Journal of Development Economics 171 (2024) 103321

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular article

College opportunity and teen fertility: Evidence from *Ser Pilo Paga* in Colombia☆

Michael D. Bloem[a], Jesús Villero[b,*]

[a] The College Board, United States of America
[b] University of Pennsylvania, United States of America

ARTICLE INFO

JEL classification:
H52
I22
I23
I24
I25
J13
O15

Keywords:
College opportunity
Teen fertility
Inequality
Financial aid

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.

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 et al., 2012; Aizer et al., 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 Fig. A.1, 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 (2020b) 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)

☆ 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, the Carolina Region Empirical Economics Day, and the California Policy Lab for helpful comments.

* Corresponding author.
E-mail addresses: mbloem@collegeboard.org (M.D. Bloem), jvillero@wharton.upenn.edu (J. Villero).

https://doi.org/10.1016/j.jdeveco.2024.103321
Received 13 January 2023; Received in revised form 9 April 2024; Accepted 31 May 2024
Available online 22 June 2024
0304-3878/

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.

The above evidence suggests teens may additionally adjust their behavior on non-academic dimensions, including decisions about childbearing. Although, the extent to which the introduction of SPP affects teen fertility rates depends on whether those who were affected academically by SPP would have also been on the margin of having a teenage pregnancy and whether students are forward-looking enough with their decisions about sexual activity to respond to a change in their opportunities for college attendance. 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 or increase secondary school attendance. 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 (Londoño-Vélez et al., 2020b). 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 itself 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.[1] 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 qualitatively similar when using a municipality's distance to the nearest SPP-eligible higher education institution instead of variation in SPP eligibility rates, and when using women aged 15–19 with completed education less than sixth grade (and thus likely a school dropout) as a comparison group instead of women aged 25–29. We also rule out that factors like differential migration or 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 suggests 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 in the United States (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

[1] Fig. A.2 shows the pre-SPP correlation between adolescent fertility and income inequality and access to higher education in Colombia.

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 et al., 2015, 2021; Muralidharan and Prakash, 2017; Bandiera et al., 2020; Giacobino et al., 2023). 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., 2020b; 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 et al., 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.[2] These "higher-than-expected" adolescent fertility rates observed in Latin American 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.[3] 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.[4] 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.[5] In 2012, the national government additionally launched a strategic framework to address the issue comprehensively, articulating different actors within the public sector.[6] 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 et al., 2021).[7]

2.2. Ser pilo paga and higher education in Colombia

Ser Pilo Paga was announced by surprise on October 1st of 2014 by Colombia's national government. 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., 2020b).[8] 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. While technically not a requirement for high school graduation, 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 was placed at approximately the 91st percentile for the first two cohorts benefiting from SPP, and for the last two cohorts, it increased to around the 95th percentile.

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 (out of 100, with higher values implying a better-off situation) 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

[2] 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 Fig. A.1. Data are from the World Bank's World Development Indicators.

[3] The age of consent in Colombia is 14 years old.

[4] See part three of Vargas Trujillo, Flórez, Cortés, and Ibarra (2019) for a recent review.

[5] *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.

[6] National Department of Planning (DNP). *Documento CONPES Social* No. 147.

[7] 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.

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

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 (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. In 2012, for example, only 11 percent of first-year undergraduate students had a government-sponsored student loan (Ferreyra et al., 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 these 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 records from the Colombian Institute for the Assessment of Education (ICFES), which contain student-level information about the national standardized high school exit exam, SABER 11, including test scores and socio-demographic characteristics. Importantly, these data include 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 (Acevedo and Bornacelly, 2014), the Ministry of Education, and DANE. Any other sources of information used throughout the text are mentioned when they are first introduced.

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. 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 almost 90 percent of births in our sample, this is equivalent to assuming that conception occurred nine months before the reported date of birth.[9]

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 (per 1,000 students).[10] We then separate the sample at the (unweighted) 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.[11] 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 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 with SABER 11 information in 2014. Our final analysis sample consists of a balanced panel of 1,106 municipalities (out of 1,122 in the country) for conception years 2008–2019.

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

[9] In practice, we observe the gestational age at delivery in week intervals. Using our Colombian data from 2010–2012, we map these intervals to 2010–2012 U.S. National Vital Statistics data (in which we observe weekly gestational age) and randomly assign a single gestational age week to each birth using the U.S. *within*-interval, age-specific distribution of weekly gestational ages. This procedure respects the interval shares in the Colombian data for each age group. We convert weeks to months by rounding $\{assigned\ weeks/4.348\}$ to the nearest integer. Finally, we subtract the resulting number from the observed year-month of birth.

[10] 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.

[11] SPP's official regulations required students to be registered in the SISBEN database with a score below the cutoffs mentioned earlier by the time the program was announced. SISBEN levels are associated with the thresholds to qualify for public health insurance and tend to be more salient than the actual raw SISBEN score. The public health insurance cutoffs follow the same area-specific definitions as the ones for SPP, but the maximum scores to qualify for SPP were slightly above the maximum scores to qualify for public health insurance. For the 14 main metropolitan areas, the public health insurance cutoff was 54.86 (compared to 57.21 for SPP). For other urban areas and rural areas, cutoffs were 51.57 and 37.80, respectively (compared to 56.32 and 40.75 for SPP). Using data from Londoño-Vélez et al. (2020a), we estimate that if we were to use an error-free eligibility measure based on SISBEN levels 1 and 2, we would capture 94.3% of SISBEN-eligible students overall (94.5% in main metro areas, 93.9% in other urban areas, and 95.5% in rural areas). We observe a self-reported SISBEN level.

Table 1
Summary statistics of key variables.

	Treatment municipalities	Comparison municipalities	All
Panel A. Births per 1,000 women			
Pre-SPP			
(1) Age 15–19	70.658	75.751	72.160
	(20.336)	(29.546)	(23.544)
(2) Age 25–29	82.796	79.147	81.800
	(20.601)	(27.605)	(22.786)
Ratio [(1)/(2)]	0.853	0.957	
Post-SPP			
(3) Age 15–19	57.547	67.110	60.421
	(22.024)	(29.661)	(24.956)
(4) Age 25–29	79.020	78.258	78.812
	(21.240)	(27.078)	(22.977)
Ratio [(3)/(4)]	0.728	0.858	
Percentage change (Post-SPP vs. Pre-SPP)			
Age 15–19	−18.6	−11.4	−16.3
Age 25–29	−4.6	−1.1	−3.7
Ratio	−14.7	−10.3	
Panel B. SPP eligibility rates per 1,000 female students (2014)			
	26.811	6.520	21.596
	(15.545)	(4.854)	(16.252)
Number of municipalities	553	553	1,106

Notes: This table shows means and standard deviations (in parentheses) for age-specific municipality-level birth rates (Panel A) and SPP eligibility rates (Panel B) for treatment and comparison municipalities. In Panel A, we also show the birth rate ratio between teens and non-teens. Birth rates are averages for relative years since SPP's announcement and are weighted using each municipality's average annual age-specific female population. Birth rates are adjusted for the lag between conception and birth date. The pre-SPP period includes relative years from 2008Q4-2009Q3 (−6) to 2013Q4-2014Q3 (−1), while the post-SPP period ranges from 2014Q4-2015Q3 (0) to 2018Q4-2019Q3 (4). 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.

1,000 female students in 2014. Fig. A.3 plots the full distribution of SPP eligibility rates for the municipalities in our sample. About 37 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. (2020b) 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.

Fig. A.4 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 A.1 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 the absence of SPP, these differences do not invalidate our empirical strategy. Fig. A.6 shows the evolution of selected relevant municipality characteristics using a municipality-level panel constructed from different available data sources.[12] The figures show that, for these characteristics, our treatment and comparison municipalities mostly display similar trends through the observed period, providing some confidence that our results are not explained by changing economic activity or other conditions between our municipality groups. Fig. A.4 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.

[12] Panels (a) to (d) come from the CEDE municipal panel (Acevedo and Bornacelly, 2014), panel (e) comes from DANE, panels (f) and (g) come from the Ministry of Education, and panel (h) comes from Prem, Vargas, and Namen (2023a).

Table 1 also displays means of municipality-level fertility rates during the pre-SPP and post-SPP periods 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. Relative to the comparison municipalities, treatment municipalities have lower fertility rates for women aged 15–19, but slightly higher fertility rates for women aged 25–29. Fig. A.5 plots the complete distribution of adolescent fertility rates, which shows a substantial amount of overlap between the distributions of treatment and comparison municipalities. Finally, Fig. 1 shows the raw trends in average fertility rates between age groups for both treatment and comparison municipalities. This figure mimics our empirical approaches discussed in the next section.

