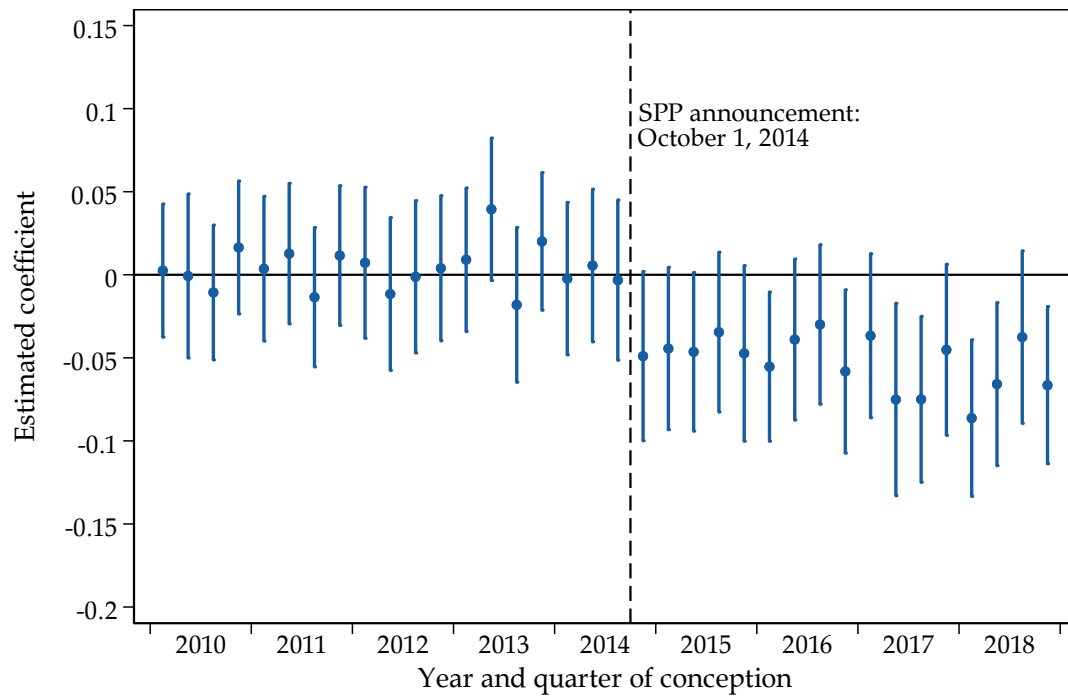


**(b) Triple difference results**



*Notes:* This figure plots the triple difference event study estimates of  $\beta_\tau$  from Equation 2 using indicators for the quartile of SPP eligibility rates instead of an indicator for being above the median in SPP eligibility. All estimates are weighted by the population of women from each age group in each municipality. Standard errors are clustered at the municipality level.

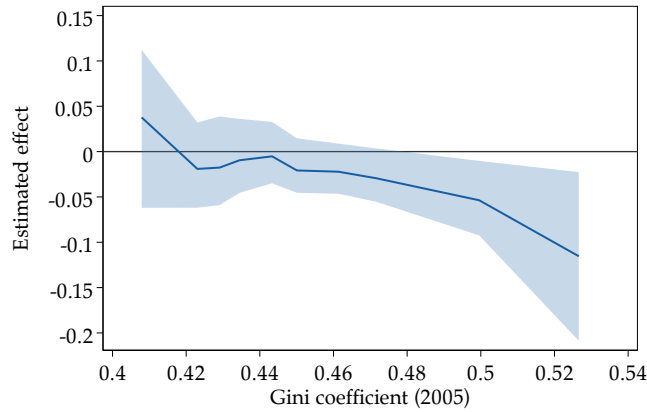
**Figure A7.** Triple difference event study estimates using quarterly data (Poisson model).



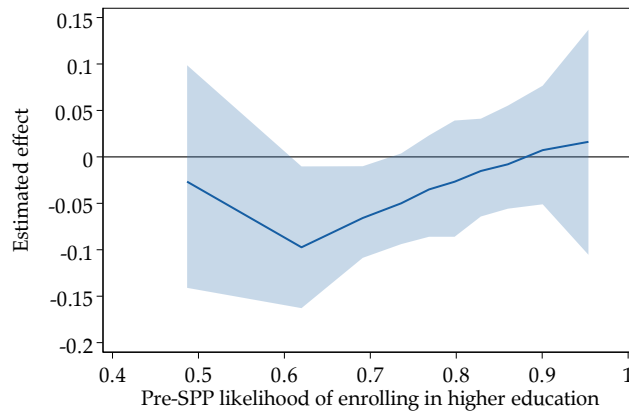
*Notes:* This figure plots the triple difference event study estimates of  $\beta_\tau$  from Equation 2 using quarterly data instead of annual data. We use the birth rate as the outcome and a Poisson model instead of the log transformation with the linear model, because at the quarterly level, some municipalities have zero births in some of the cells. For estimation, we use Correia, Guimarães and Zylkin (2020)'s Stata command. The quarters in 2008 are the reference period. Only estimates for four years around 2014 are plotted. The dots represent the estimated coefficients and the vertical lines represent 95 percent confidence intervals. All estimates are weighted by the annual population of women in each municipality and age group. Standard errors are clustered at the municipality level.