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 describes the identifying assumptions behind our research designs, introduces our target parameter of interest, and 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 in our setting is that the proportional change in the average teenage birth rate observed in comparison municipalities provides a good counterfactual for the change in treatment municipalities

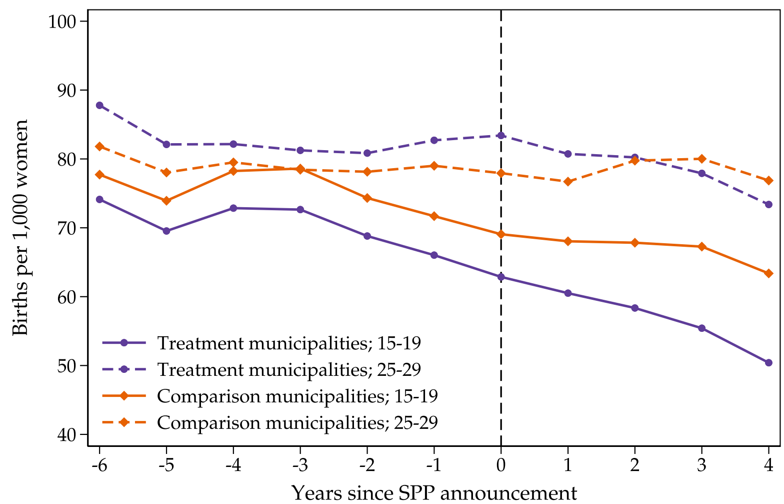


Fig. 1. Trends in Fertility Rates for 15–19 Year Olds and 25–29 Year Olds by Treatment and Comparison Municipalities.
Notes: This figure plots trends in average fertility rates between age groups for both treatment and comparison municipalities. Treatment municipalities are above the median in female eligibility rates for SPP in 2014, while comparison municipalities are below the median. The horizontal axis is measured in years relative to October 1, 2014, the date in which SPP was first announced. For example, the relative year $t = 0$ (the announcement year) consists of the last quarter of 2014 and the first three quarters of 2015. The annual means weight municipalities by the average annual population of women in each municipality and age group.

in the absence of SPP. More explicitly, using the potential outcomes framework, let $Y_{mt}(0)$ and $Y_{mt}(1)$ denote the potential adolescent fertility rates for municipality m in period t in a world with and without SPP, respectively. In addition, let $SPP^*_m = \mathbb{1}\left[SPP_m > \text{median}\left(SPP_m\right)\right]$ denote treatment and comparison municipalities, with SPP_m being the rate of female students eligible for the program in a given municipality in 2014. Finally, let the indicator variable $Post_t$ take the value of one for the periods after the announcement of SPP. Following Wooldridge (2023) and Chen and Roth (2023), our parallel trends assumption for the difference-in-differences strategy can formally be expressed as:

$$\frac{\mathrm{E}[Y_{mt}(0) \mid SPP^*_m = 1, Post_t = 1]}{\mathrm{E}[Y_{mt}(0) \mid SPP^*_m = 1, Post_t = 0]} = \frac{\mathrm{E}[Y_{mt}(0) \mid SPP^*_m = 0, Post_t = 1]}{\mathrm{E}[Y_{mt}(0) \mid SPP^*_m = 0, Post_t = 0]}. \tag{1}$$

Eq. (1) represents a parallel trends assumption in terms of growth in average fertility rates in the absence of SPP (Wooldridge, 2023). The post-SPP average adolescent fertility rate in treatment municipalities in the absence of the program (an unobserved quantity), $\mathrm{E}[Y_{mt}(0) \mid SPP^*_m = 1, Post_t = 1]$, is identified using the evolution of fertility rates in comparison municipalities (the ratio on the right side of the equation) and a standard no anticipation assumption:

$$\mathrm{E}[Y_{mt}(0) \mid SPP^*_m = 1, Post_t = 0] = \mathrm{E}[Y_{mt}(1) \mid SPP^*_m = 1, Post_t = 0]. \tag{2}$$

Once we recover the counterfactual $\mathrm{E}[Y_{mt}(0) \mid SPP^*_m = 1, Post_t = 1]$, our target parameter is the average proportional treatment effect on the treated in the post-period (Chen and Roth, 2023):

$$ATT\% = \frac{\mathrm{E}[Y_{mt}(1) \mid SPP^*_m = 1, Post_t = 1] - \mathrm{E}[Y_{mt}(0) \mid SPP^*_m = 1, Post_t = 1]}{\mathrm{E}[Y_{mt}(0) \mid SPP^*_m = 1, Post_t = 1]}, \tag{3}$$

which quantifies the percentage change in the average adolescent fertility rate between treatment and comparison municipalities.[13]

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 relative 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 relative outcomes in municipalities with lower SPP eligibility rates (comparison municipalities). This is an extension of the parallel trends assumption underlying the difference-in-differences design we discussed before. Again, more formally, our identifying assumption for the triple difference strategy is:

$$\frac{\dfrac{\mathrm{E}[Y_{amt}(0) \mid a = \text{15-19}, SPP^*_m = 1, Post_t = 1]}{\mathrm{E}[Y_{amt}(0) \mid a = \text{25-29}, SPP^*_m = 1, Post_t = 1]}}{\dfrac{\mathrm{E}[Y_{amt}(0) \mid a = \text{15-19}, SPP^*_m = 1, Post_t = 0]}{\mathrm{E}[Y_{amt}(0) \mid a = \text{25-29}, SPP^*_m = 1, Post_t = 0]}} = \frac{\dfrac{\mathrm{E}[Y_{amt}(0) \mid a = \text{15-19}, SPP^*_m = 0, Post_t = 1]}{\mathrm{E}[Y_{amt}(0) \mid a = \text{25-29}, SPP^*_m = 0, Post_t = 1]}}{\dfrac{\mathrm{E}[Y_{amt}(0) \mid a = \text{15-19}, SPP^*_m = 0, Post_t = 0]}{\mathrm{E}[Y_{amt}(0) \mid a = \text{25-29}, SPP^*_m = 0, Post_t = 0]}}, \tag{4}$$

where now Y_{amt} denotes the fertility rate of age group $a \in \{15\text{–}19, 25\text{–}29\}$ in municipality m and period t, and everything else is defined as in Eq. (1). Note that, as before, Eq. (4) only requires *one* parallel trends assumption to hold, but this time, it is in terms of the relative fertility rate between the two age groups in the absence of the policy (Olden and Møen, 2022). Our counterfactual of interest is still the post-SPP mean adolescent fertility rate for municipalities above the median eligibility rate, denoted here as $\mathrm{E}[Y_{amt}(0) \mid a = 15\text{–}19, SPP^*_m = 1, Post_t = 1]$.[14]

[13] This $ATT\%$ is not an average of municipality-level percentage changes. See the discussion in Chen and Roth (2023) for reference. Using the summary statistics in Table 1, our identifying assumption in Eq. (1) implies a raw difference-in-differences $ATT\%$ of −8.1%, calculated as $(57.547/(70.658 \times (1 - 0.114))) - 1$.

[14] Using the summary statistics in Table 1, our identifying assumption in Eq. (4) implies a raw triple difference $ATT\%$ of −4.8%, calculated as $(0.728/(0.853 \times (1 - 0.103))) - 1$.

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 likely 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.

We highlight that our parallel trends assumptions in Eqs. (1) and (4) could have alternatively been introduced in terms of *absolute* changes (*i.e.*, in levels).[15] We see our assumptions in terms of proportional changes as more plausible considering the differences in average fertility rates in the pre-SPP period shown in Table 1. This is especially relevant when we further disaggregate our treatment and comparison groups to conduct our heterogeneity analysis, and pre-SPP average teen fertility rates become more dissimilar across subgroups. As mentioned by Chen and Roth (2023), time-varying factors likely have unequal *level* effects on the outcome when pretreatment means are different across groups. Nonetheless, we discuss the robustness of our results to this alternative parallel trends assumption in Section 6.

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 major events that occurred in the country around 2014 cannot explain our results in Section 6. In terms of policies, there are also other Colombian programs that launched around the same time and share the same eligibility cutoff on the SISBEN. In particular, these include Vivienda Rural, which provided rural housing building; BEPS, a savings program for the elderly without a pension; and Access-Icetex, a long-term credit program for tertiary education. Importantly, none of these programs shared the same start year as SPP and the number of beneficiaries for these programs did not change substantially when SPP began.[16] We also note that while SPP shares a similar eligibility cutoff on the SISBEN with these other programs, SPP also included an eligibility cutoff on the SABER 11 exam which is not shared by these other programs and was generally the more binding criteria for eligibility in the sense that single dimensional eligibility rates are much lower for the SABER 11 criteria than the SISBEN criteria.

Additionally, differential migration patterns between our groups of municipalities could be problematic for our research design. In Table A.6, we document how migration patterns differ between treatment and comparison municipalities using cross-sectional information from the full-count 2018 Colombian census. The results show that the likelihood of having moved municipalities within the last year or five years (relative to 2018) is very similar between residents of treatment and comparison municipalities, even for those in the age groups we study in our primary analyses.

4.2. Estimation approaches

We implement the difference-in-differences design with an event-study specification using an exponential mean function and Poisson regression on the sample of 15–19 year olds. More specifically, we estimate the following model using Poisson quasi-maximum likelihood (QMLE), which allows us to consistently estimate our target parameter (Wooldridge, 2023; Chen and Roth, 2023):[17]

$$Y_{mt} = \exp\left(SPP_m^* \times \sum_{\substack{\tau=-6 \\ \tau\neq-1}}^{4} \alpha_\tau \mathbb{1}[t=\tau] + \theta_m + \theta_{d(m)t} \right) \epsilon_{mt}, \tag{5}$$

where, following the notation introduced earlier, Y_{mt} is the teen fertility rate in municipality m and year $t \in [-6, 4]$, which is measured in years relative to October 1, 2014, the date in which SPP was first announced. For example, the relative year $t = 0$ (the announcement year) consists of the last quarter of 2014 and the first three quarters of 2015. Again, $SPP_m^* = \mathbb{1}\left[SPP_m > \text{median}\left(SPP_m\right)\right]$ denotes treatment and comparison municipalities. The terms θ_m and $\theta_{d(m)t}$ are municipality fixed effects and year fixed effects that we allow to be department-specific, respectively.[18] Finally, ϵ_{mt} is an error term.

In a similar fashion, we implement our preferred triple difference design with the following model that we also estimate by Poisson QMLE but using the sample of 15–19 *and* 25–29 year olds:

$$Y_{amt} = \exp\left(Teen_a \times SPP_m^* \times \sum_{\substack{\tau=-6 \\ \tau\neq-1}}^{4} \beta_\tau \mathbb{1}[t=\tau] + \gamma_{am} + \gamma_{mt} + \gamma_{ad(m)t} \right) \varepsilon_{amt}, \tag{6}$$

where Y_{amt} is the fertility rate of age group $a \in \{15\text{–}19, 25\text{–}29\}$ in municipality m and year $t \in [-6, 4]$, again measured in years relative to the announcement of SPP. $Teen_a$ is an indicator for age group defined as $Teen_a = \mathbb{1}\left[a = 15\text{–}19\right]$. The terms γ_{am} and γ_{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, ε_{amt} is an error term.

From Eqs. (5) and (6), $\exp\left(\alpha_{\tau,t\geq0}\right) - 1$ and $\exp\left(\beta_{\tau,t\geq0}\right) - 1$ represent the proportional reduction in the average adolescent fertility rate in treatment municipalities at time $t = \tau$ after the introduction of SPP (*i.e.*, the $ATT\%$ at $t = \tau$). For the estimation, we use the year before the announcement of SPP ($t = -1$) as our reference period.

For Eq. (6), 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 $\left(\gamma_{am}\right)$ 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 $\left(\gamma_{mt}\right)$ control for municipality-specific trends in fertility rates common to all age groups. Finally, the age group by department year effects $\left(\gamma_{ad(m)t}\right)$ 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 Eqs. (5) and (6) that replaces the

[15] For instance, Eq. (1) would be $\mathrm{E}[Y_{mt}(0) \mid SPP_m^* = 1, Post_t = 1] - \mathrm{E}[Y_{mt}(0) \mid SPP_m^* = 1, Post_t = 0] = \mathrm{E}[Y_{mt}(0) \mid SPP_m^* = 0, Post_t = 1] - \mathrm{E}[Y_{mt}(0) \mid SPP_m^* = 0, Post_t = 0]$ under this alternative assumption.

[16] See Table 2 in Laajaj et al. (2022) who provide a comprehensive overview of the other Colombian programs that share the same SISBEN eligibility cutoff as SPP, including the number of annual beneficiaries for these programs.

[17] Recent studies in the literature implementing difference-in-differences designs with similar modeling approaches include (Lindo and Packham, 2017), Fischer, Royer, and White (2018), and Farin, Hoehn-Velasco, and Pesko (2023). More generally, recent applied literature studying birth rates using log-linear (or log-like-linear) specifications include Dettling and Kearney (2023), Kelly, Lindo, and Packham (2020), and Kearney and Levine (2015). A more methodological discussion can be found in Wooldridge (2023), Chen and Roth (2023), and Kahn-Lang and Lang (2020).

[18] Departments in Colombia are similar to states in the United States. A group of municipalities forms each department.

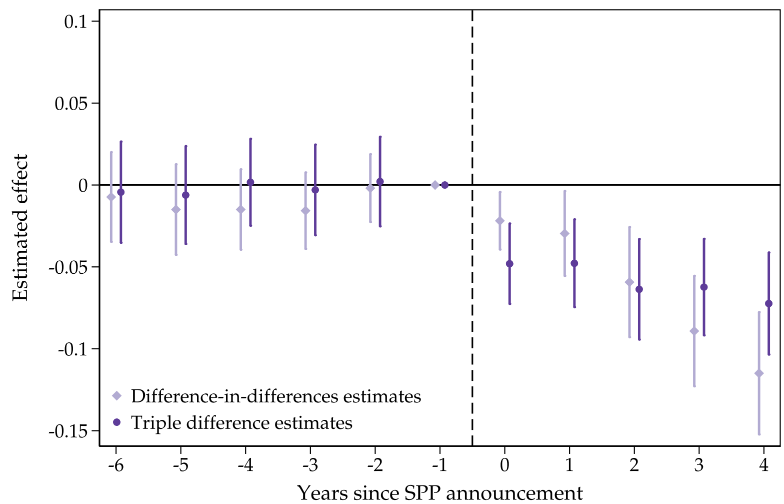


Fig. 2. Event Study Estimates of the Effect of *Ser Pilo Paga* on Teen Fertility Rates.
Notes: This figure plots the difference-in-differences estimates of $\exp(\alpha_\tau) - 1$ from Eq. (5) and the triple difference event study estimates of $\exp(\beta_\tau) - 1$ from Eq. (6). The dots and diamonds represent the estimated proportional effects and the vertical lines represent 95 percent confidence intervals. All estimates use the number of births as the outcome and the population of women as the exposure variable. Standard errors are clustered at the municipality level.

year indicators with a single post-SPP indicator variable. Specifically, we estimate the following two equations by Poisson QMLE:

$$Y_{mt} = \exp\left(\alpha\left(SPP_m^* \times Post_t\right) + \theta_m + \theta_{d(m)t}\right)\xi_{mt}, \text{ and} \tag{7}$$

$$Y_{amt} = \exp\left(\beta\left(Teen_a \times SPP_m^* \times Post_t\right) + \gamma_{am} + \gamma_{mt} + \gamma_{ad(m)t}\right)u_{amt}, \tag{8}$$

where $Post_t = \mathbb{1}[t \geq 0]$ and everything else is defined as in Eqs. (5) and (6). In Eqs. (7) and (8), $\exp(\alpha) - 1$ and $\exp(\beta) - 1$ are the summary difference-in-differences and triple difference proportional effects across all post-SPP years, respectively.

We use the number of births as the outcome and the population of women in each municipality and age group as the exposure variable when modeling the birth rates in all our Poisson regressions. Including the population exposure allows us to account for the large population variation across municipalities in Colombia and, therefore, for the underlying differential potential for births (Fischer et al., 2018). An alternative would be to directly use the birth rate as the outcome and the population of women as weights during estimation. These two approaches lead to numerically equivalent results, as noted in Farin et al. (2023). In all our regressions, we cluster the standard errors at the municipality level (the level of the treatment). Finally, we estimate all our Poisson models using Correia, Guimarães, and Zylkin (2020)'s routine.

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

Fig. 2 shows the event study estimates of $\exp(\alpha_\tau) - 1$ from the simple difference-in-differences specification in Eq. (5) 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 five percent level. We conduct a test of pre-period trends by estimating the difference-in-differences specification in Eq. (5) using only pre-SPP observations and running a joint test of these coefficients.[19] With a p-value of 0.763, this Wald-type test cannot reject the null hypothesis that the pre-period placebo effects are jointly equal to zero. Additionally, using the methods from Roth (2022), we have a high level of power to detect relatively small hypothetical pre-period trends.[20]

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 five percent level. Column 1 of Table 2 presents the summary estimate from the difference-in-differences design, indicating that after SPP was introduced fertility rates of women aged 15–19 in treatment municipalities decreased by about 5.3 percent relative to comparison municipalities.

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. The student-weighted mean SPP eligibility rates for the 2nd through 4th quartile are 9.128, 20.156, and 47.550, respectively. Column 3 of Table 2 reports the summary results, but we also show the event-study equivalent of this specification in Fig. A.8, Panel (a).

Relative to municipalities in the 1st quartile of SPP eligibility, we estimate that 2nd quartile municipalities experienced a 1.4 percent decrease (not statistically significant) in adolescent fertility. Meanwhile, we estimate that the 3rd and 4th quartile municipalities experienced a 5.8 and 7.1 percent decrease, respectively. The estimates for the

[19] Formally, we test $\exp(\alpha_\tau) - 1 = 0$ for $\tau \in [-5, -1]$, with $t = -6$ as the reference group. The estimation sample is $t \in [-6, -1]$.

[20] Specifically, we calculate the least extreme linear pre-trend where the conditional expectations after pre-testing are within at least three of the confidence intervals of our event study estimates. The power to detect such a hypothetical trend is 0.88—above the standard 0.80 used as a benchmark in power calculation analyses (Roth, 2022). Thus, we conclude that the type of pre-trends needed to explain away our results are unlikely to be undetected using standard pre-testing methods in our setting.

Table 2
Summary difference-in-differences and triple difference estimates of the effect of *Ser Pilo Paga* on teen fertility rates.