**Figure A8.** Triple difference estimates by municipality characteristics (non-parametric analysis).

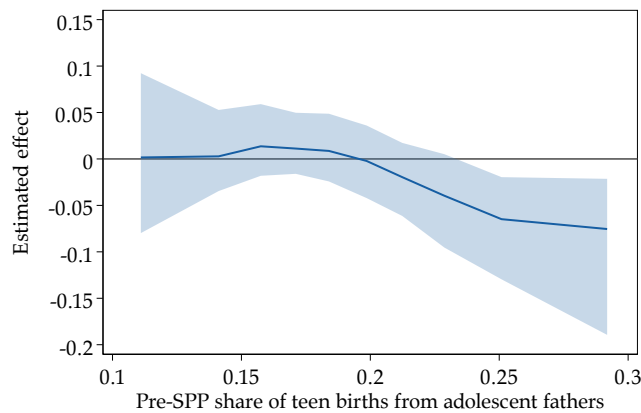
**(a)** Baseline income inequality.



**(b)** Baseline college-going expectations.

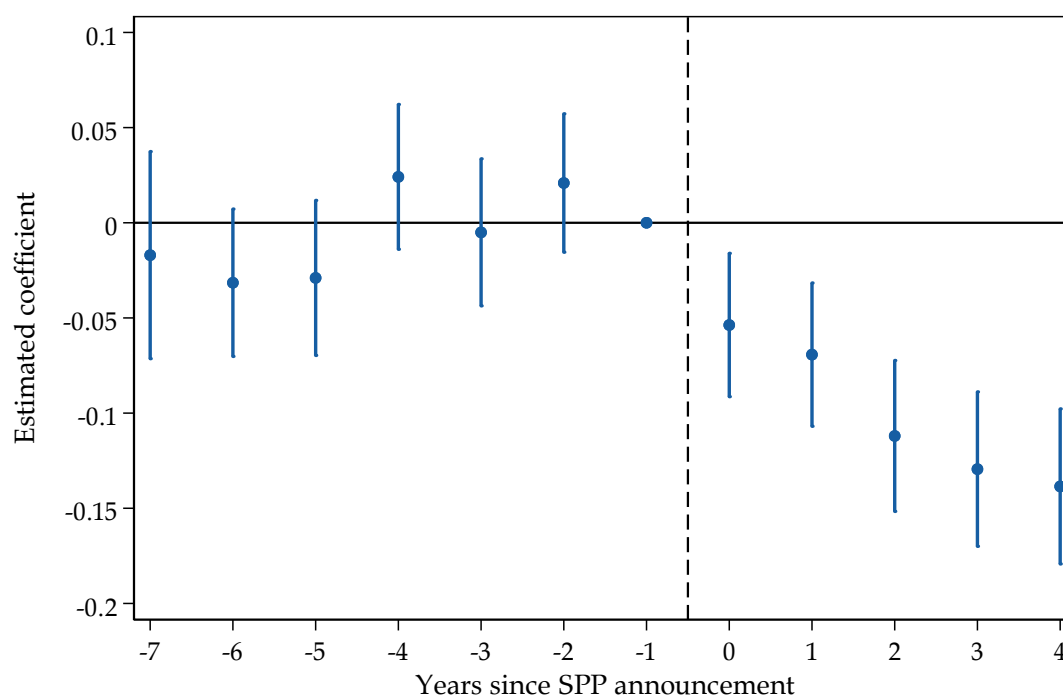


**(c)** Baseline share of teenbirths from adolescent fathers.



*Notes:* This figure presents triple difference estimates by municipality characteristics indicated in each panel. To do this, we first divide municipalities in deciles according to the level of each characteristic. After recasting the triple difference as a simple difference-in-differences by calculating the within-municipality difference between the (log) fertility rate of teens and non-teens, we use Equation 3 to estimate the triple difference reduction in fertility for each decile. We then plot these estimates using a lowess regression smoothing of the triple difference coefficients on the average level of the characteristic for the deciles. The shaded area represents 95% confidence interval based on 10,000 bootstrap replications of this procedure.