	Fertility rate			
	(1)	(2)	(3)	(4)
$SPP \times Post$	−0.053*** (0.013)			
$Teen \times SPP \times Post$		−0.057*** (0.010)		
$SPPQuartile \times Post$				
1st quartile			[Reference]	
2nd quartile			−0.014 (0.020)	
3rd quartile			−0.058*** (0.015)	
4th quartile			−0.071*** (0.017)	
$Teen \times SPPQuartile \times Post$				
1st quartile				[Reference]
2nd quartile				−0.035** (0.016)
3rd quartile				−0.077*** (0.013)
4th quartile				−0.087*** (0.015)
Observations	12,166	24,332	12,166	24,332
Treatment municipalities	553	553	–	–
Comparison municipalities	553	553	–	–
Pre-trends test p-value	0.763	0.988	–	–

Notes: This table presents summary difference-in-differences and triple difference estimates using Eq. (7) and Eq. (8), 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. The student-weighted mean eligibility rates (per 1,000 female students) for the 1st through 4th quartile are 0, 9.128, 20.156, and 47.550, respectively. In all cases, reported estimates correspond to proportional reductions calculated as $\exp(\hat{\alpha}) - 1$ and $\exp(\hat{\beta}) - 1$ for difference-in-differences and triple difference models, respectively. The pre-trends test p-values in columns 1 and 2 come from a joint test of the pre-period placebo effects conducted using only pre-SPP observations. The Poisson models are estimated using the number of births as the outcome and the annual population of women in each municipality and age group as the exposure. Standard errors are clustered at the municipality level and presented in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

3rd and 4th quartile municipalities are both statistically different from the estimate for the 2nd quartile at the one percent significance level. These results highlight that our estimates are not dependent upon how we split municipalities into treatment and comparison groups. These results also suggest the effects are non-linear in eligibility rates: fertility rate effects are larger for municipalities with a higher eligibility rate, but at a decreasing rate.

5.2. Triple difference estimates

Fig. 2 also displays our preferred estimates of the fertility impacts of SPP ($\exp(\beta_\tau) - 1$) using the triple difference event-study specification in Eq. (6). Similar to the difference-in-differences results, the pre-period coefficients are again close to zero and not statistically significant. A Wald-type test of the null hypothesis that the pre-period effects are jointly zero using only pre-SPP observations produces a p-value of 0.988.[21] Also, using the methods from Roth (2022), we again conclude that it would be unlikely that we would be unable to detect a relatively conservative differential pre-period trend.[22] These results provide 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 five percent level. We report the summary estimates using Eq. (8) 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. 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 Eq. (6) using a quarterly-level data set. 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 Fig. 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 Fig. A.7 and in Table 2. The estimates from the event-study equivalent of this specification are shown in Fig. A.8, Panel (b). Relative to municipalities in the 1st quartile of eligibility, 2nd quartile municipalities experienced

[21] Formally, we test $\exp(\beta_\tau) - 1 = 0$ for $\tau \in [-5, -1]$, with $t = -6$ as the reference group. The estimation sample is $t \in [-6, -1]$.

[22] Similar to how we approached the difference-in-differences results, we calculated the least extreme linear pre-trend where the conditional expectations after pre-testing are within at least three of the confidence intervals of our event study estimates. We calculate that we have a power of 0.89 to detect such a hypothetical pre-trend.

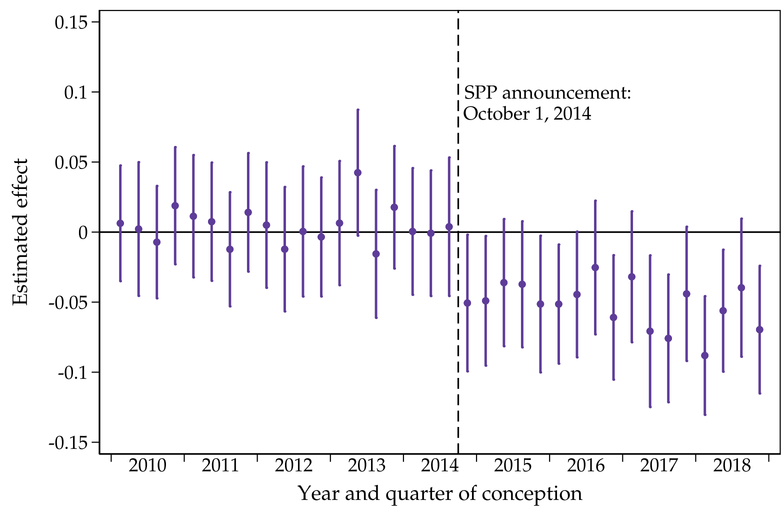


Fig. 3. Triple Difference Event Study Estimates of the Effect of *Ser Pilo Paga* on Teen Fertility Rates Using Quarterly Data.
Notes: This figure plots the triple difference event study estimates of $\exp(\beta_\tau) - 1$ from Eq. (6) using quarterly data instead of annual data. The quarters in 2008 are the reference period. Only estimates for four years around 2014 are plotted. The dots represent the estimated effects and the vertical lines represent 95 percent confidence intervals. All estimates use the number of births as the outcome and the population of women as the exposure variable. Standard errors are clustered at the municipality level.

a 3.5 percent decrease in fertility rates, while the 3rd and 4th quartile municipalities experienced a 7.7 and 8.7 percent decrease, respectively. The estimates for each quartile are statistically different from the 1st quartile at the five percent level (one 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 one percent level.

These results highlight a few important points. First, our main estimates are 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. Finally, the effects of SPP are non-linear in eligibility rates.

5.3. Mechanisms

In this section, we aim to assess the extent to which our estimates are explained by direct effects—where receipt of the scholarship itself drives the results through incapacitation (via college attendance) or income effects—or by indirect effects—where results are driven by behavioral responses before being able to receive the scholarship, such as motivational effects, changing opportunity costs, peer effects, or incapacitation effects from increased secondary school attendance. 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 number of adolescent births 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 24,709 fewer births to teenage mothers as a result of the introduction of SPP in treatment municipalities.[23] 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 can, at most, explain around 70% of the effects we observe.[24]

Next, returning to Fig. 3, the timing of the effects we observe is also informative for the mechanisms driving the results. If incapacitation effects or income effects from receiving the scholarship 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 effects directly from receiving the scholarship, including within-household transfers from scholarship recipients to younger siblings, 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 Eq. (7) except with the age-specific fatherhood rates per 1,000 young

[23] We calculate this number as:

$$\sum_{\tau=0}^{4}\left(AdolescentBirths_t \mid SPP_m^* = 1, \right.$$
$$\left. t = \tau - \frac{AdolescentBirths_t \mid SPP_m^* = 1, t = \tau}{1 + \widehat{ATT\%}_\tau} \right), \tag{9}$$

with $\widehat{ATT\%}_\tau = \exp\left(\hat{\beta}_\tau\right) - 1$ from Eq. (6).

[24] We estimate that only about 86% of the overall number of SPP recipients came from treatment municipalities and, therefore, we are being conservative with the numerator in this calculation.

Table 3
Summary difference-in-differences estimates of the effect of *Ser Pilo Paga* on teen fatherhood.

	Births per 1,000 men			
Age group:	All teens (15–19)	15–17	18	19
	(1)	(2)	(3)	(4)
$SPP \times Post$	−0.075***	−0.070**	−0.104***	−0.059***
	(0.019)	(0.030)	(0.023)	(0.019)
Observations	12,166	12,166	12,166	12,166
Treatment municipalities	553	553	553	553
Comparison municipalities	553	553	553	553
Pre-trends testing p-value	0.486	0.338	0.498	0.312
Pre-SPP share of teen fathers	100	25.9	33.6	40.5

Notes: This table presents summary difference-in-differences estimates using Eq. (7) with the number of births per 1,000 men in each age group as the outcome. 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. In all cases, reported estimates correspond to proportional reductions calculated as $\exp(\hat{\alpha}) - 1$. The pre-trends test p-values in all columns come from a joint test of the pre-period placebo effects conducted using only pre-SPP observations. The Poisson models are estimated using the number of births as the outcome and the annual population of men in each municipality and age group as the exposure. Standard errors are clustered at the municipality level and presented in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

men as the outcome. 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 7.5 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 about 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 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. Our data limit us from precisely identifying the most relevant source of the possible indirect effects we describe. But, our results are 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.

Kearney and Levine (2014) show that rates of teen childbearing are closely related to income inequality, theorizing that feelings of greater economic hopelessness arise in the context of high income inequality, leading economically disadvantaged young women to perceive the opportunity costs of early childbearing to be low. Following this logic, we expect the fertility effects of SPP to be larger in municipalities that ex-ante had higher levels of inequality, where the chances of economic mobility were arguably more limited. We use a Gini coefficient as a proxy of income inequality, measured in 2005 at the municipality level. Since in Kearney and Levine (2014)'s theoretical framework the relationship between inequality and teen childbearing holds after conditioning on socioeconomic status (*i.e.*, low-income women in more unequal places have worse perceptions of economic success than low-income women in less unequal places), we first residualize the Gini coefficient by regressing it on the 2005 municipal poverty incidence and then rank municipalities using this residualized measure. We then separately estimate Eq. (8) for those below the median (lower inequality) and above the median (higher inequality). These analyses also inform our main results, where we largely interpret the fertility effects of SPP as an increase in economic opportunities.

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 effects 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 (1,040 out of 1,106 in the main estimation sample). The estimated reduction in teen fertility in municipalities with higher levels of income inequality is 2.8 percentage points greater than in those with lower inequality (column 3 versus column 2). However, we cannot reject the null hypothesis that these two effects are the same using a two-sided test at conventional levels (p-value = 0.149).

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 five percent of the test takers and at least ten students in this analysis. This limits the estimation sample to 658 municipalities.[25]

[25] 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 of the effect of *Ser Pilo Paga* on teen fertility rates by group of municipalities.

	Fertility rate		
Municipalities:	All	Below median	Above median
	(1)	(2)	(3)
Panel A. Baseline income inequality			
$Teen \times SPP \times Post$	−0.056***	−0.028**	−0.056***
	(0.010)	(0.014)	(0.013)
Observations	22,880	11,484	12,848
Treatment municipalities	522	288	265
Comparison municipalities	518	234	319
Pre-trends test p-value	0.975	0.708	0.987
Panel B. Baseline college-going expectations			
$Teen \times SPP \times Post$	−0.057***	−0.073***	−0.040***
	(0.012)	(0.017)	(0.014)
Observations	14,476	7,238	7,238
Treatment municipalities	390	179	211
Comparison municipalities	268	150	118
Pre-trends test p-value	0.963	0.245	0.325
Panel C. Baseline share of teen births from adolescent fathers			
$Teen \times SPP \times Post$	−0.059***	−0.002	−0.069***
	(0.011)	(0.015)	(0.015)
Observations	24,332	12,100	12,232
Treatment municipalities	553	243	310
Comparison municipalities	553	307	246
Pre-trends test p-value	0.964	0.446	0.903

Notes: This table presents summary triple difference estimates using Eq. (8) for a subgroup of municipalities indicated in each column. In all cases, reported estimates correspond to proportional reductions calculated as $\exp(\hat{\beta}) - 1$. In Panel A, income inequality is measured by the residualized 2005 municipal Gini obtained from CEDE's municipal panel (see text for details). 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 2009–2013. Two-sided p-value from a χ^2-test of equality of the percentage reduction in fertility rates, H_0 [2] = [3]: Panel A p-value = 0.149; Panel B p-value = 0.138; Panel C p-value = 0.001. The pre-trends test p-values in all panels/columns come from a joint test of the pre-period placebo effects conducted using only pre-SPP observations. The Poisson models are estimated using the number of births as the outcome and the annual population of women in each municipality and age group as the exposure. Standard errors are clustered at the municipality level and presented in parentheses. (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Finally, we group municipalities by whether they are above or below the median share of students who meet this criteria and estimate Eq. (8) separately for these two subsets of municipalities.

The triple difference coefficients are presented in Panel B of Table 4. First, we cannot reject the null that all pre-period effects for both subsets of municipalities are zero. Second, the overall coefficient in column 1 is equal to the one in our main analysis. We then estimate a reduction in fertility over three percentage points larger 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 test at conventional levels (p-value = 0.138).

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 Fig. A.9 a slightly more detailed version of the results in Table 4. There, we divide municipalities into deciles according to the level of each characteristic discussed before. We then use Eq. (8) to estimate the triple difference reduction in fertility for each decile. Similar to Muralidharan and Prakash (2017), we plot these estimates together with a lowess-smoothed line of the triple difference effects 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 supports 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 comparison 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 A.3 also shows that these results hold when using a linear model estimated with ordinary least squares (OLS) to calculate the reduction in the average teen fertility rate and the associated $ATT\%$. Our linear model implies an $ATT\%$ of −5.3%, very similar to our main result of −5.7%.[26] 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

[26] The implied $ATT\%$ is $-5.3\% = ATTinLevels/Counterfactual = ATTinLevels/(Observed - ATTinLevels) = -3.243/(57.547 + 3.243)$. See also Fig. A.10 with the event study estimates using the linear model estimated by OLS.

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 that uses whether a municipality is above the median SPP recipient rate in 2014 (not gender-specific), which we observe more precisely than eligibility rates. Our SPP eligibility rates rely on a component that is directly observable to us (the SABER 11 eligibility) and one that is a proxy likely measured with some error (the SISBEN eligibility). We obtained the actual number of SPP recipients among the 2014 cohort of students from the Ministry of Education for each municipality and calculated an equivalent recipient rate as the number of SPP recipients per 1,000 SABER 11 test takers. The summary estimate, reported in column 2 of Table A.4, is very similar to our main estimate.[27] We plot the event study estimates from this analysis in Fig. A.11. Our eligibility rates have the advantage of being disaggregated by gender, which is important given the geographic variation in gender performance gaps in Colombia (Abadía and Bernal, 2017), and also of not being affected by final take-up or enrollment decisions. Our main interest is to capture the underlying possibility of receiving the scholarship and this is better captured by eligibility rates.

Second, we also estimate a triple difference model similar to Eq. (6) which replaces SPP^*_m with an indicator for whether a municipality is below the median distance (closer) to the nearest SPP-eligible institution.[28] 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 Fig. A.12. The results show that the age group differentials in birth rates trend similarly, with no clear pattern, between municipalities that are closer and farther from SPP-eligible institutions before the scholarship announcement. 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 3 of Table A.4, implies that municipalities closer to SPP-eligible institutions experienced a 9.2 percent decrease in teen fertility rates after SPP was introduced relative to municipalities that are farther away.

Finally, 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 fifth grade or less as a comparison group. We replace $Teen_a$ in Eq. (6) with an indicator for having completed sixth grade or more using only data on births to women aged 15–19.[29] The intuition is that any woman aged 15–19 whose highest grade completed is fifth grade (*i.e.*, primary school) 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 some years of secondary education (*i.e.*, sixth grade or higher).

The results of this approach are reported in Fig. A.13. Panel (a) shows, separately for treatment and comparison municipalities, the trends in the differential in number of births between women age 15–19 who have completed sixth grade or higher and those who have completed only fifth grade or lower.[30] The plot shows that births to women with sixth grade or higher education are increasing over time relative to women with less than fifth grade education. But these differentials trend very similarly between treatment and comparison municipalities before SPP. Right after SPP's introduction, 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 compare the changes in the number of births to women aged 15–19 overtime between those completing sixth grade or more versus fifth 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 after October 2014. The summary estimate, presented in column 4 of Table A.4, implies that births to adolescent women who have completed at least sixth grade decreased by almost 12 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.

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 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 C.1. 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

We have shown robust evidence that relative 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. We perform an additional 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 A.5, we use 2008–2014 data and estimate the same specification in Eq. (8), pretending that SPP was introduced in the years 2008–2013. The overall estimated effects after each of the placebo treatment years are always statistically insignificant and close to zero.

The assumption of parallel trends is an assumption about counterfactuals and, therefore, ultimately untestable. We can, however, test the sensitivity of our main estimates in Fig. 2 to potential deviations from the parallel trends assumption. We do this in a systematic way using the approaches developed by Rambachan and Roth (2023). We

[27] The correlation between this SPP recipient rate (not gender-specific) and our main SPP eligibility rate (female-specific) is 0.6. Both are measured using only the 2014 cohort of students.

[28] See Fig. A.4 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.

[29] We also restructure the analysis data set to have one observation per year per municipality per grade-level (*i.e.*, sixth grade or above, or fifth grade or less).

[30] We model the number of births instead of a birth rate because we cannot reliably calculate the number of young women in each municipality above or below a sixth grade level of education over time.

report the results in Fig. A.16 for the sensitivity of our estimates to (1) parallel trends violations of various sizes relative to the maximum violation observed in the pre-SPP period, and (2) allowing different extrapolations of trends between the pre and post periods. The results show that our estimates are robust to assumptions with relatively large deviations from the observed trends. For example, we would need a violation of the parallel trends of 70% the maximum observed in the pre-SPP years *across all consecutive post-SPP years* to make the average $ATT\%$ statistically insignificant. We consider this to be a very extreme breakdown value. Our results are also robust to linear violations of parallel trends and to violations that deviate from linearity by as much as 0.3 percentage points across all consecutive post-SPP years.

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 Eq. (8). We report the results in Fig. A.15. 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 Fig. A.14). We also estimate our main specification by excluding all municipalities that include a department's capital city (reported in column 2 of Table A.3). The results from each of these regressions produce estimates that are very similar to our main estimates.

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.

To aid a comparison of our estimates to existing research on the effects of other programs on teen fertility in similar settings, we divide our main estimate by the difference in eligibility rates between our treatment and comparison municipalities to get an approximation for how a change in eligibility rates would lead to a change in adolescent fertility rates. As a benchmark comparison, Berthelon and Kruger (2011)'s estimates imply that a 20 percentage point (about a 0.5 standard deviation) increase in student enrollment in full-day schools in Chile (as opposed to half-day schools) reduced teenage motherhood by about 2.8 percent. In our setting, an "equivalent" 0.5 standard deviation increase in SPP eligibility rates reduces adolescent birth rates by about 2.3 percent. Thus, we consider the impacts of SPP on adolescent fertility to be comparable in magnitude to other educational interventions, but substantially smaller than the effects of *Familias en Acción*—Colombia's conditional cash transfer program—which reduced teenage pregnancy rates by 27 percent (Attanasio et al., 2021).

Overall, 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, Michelmore, and Owen (2021)).

In 2018, under a new presidential administration, the Colombian government announced that the *Ser Pilo Paga* program would no longer accept new beneficiaries and the program would be replaced. The program gained controversy during its four years due to its high cost to the government and the fact that most SPP beneficiaries attended private institutions. Our findings illustrate important indirect benefits of SPP. Taking at face value contemporary estimates from the United Nations Population Fund regarding the overall per capita costs to society of teen pregnancies in six Latin American countries (UNFPA, 2020), including Colombia, a back-of-the-envelope calculation suggests that our estimated number of births averted in treatment municipalities would have involved a one-time cost of USD 22.3 million to the Colombian health care system (around USD 901 per pregnancy). Moreover, each potential adolescent mother would earn around USD 1,046 less each year in the labor market during adulthood due to lower education attainment and labor force participation (relative to women who became mothers later in life) (UNFPA, 2020). This last cost translates into USD 26.8 million annually, considering our estimated number of averted births. These benefits should be included in a full accounting of the program's costs and benefits, including the effects on college enrollment outcomes documented in Londoño-Vélez et al. (2020b).