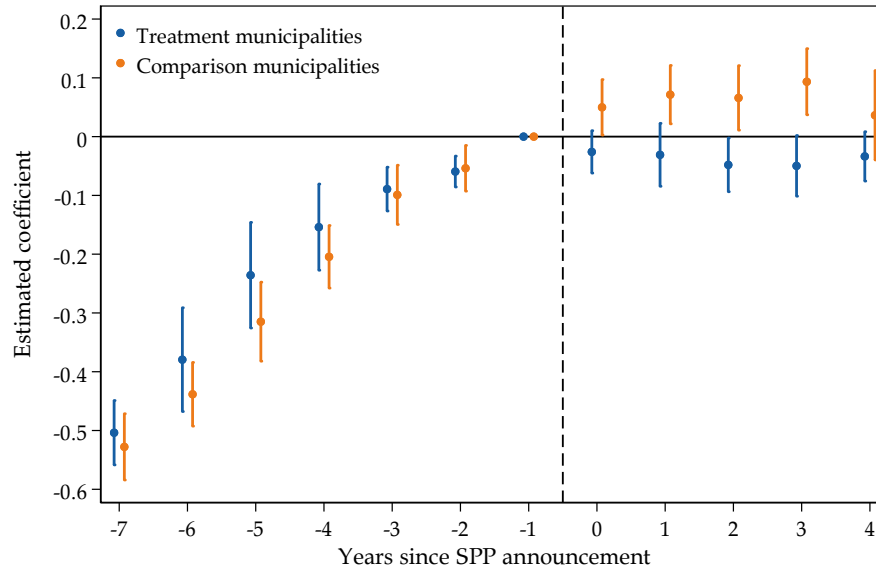
**Figure A9.** Event study estimates exploiting distance to nearest SPP-eligible institution.



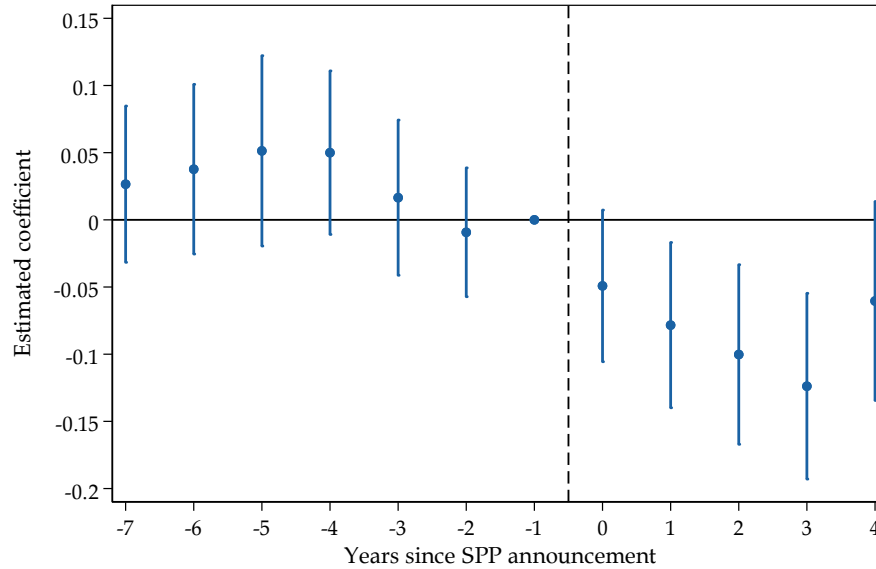
*Notes:* This figure plots the triple difference event study estimates from [Equation 2](#), where  $SPP_m^*$  is replaced with an indicator for whether a municipality is below the median distance to the nearest SPP-eligible institution. The dots represent the estimated coefficients and the vertical lines represent 95 percent confidence intervals. All estimates are weighted by the number of women for each age group in each municipality. Standard errors are clustered at the municipality level.

**Figure A10.** Event study estimates exploiting mothers' grade level.

**(a)** Trends in IHS birth differentials between grade levels.



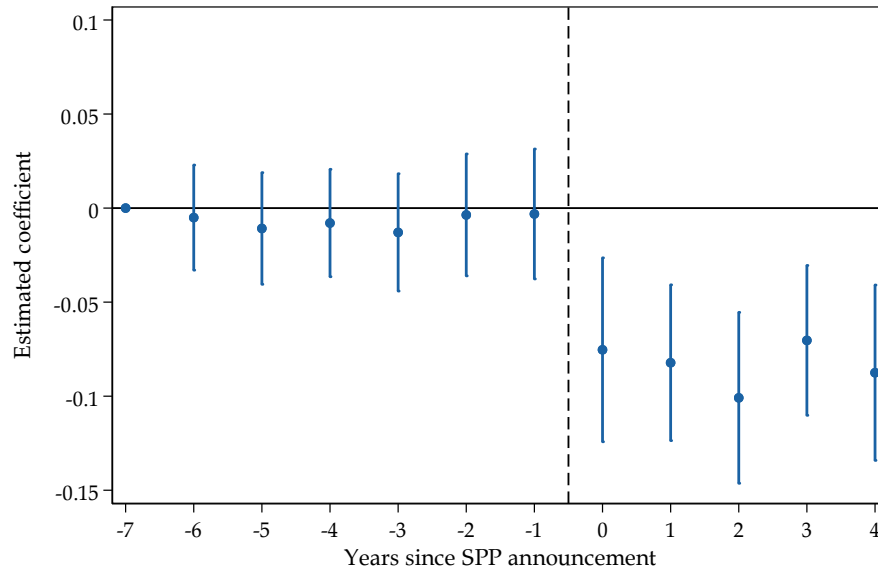
**(b)** Triple difference estimates by grade level and SPP eligibility.