CRediT authorship contribution statement

Michael D. Bloem: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Jesús Villero:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Data availability

Data will be made available on request.

Appendix A. Appendix tables and figures

See Tables A.1–A.6 and Figs. A.1–A.15.

Appendix 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

Table A.1
Municipality characteristics.

	Comparison municipality mean (1)	Treatment municipality difference (2)
Ln(population) (2009)	9.298	0.449***
	(0.042)	(0.066)
Rural share of population (2009)	0.620	−0.085***
	(0.009)	(0.014)
Distance to department's capital (km)	88.106	−20.000***
	(2.633)	(3.369)
Distance to nearest SPP eligible institution (km)	111.265	−26.132***
	(4.697)	(5.965)
Poverty incidence (2005)	0.544	−0.064***
	(0.005)	(0.006)
Gini coefficient (0–1) (2005)	0.461	−0.015***
	(0.002)	(0.002)
Public expenditure per capita (2009)	1,618.517	−127.155**
	(42.956)	(58.684)
Public revenue per capita (2009)	1,540.250	−155.904***
	(40.232)	(51.988)
Tax revenue per capita (2009)	145.050	66.401***
	(6.315)	(12.240)
Public investment in education per capita (2009)	436.061	−116.908***
	(273.679)	(308.257)
Gross enrollment rate 6th–9th grade (2011)	0.971	0.096***
	(0.012)	(0.016)
Gross enrollment rate 10th–11th grade (2011)	0.645	0.141***
	(0.011)	(0.015)
Dropout rate 6th–9th grade, public schools (2011)	0.050	0.002***
	(0.001)	(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)	0.136	−0.038**
	(0.015)	(0.019)
Number of municipalities	553	553

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 2023 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 A.2
Summary difference-in-differences estimates of the effect of *Ser Pilo Paga* on teen fatherhood (Linear model)

	Births per 1,000 men			
Age group:	All teens (15–19)	15–17	18	19
	(1)	(2)	(3)	(4)
$SPP \times Post$	−1.318***	−0.496*	−3.106***	−2.118***
	(0.411)	(0.265)	(0.842)	(0.819)
	[-0.070***]	[-0.059**]	[-0.072***]	[-0.073***]
	(0.020)	(0.030)	(0.020)	(0.020)
Observations	12,166	12,166	12,166	12,166
Treatment municipalities	553	553	553	553
Comparison municipalities	553	553	553	553
Pre-trends testing p-value	0.381	0.469	0.584	0.272
Pre-SPP share of teen fathers	100	25.9	33.6	40.5

Notes: This table presents summary difference-in-differences estimates from a linear model version of Eq. (7) estimated by Ordinary Least Squares (OLS). The outcome is the number of births per 1,000 men in each age group. 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. The implied average proportional treatment effects on the treated are shown in brackets. The pre-trends test p-values in all columns come from a joint test of the pre-period placebo effects conducted using only pre-SPP observations. 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 (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

treatment measure. We use individual-level test scores on the SABER 11 exam for female students between 2010 and 2016.[31] We use a triple difference empirical approach that leverages the same municipality-level variation in SPP eligibility as in our fertility analysis (see Eq. (8))

[31] 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. Following Londoño-Vélez et al. (2020b), our sample includes test-takers aged 14–23 to maximize their probability of being high school seniors.

Table A.3
Robustness to Alternative Specifications and Sample of Municipalities.

	Fertility rate			
	(1)	(2)	(3)	(4)
$Teen \times SPP \times Post$	−0.057***	−0.056***	−3.243***	−3.061***
	(0.010)	(0.012)	(0.684)	(0.753)
			[-0.053***]	[-0.047***]
			(0.011)	(0.011)
Model	Exponential	Exponential	Linear	Linear
Estimation	Poisson	Poisson	OLS	OLS
Exclude capital cities	No	Yes	No	Yes
Observations	24,332	23,628	24,332	23,628
Treatment municipalities	553	526	553	526
Comparison municipalities	553	548	553	548
Pre-trends test p-value	0.988	0.953	0.954	0.802

Notes: This table reports results from variations of our main specification and sample of municipalities. Column 1 replicates our main results from Eq. (8). In column 2, we exclude the group of 32 municipalities corresponding to the capitals of departments. For columns 1 and 2, reported estimates correspond to proportional reductions calculated as $\exp(\hat{\beta}) - 1$. Column 3 uses a linear model version of Eq. (8) estimated by Ordinary Least Squares (OLS). For columns 3 and 4, the implied average proportional treatment effects on the treated are shown in brackets. The pre-trends test p-values in all columns come from a joint test of the pre-period placebo effects conducted using only pre-SPP observations. The Poisson models are estimated using the number of births as the outcome and the annual population of women in each municipality and age group as the exposure. For columns 3–4, 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 A.4
Robustness to Alternative Treatment Definitions.

	Fertility rate			Births
	(1)	(2)	(3)	(4)
Triple difference coefficient	−0.057***	−0.053***	−0.092***	−0.119***
	(0.010)	(0.010)	(0.012)	(0.029)
Across-municipality treatment	Above median SPP eligibility rate	Above median SPP recipient rate	Below median distance SPP-eligible HEI	Above median SPP eligibility rate
Within-municipality comparison	25–29	25–29	25–29	15–19, < 6th grade
Sample (age groups)	15–19 & 25–29	15–19 & 25–29	15–19 & 25–29	15–19
Observations	24,332	24,332	24,332	24,332
Treatment municipalities	553	554	561	553
Comparison municipalities	553	552	545	553
Pre-trends test p-value	0.988	0.187	0.022	0.228

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 a recipient rate calculated with the number of actual SPP recipients (of any gender) from each municipality expressed per 1,000 SABER 11 test takers in 2014. Treatment municipalities are those below the median recipient rate, while comparison municipalities are those above the median. In column 3, 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 4, 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 sixth grade. For columns 1-3, the Poisson models are estimated using the number of births as the outcome and the annual population of women in each municipality and age group as the exposure. In column 4, 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 a sixth grade level of education over time. Standard errors are clustered at the municipality level and are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Table A.5
Robustness to placebo treatment years.

	Fertility rate				
Placebo year (h):	2009	2010	2011	2012	2013
	(1)	(2)	(3)	(4)	(5)
$Teen \times SPP \times Post$	0.002	0.004	0.001	0.003	0.001
	(0.012)	(0.010)	(0.009)	(0.010)	(0.012)
Observations	13,272	13,272	13,272	13,272	13,272
Treatment municipalities	553	553	553	553	553
Comparison municipalities	553	553	553	553	553

Notes: Each column in this table assumes that SPP was introduced in year h instead of 2014 and estimates Eq. (8) with $Post_t = \mathbb{1}[t > h]$ using only pre-SPP observations. All regressions are estimated using the number of births as the outcome and the annual population of women in each municipality and age group as the exposure. Standard errors are clustered at the municipality level and are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Table A.6
Recent migration patterns across municipalities.

	Resided in a different municipality				
Women:	All 15 and older (1)	15–19 (2)	20–24 (3)	25–29 (4)	30-34 (5)
Panel A. Five years ago (relative to 2018)					
$SPP(\times 100)$	0.058 (0.725)	0.789 (0.535)	0.016 (0.727)	−0.320 (0.959)	0.001 (1.059)
Observations	17,448,967	1,846,601	1,924,137	1,827,670	1,676,189
Treatment municipalities	553	553	553	553	553
Comparison municipalities	553	553	553	553	553
Mean comparison municipalities (%)	9.102	9.565	12.893	13.859	12.619
Panel B. One year ago (relative to 2018)					
$SPP(\times 100)$	−0.135 (0.201)	−0.072 (0.174)	−0.440 (0.277)	−0.130 (0.304)	−0.046 (0.288)
Observations	17,448,967	1,846,601	1,924,137	1,827,670	1,676,189
Treatment municipalities	553	553	553	553	553
Comparison municipalities	553	553	553	553	553
Mean comparison municipalities (%)	2.980	3.597	5.044	4.736	3.894

Notes: This table presents differential migration patterns between treatment and comparison municipalities using cross-sectional information from the full-count 2018 Colombian census. Treatment municipalities are above the median in female eligibility rates for SPP in 2014, while comparison municipalities are below the median. We report coefficients on a treatment indicator representing the mean difference between treatment and comparison municipalities. In Panel A, the outcome variable is an indicator of having resided in a municipality (or country) different from the current one five years ago. In Panel B, the outcome variable is an indicator of having resided in a municipality (or country) different from the current one 12 months ago. In line with our main specification, all regressions include department fixed effects. Standard errors are clustered at the municipality level and presented in parentheses. (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