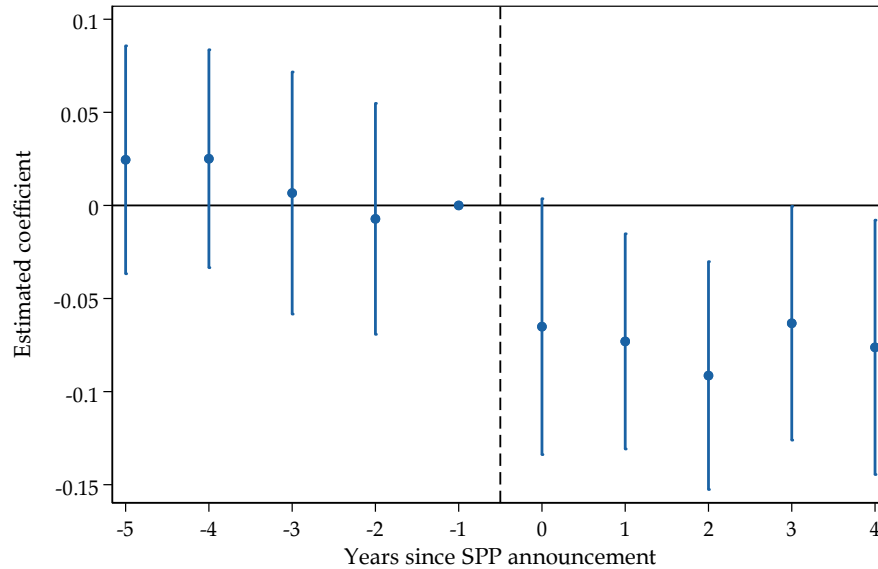
*Notes:* This figure plots event study estimates from [Equation 2](#), where  $Teen_a$  is replaced with an indicator for mothers who have completed eighth grade or higher. Mothers aged 15-19 with less than an eighth grade education is used as the comparison group instead of women aged 25-29. We use the number of births instead of a birth rate because we cannot reliably calculate the number of young women in each municipality with above or below an eighth grade level of education. Also, we use an inverse hyperbolic sine (IHS) transformation instead of a log transformation since some municipalities have zero recorded births in these year by grade-level cells. Panel (a) plots trends in the differentials between grade levels, separately for treatment and comparison municipalities. Panel (b) plots the triple difference coefficients. The dots represent the estimated coefficients and the vertical lines represent 95 percent confidence intervals. All estimates are weighted by the number of women for each age group in each municipality. Standard errors are clustered at the municipality level.

**Figure A11.** Tripple difference event study estimates using alternative estimators.

**(a)** [Borusyak et al. \(2021\)](#)'s imputation estimator.

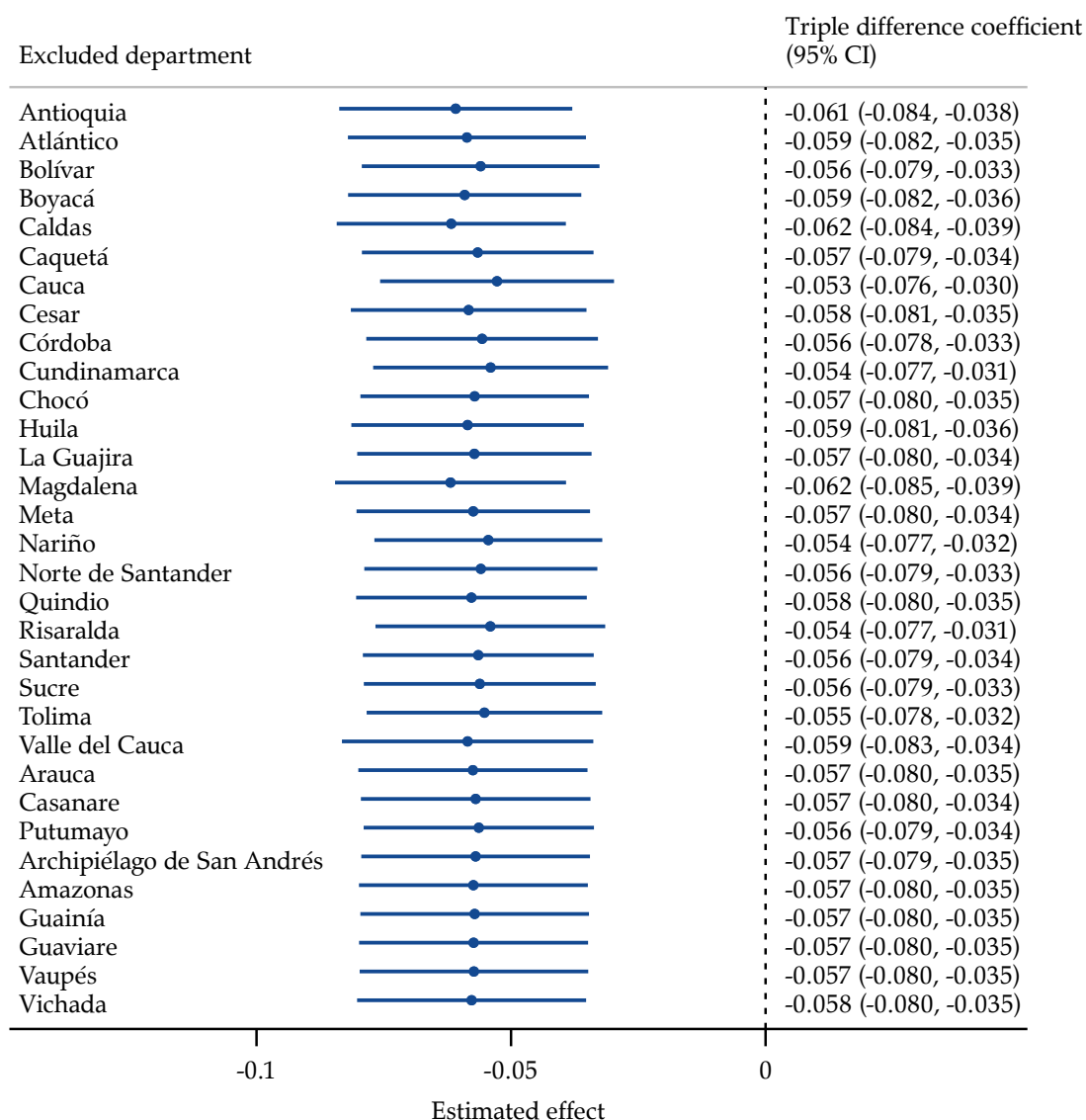


**(b)** [de Chaisemartin and D'Haultfœuille \(2022a\)](#)'s DID estimator.



*Notes:* This figure plots the triple difference event study estimates using [Borusyak et al. \(2021\)](#)'s imputation estimator and [de Chaisemartin and D'Haultfœuille \(2022a\)](#)'s DID estimator. In panel (a), the estimates to the left of the dotted line correspond to OLS estimates for pre-trends and the ones to the right represent treatment effects. Vertical lines represent 95 percent confidence intervals. See [Borusyak et al. \(2021\)](#) for more details. In panel (b), the estimates to the left of the dotted line correspond to placebo estimates and the ones to the right represent DID treatment effect estimates. Vertical lines represent 95 percent confidence intervals. See [de Chaisemartin and D'Haultfœuille \(2022a\)](#) for details. All estimates are weighted by the annual population of women in each municipality and age group. Standard errors are clustered at the municipality level and are based on 5,000 bootstrap replications for the DID estimator.

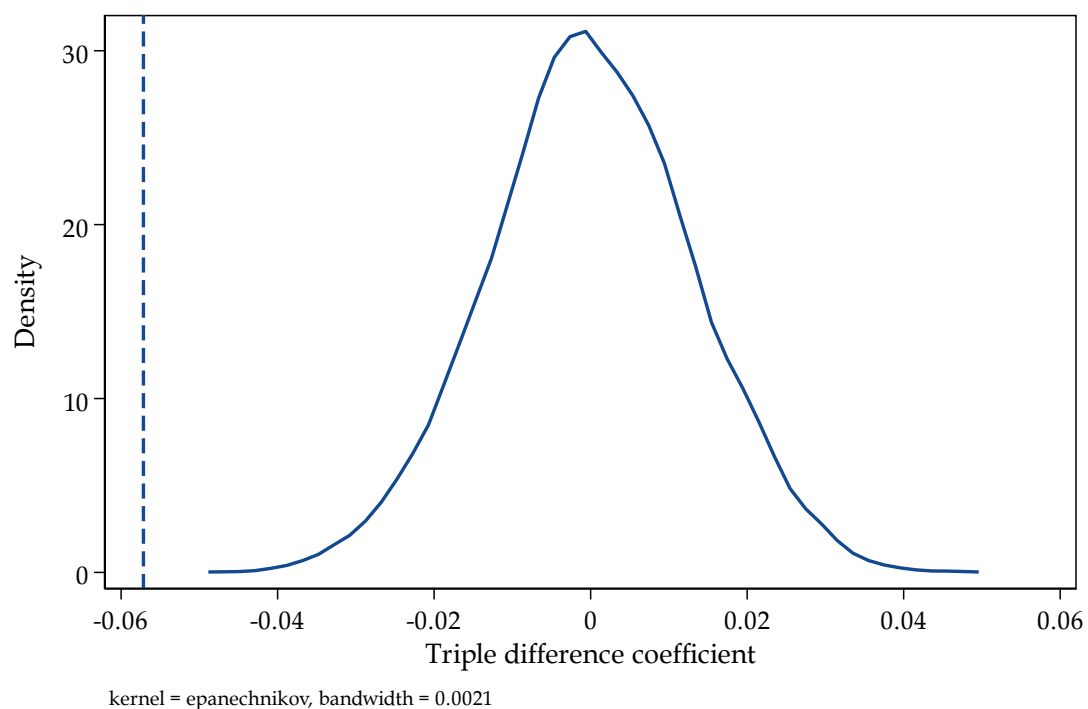
**Figure A12.** Robustness to excluding all municipalities in a given department.