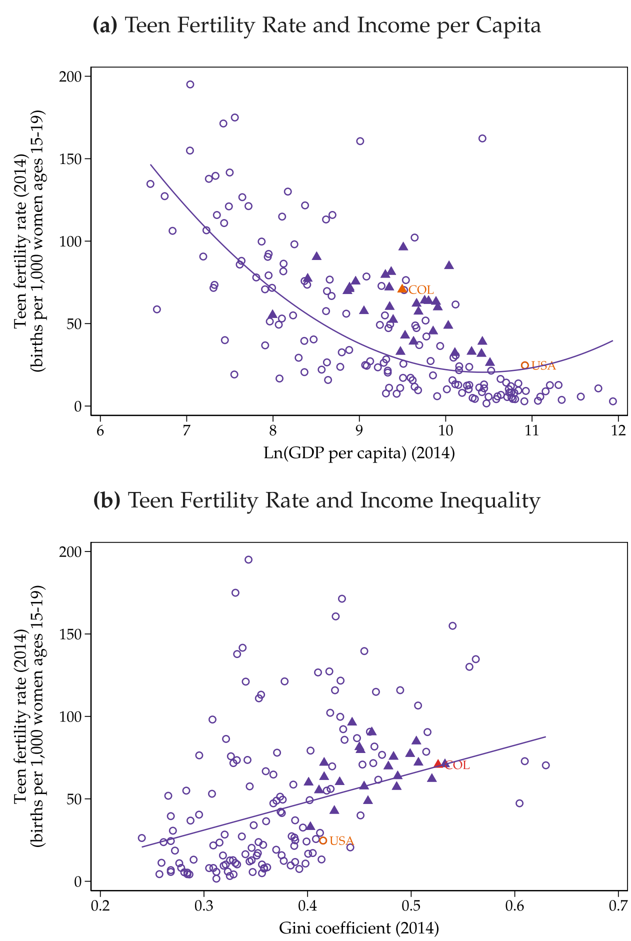
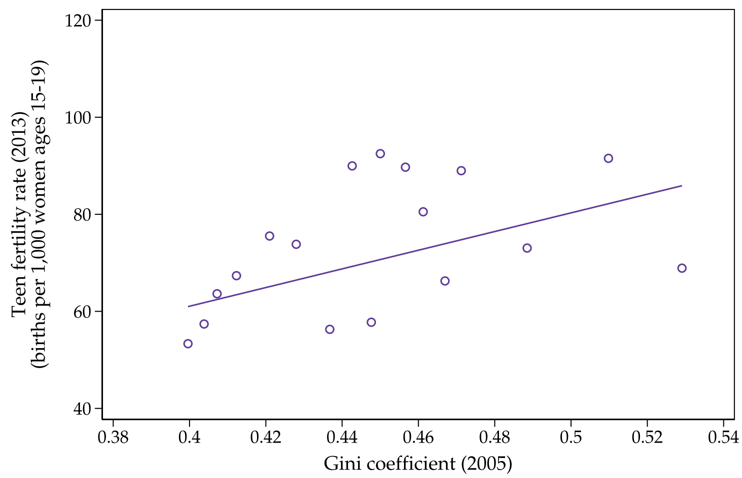


Fig. A.1. Correlates of Teen Fertility Rates Across Countries.
Notes: This figure shows some correlates of teen fertility across countries. Panel (a) shows the relationship between teen fertility rates and GDP per capita (PPP) in 2014 for a sample of 183 countries. Panel (b) shows the relationship between 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 orange, 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).

(a) Teen Fertility Rate and Income Inequality Across Municipalities



(b) Teen Fertility Rate and College Enrollment Across Departments

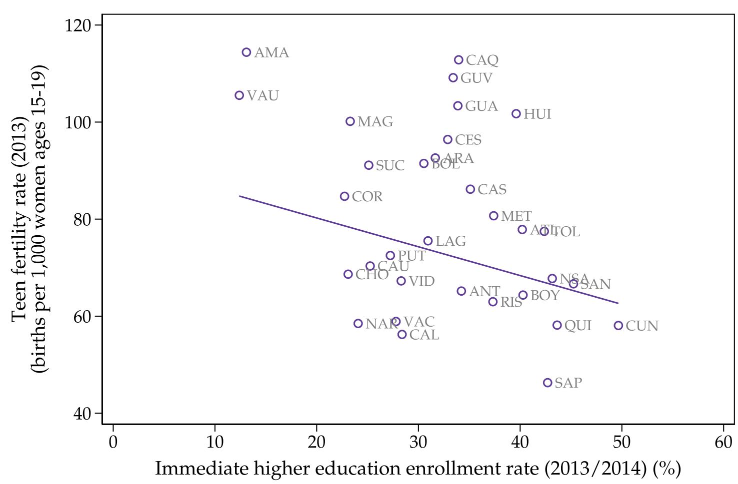


Fig. A.2. Correlates of Teen fertility Rates in Colombia.
Notes: This figure shows some correlates of teen fertility across municipalities and departments in Colombia. Departments in Colombia are similar to states in the United States. A group of municipalities forms each department. Panel (a) shows 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 192.392 (s.e. = 43.725). 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 −0.593 (s.e. = 0.290). 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.

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 ordinary least squares (OLS):

$$StdTestScore_{it} = \phi\left(SISBEN_i^{1\text{-}2} \times SPP^*_{m(i)} \times Post_t\right) + X_{it}\Gamma_t + \psi_{s(i)m(i)} + \psi_{m(i)t} + \psi_{s(i)d(i)t} + v_{it}, \tag{B.1}$$

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.[32] $SISBEN_i^{1\text{-}2}$ indicates whether the student is categorized as SISBEN

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

$$TestScore_i = \frac{Chem_i + Bio_i + Phys_i + 2SocSci_i + Philo_i + 3Lang_i + 3Math_i + Eng_i}{13}$$

for 2010–2013, and

$$TestScore_i = 5 \times \left(\frac{3Math_i + 3Reading_i + 3NatSci_i + 3SocSci_i + Eng_i}{13}\right)$$

for 2014–2016.

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.

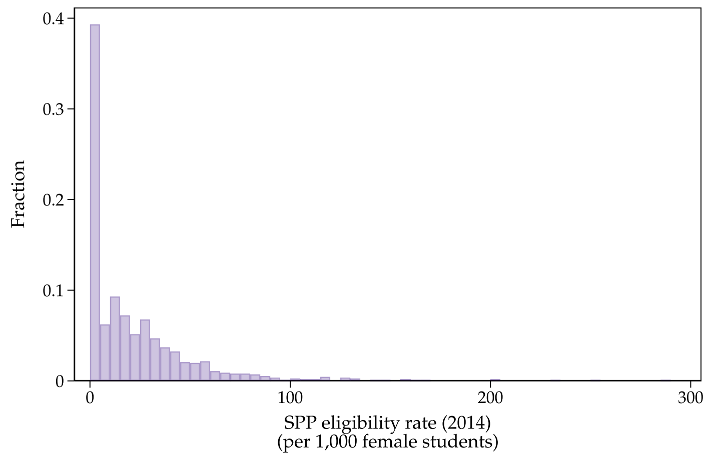


Fig. A.3. Distribution of SPP Eligibility Rates.
Notes: This histogram shows the distribution of SPP eligibility rates in 2014 for the 1106 municipalities in our main sample. For the overall sample of municipalities, the minimum value is 0 and the maximum value is 285.7 (eligible female students per 1,000 female students). The median eligibility rate is 12.6. The 25th and 75th percentiles are 0 and 30.8, respectively. These are unweighted percentiles.

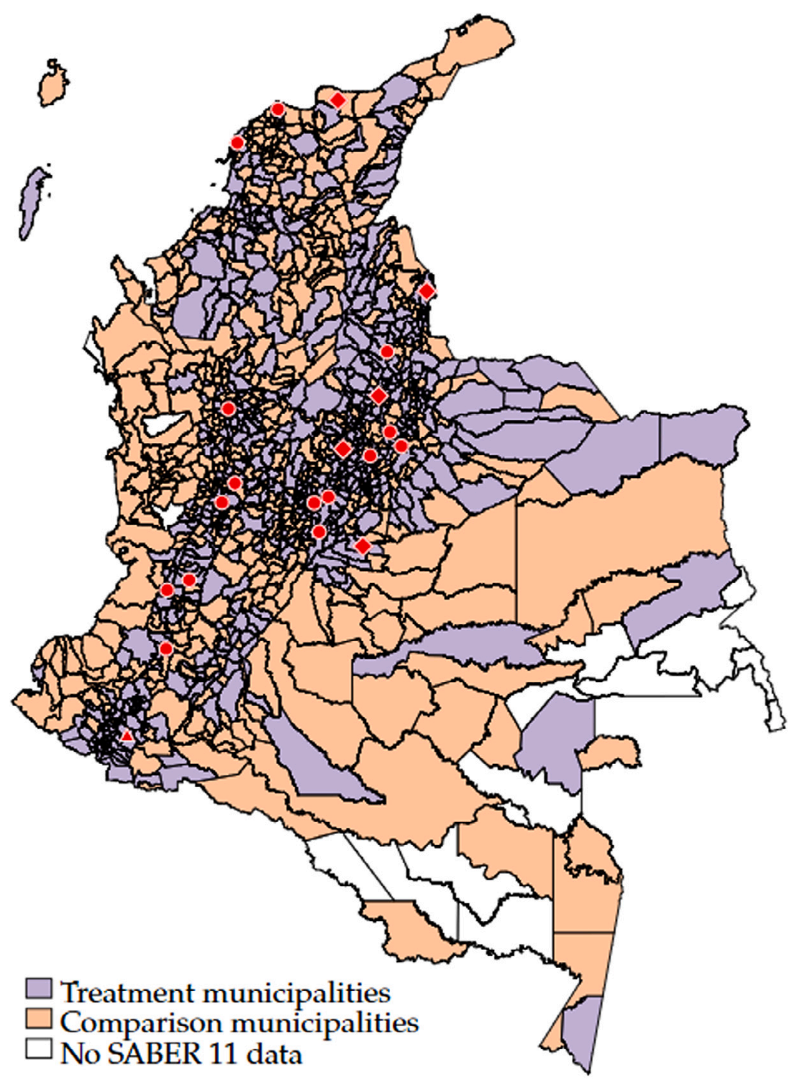


Fig. A.4. Municipality-Level Variation in SPP Eligibility.
Notes: This map displays the geographic distribution of treatment and comparison municipalities. Treatment municipalities are above the median in female eligibility rates for SPP in 2014, while comparison municipalities are below the median. Red markers indicate municipalities with at least one SPP-eligible HEI (circle: from SPP 1; diamond: added in SPP 3; triangle: added in SPP 4).