*Notes:* This figure reports triple difference estimates of  $\beta$  from Equation 4 where a single department is excluded in each regression. Dots represent estimated coefficients and horizontal lines represent 95 percent confidence intervals. All estimates are weighted by the annual population of women in each municipality and age group. Standard errors are clustered at the municipality level.



**Figure A13.** Randomization Inference (RI): Distribution of placebo treatments.



*Notes:* This figure presents the distribution of placebo treatments after 5,000 random permutations of the treatment assignment (i.e., we randomize municipalities to be treatment or comparison municipalities). We run the regressions using our summary specification in [Equation 4](#). The vertical dashed line represents the original estimated coefficient. RI-based  $p$ -value = 0.000. The procedure was implemented using the routine by [Heß \(2017\)](#).

## B Validation of Treatment Intensity Measure

We attempt to probe the validity of our treatment intensity measure in the context of the previous literature by estimating whether it is associated with an increase in SABER 11 test scores after SPP is introduced. This is essentially testing whether we can replicate the results from [Bernal and Penney \(2019\)](#) and [Laajaj et al. \(2022\)](#) using our treatment measure. We use individual-level test scores on the SABER 11 exam for 15-19 years old female students between 2010 and 2016.<sup>21</sup> We use a triple difference empirical approach that leverages the same municipality-level variation in SPP eligibility as in our fertility analysis (see [Equation 4](#)) and variation between students who are eligible for SPP on the SISBEN margin and those who are not.

Specifically, we estimate the following equation by OLS:

$$\begin{aligned} StdTestScore_{it} = & \phi \left( SISBEN_i^{1-2} \times SPP_{m(i)}^* \times Post_t \right) + X_{it}\Gamma_t \\ & + \psi_{s(i)m(i)} + \psi_{m(i)t} + \psi_{s(i)d(i)t} + v_{it}, \end{aligned} \quad (B1)$$

where  $i$  denotes student,  $m$  denotes municipality,  $d$  denotes department,  $s$  denotes SISBEN level, and  $t$  denotes year. The  $StdTestScore_{it}$  variable is students' SABER 11 test score standardized by test year.<sup>22</sup>  $SISBEN_i^{1-2}$  indicates whether the student is categorized as SISBEN level 1 or 2. A SISBEN level of 1 or 2 is roughly equivalent to being eligible for SPP on the SISBEN margin, whereas students with higher SISBEN levels or not categorized are ineligible.  $SPP_m^*$  denotes treatment and comparison municipalities and is defined as  $SPP_m^* = 1 [SPP_m > \text{median}(SPP_m)]$  with  $SPP_m$  being the rate of female students eligible for the program in a given municipality. [Equation B1](#) contains a set of controls ( $X_{it}$ ), including the average ranking for the school the student attends (as a proxy for school quality), indicators for the student's

<sup>21</sup>A consistent SISBEN level variable is only available in the SABER 11 data for these years. The SISBEN level is self-reported by the student. We use data from fall semesters only.

<sup>22</sup>The overall individual SABER 11 score is a linear combination of scores in different subjects. We follow ICFES and [Londoño-Vélez et al. \(2020\)](#) (see their Online Appendix) and calculate this individual score as follows:

$$\begin{aligned} TestScore_i = & \frac{Chem_i + Bio_i + Phys_i + 2SocSci_i + Philo_i + 3Lang_i + 3Math_i + Eng_i}{13} \text{ for 2010-2013, and} \\ & TestScore_i = 5 \times \left( \frac{3Math_i + 3Reading_i + 3NatSci_i + 3SocSci_i + Eng_i}{13} \right) \text{ for 2014-2016.} \end{aligned}$$

We then normalize these scores by year using the mean and standard deviation from the whole sample of students (males and females) in each year. We only use data from the fall semester each year.

age, whether the student is enrolled in a public school, whether the student is enrolled in a rural school, the student's school schedule, and the parents' education levels. We interact all these indicators with year dummies. In the standard way for triple difference specifications, we include the three two-way interactions, denoted by  $\psi$ , between fixed effects for SISBEN levels ( $s \in \{1-2, \text{Other}\}$ ), municipalities, and years. Similar to our main fertility specification (see Equation 4), we allow the SISBEN-specific year effects to vary by region (i.e., department). Finally,  $v_{it}$  is an error term. In Equation B1,  $\phi$  is our parameter of interest, measuring the effect of SPP on test scores.

**Table B1.** Triple difference estimates on SABER 11 test scores.

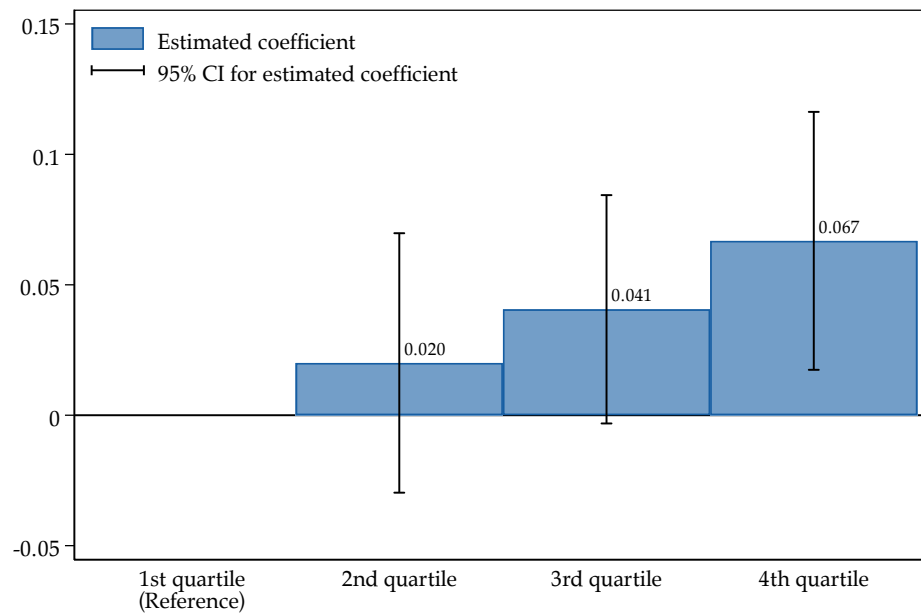
	Standardized SABER 11 test score	
	(1)	(2)
$SISBEN^{1-2} \times SPP \times Post$	0.029** (0.015)	
$SISBEN^{1-2} \times SPP_{Quartile} \times Post$		
1st quartile		[Reference]
2nd quartile		0.020 (0.025)
3rd quartile		0.041* (0.022)
4th quartile		0.067*** (0.025)
Observations	1,863,745	1,863,745
Treatment municipalities	541	–
Comparison municipalities	526	–
Pre-trends test $p$ -value	0.055	–
Pre-SPP socioeconomic achievement gap	0.701	–

*Notes:* The table above reports triple difference estimates of  $\phi$  from Equation B1. Column 2 uses indicators for the quartile of SPP eligibility rates instead of an indicator for being above the median in SPP eligibility. The reference group here are municipalities in the first (lowest) quartile of SPP eligibility. Standard errors are clustered at the municipality level and are reported in parentheses (\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).

Table B1 presents the results from this regression. We find that, after the introduction of SPP, test scores increased for SISBEN-eligible female students in treatment municipalities by 0.029 standard deviations relative to comparison municipalities ( $p$ -value = 0.049). This increase represents about 4.1 percent of the raw pre-SPP test score gap between SISBEN levels 1-2 and higher SISBEN levels. This

estimate is qualitatively similar to the estimates in [Bernal and Penney \(2019\)](#) and [Laajaj et al. \(2022\)](#). Both [Bernal and Penney \(2019\)](#) and [Laajaj et al. \(2022\)](#) use variations of regression discontinuity designs as their main strategies, and therefore their estimates reflect local average treatment effects. Since our triple difference estimates represent average treatment effects, it is reasonable to expect somewhat different results. However, the relative reduction in the socioeconomic gap here is remarkably similar to [Laajaj et al.](#)'s results.

**Figure B1.** Triple difference estimates on SABER 11 test scores by quartile of SPP eligibility.



*Notes:* This figure plots summary triple difference estimates of  $\phi$  from [Equation B1](#) using indicators for the quartile of SPP eligibility rates instead of an indicator for being above the median in SPP eligibility. Standard errors are clustered at the municipality level.

Like for our main fertility results (see [section 5](#)), in [Table B1](#) and [Figure B1](#), we also report estimates where we use indicators for the quartile of SPP eligibility rates instead of being above the median SPP eligibility. A similar pattern of results emerges here. We observe bigger increases in test scores moving from the second (0.020 SD or 2.9 percent, not statistically significant) to the fourth quartile (0.067 SD or 9.6 percent,  $p$ -value = 0.008).

## C Robustness to Possibly Confounding Events

This appendix provides a full description of the analyses we conduct to assess whether events and policies that occurred around the time SPP was introduced are driving our results. We consider three events: 1) the unilateral permanent ceasefire by the Revolutionary Armed Forces of Colombia (FARC, from the Spanish acronym) in December 2014 as part of the by then ongoing peace process between the guerrilla group and the Colombian government, 2) the Zika virus epidemic, which occurred from October 2015 to July 2016, and 3) the *Jornada Única* initiative, which gradually transitioned some public secondary schools that were operating half-day shifts into full school days beginning in 2015.

Guerra-Cújar, Prem, Rodríguez-Lesmes and Vargas (2020) finds evidence that the peace agreement with FARC led to a “baby boom” in municipalities that experienced more FARC conflict before the peace agreement, and other studies find effects of the peace agreement on educational outcomes and deforestation (Prem, Vargas and Namen, 2021; Prem, Saavedra and Vargas, 2020). While Guerra-Cújar et al. (2020) find that the relative increase in fertility rates does not seem to be driven by any particular age group, we assess whether the effects of the FARC peace agreement are driving our results. We use data from Prem et al. (2020) on the locations of FARC presence in the years before the ceasefire and estimate our main specification with the subset of municipalities that did not experience any FARC-related violence in the period 2011–2014. We report these results in column (2) of Table C1. The triple difference estimate for this subset of municipalities is -0.053, nearly identical to the estimate with all municipalities.

The Zika virus can be spread from a pregnant woman to her baby, which can result in birth defects. Gamboa and Rodríguez-Lesmes (2019) studies the effect of the Zika virus epidemic in Colombia on birth rates, finding a 10 percent decline. Using municipality-level data from the Colombian National Institute of Health, we assess whether the Zika virus could be driving our results by estimating our main specification on the subset of municipalities that experienced a low incidence of Zika during 2016, the peak year of the epidemic. These results are reported in column (3) in Table C1. Our estimate for this subset of municipalities is -0.067, even larger than our main estimate. Together, these results indicate that our estimated teen fertility impacts of SPP are not driven by the FARC ceasefire or the Zika virus epidemic.

**Table C1.** Robustness to possible confounding events.

Municipalities:	Log fertility rate			
	All	No FARC	Low Zika incidence	No <i>Jornada Única</i>
	(1)	(2)	(3)	(4)
<i>Teen × SPP × Post</i>	-0.057*** (0.011)	-0.053*** (0.013)	-0.067*** (0.024)	-0.063*** (0.017)
Observations	25,608	22,392	12,504	15,888
Treatment municipalities	541	485	258	304
Comparison municipalities	526	448	263	358
Pre-trends testing <i>p</i> -value	0.985	0.875	0.918	0.736

Notes: Column 1 reproduces the preferred triple difference results. Column 2 reports results from our summary specification in Equation 4 for municipalities that did not experience any violent events by FARC from 2011 to 2014 using data from Prem et al. (2020). Column 3 reports results from our summary specification in Equation 4 for municipalities below the median incidence of Zika in 2016. The mean incidence of Zika in these municipalities was 6.5 cases (including both confirmed and probable cases) per 100,000 inhabitants, versus 313.0 in the top half of municipalities. Column 4 excludes municipalities in which female students were exposed to *Jornada Única* (“Full School Day”) at any point between 2015-2018. All estimates are weighted by the annual population of women in each municipality and age group. Standard errors are clustered at the municipality level and are reported in parentheses (\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).

Finally, since 2015 the Colombian Ministry of Education has been gradually implementing an initiative to transition public schools from half-day shifts (morning and afternoon) to full school days to extend the duration and the quality of instruction. This policy is called *Jornada Única* or “Full School Day.”<sup>23</sup> There is evidence from other contexts that lengthening the school day can reduce adolescent pregnancies via an incapacitation effect (Berthelon and Kruger, 2011). Accordingly, we also test that our results are robust to the expansion of *Jornada Única*. Given the high costs associated with this strategy, its expansion has been very gradual over time and was not adopted in all municipalities during our period of interest. In 2015, less than 0.04 percent of the female students taking the SABER 11 test attended a school with *Jornada Única*. This share increased to 0.46 percent in 2016, 6 percent in 2017, and 8 percent in 2018. We, therefore, do not expect this policy

<sup>23</sup>See Hincapié (2016) for a review of the length of the school day in Colombia around the time of the implementation of *Jornada Única*. But, shortly, in many public schools, two separate groups of students attend the same institution (i.e., use the same physical infrastructure), one in the morning and one in the afternoon. So, there are two “shifts,” particularly in schools serving basic secondary (grades 6 to 9) and mid secondary (grades 10 and 11) students.

to explain the sharp decline in teen fertility observed right after the introduction of SPP in 2014. Column 4 of [Table C1](#) corroborates this. It presents our summary triple difference estimate excluding the municipalities in which female students were exposed to *Jornada Única* at any point between 2015-2018. We still find a big, negative and significant impact of SPP on the sample of municipalities not exposed to full-day shifts due to the *Jornada Única* initiative (-6.3 percent).