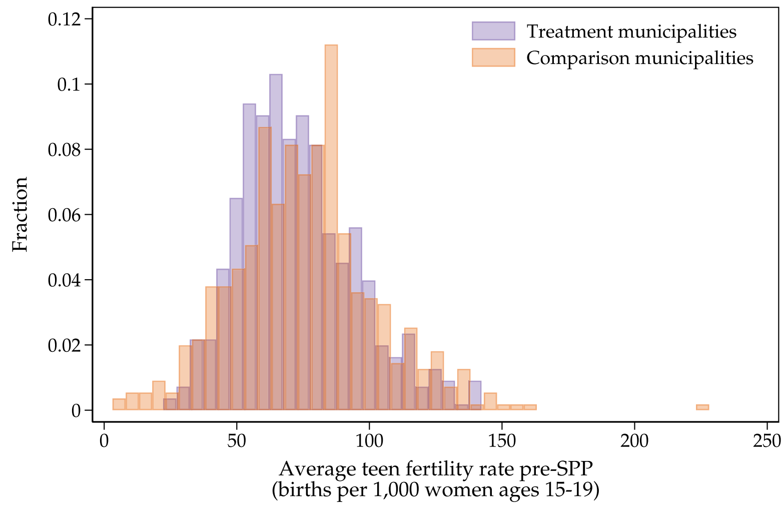


Fig. A.5. Distribution of Teen Fertility Rates Pre-SPP.
Notes: This histogram shows the distribution of average teen fertility rates in the pre-SPP period for the 1106 municipalities in our main sample separately for treatment and comparison municipalities. The pre-period includes relative years from 2008Q4-2009Q3 to 2013Q4-2014Q3. Treatment municipalities are above the median in female eligibility rates for SPP in 2014, while comparison municipalities are below the median.

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.[33] SPP_m^* denotes treatment and comparison municipalities and is defined as $SPP_m^* = \mathbb{1}\left[SPP_m > \text{median}\left(SPP_m\right)\right]$ with SPP_m being the rate of female students eligible for the program in a given municipality in 2014. Eq. (B.1) 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 age, whether the student is enrolled in a public school, whether the student is enrolled in a rural school, the student's school schedule, the parents' education levels, and the family size. 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 ψ, between fixed effects for SISBEN levels ($s \in \{1\text{–}2, \text{Other}\}$), municipalities, and years. Similar to our main fertility specification (see Eq. (8)), we allow the SISBEN-specific year effects to vary by region (*i.e.*, department). Finally, v_{it} is an error term. In Eq. (B.1), ϕ is our parameter of interest, measuring the effect of SPP on test scores.

Table B.1 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 0.028 standard deviations relative to comparison municipalities (p-value = 0.047). This increase represents about 4 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.

Like for our main fertility results (see Section 5), in Table B.1 and Fig. B.1, 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 (0.017 SD or 2.4 percent, not statistically significant) to the fourth quartile (0.064 SD or 9 percent, p-value = 0.008).

[33] See Section 3 for more information on SISBEN levels.

Appendix 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 (2023) 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, deforestation, and entrepreneurial activity (Prem et al., 2023b, 2020; Bernal et al., 2024). While Guerra-Cújar et al. (2023) 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 C.1. The triple difference estimate for this subset of municipalities is −0.056, 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 C.1. Our estimate for this subset of municipalities is −0.085, 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.

Finally, since 2015, the Colombian Ministry of Education has been gradually implementing an initiative to transition public schools from

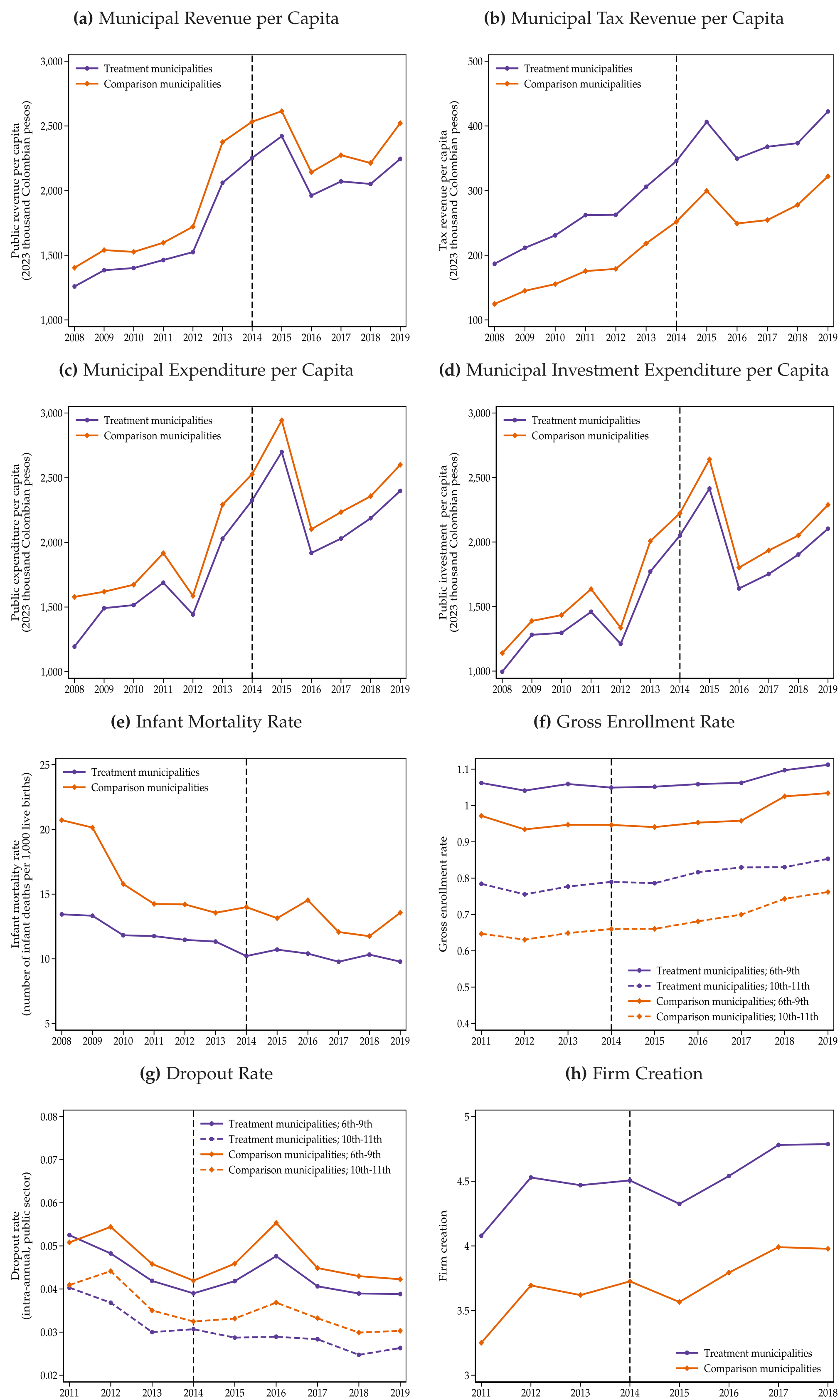


Fig. A.6. Evolution of Other Municipality Characteristics.
Notes: This figure presents the evolution of municipality characteristics between treatment and comparison municipalities during our period of analysis.

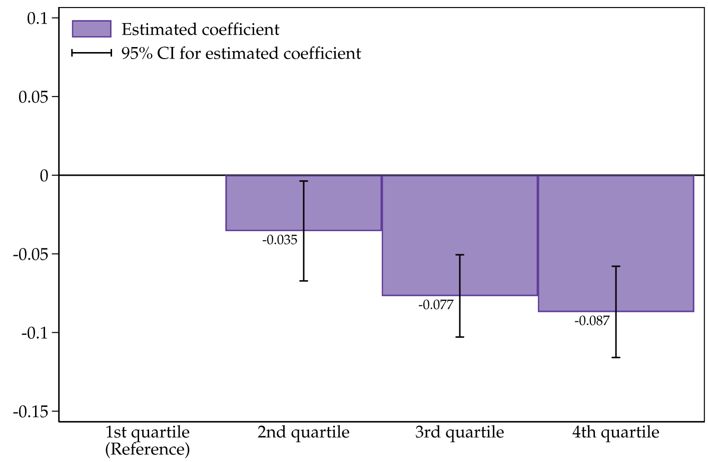


Fig. A.7. Summary Triple Difference Estimates by Quartile of Initial SPP Eligibility.
Notes: This figure plots summary triple difference estimates of $\exp(\beta_\tau) - 1$ from Eq. (8) using indicators for the quartile of SPP eligibility rates instead of an indicator for being above the median in SPP eligibility. The student-weighted mean eligibility rates (per 1,000 female students) for the 1st through 4th quartile are 0, 9.128, 20.156, and 47.550, respectively. All estimates use the number of births as the outcome and the population of women as the exposure variable. Standard errors are clustered at the municipality level.

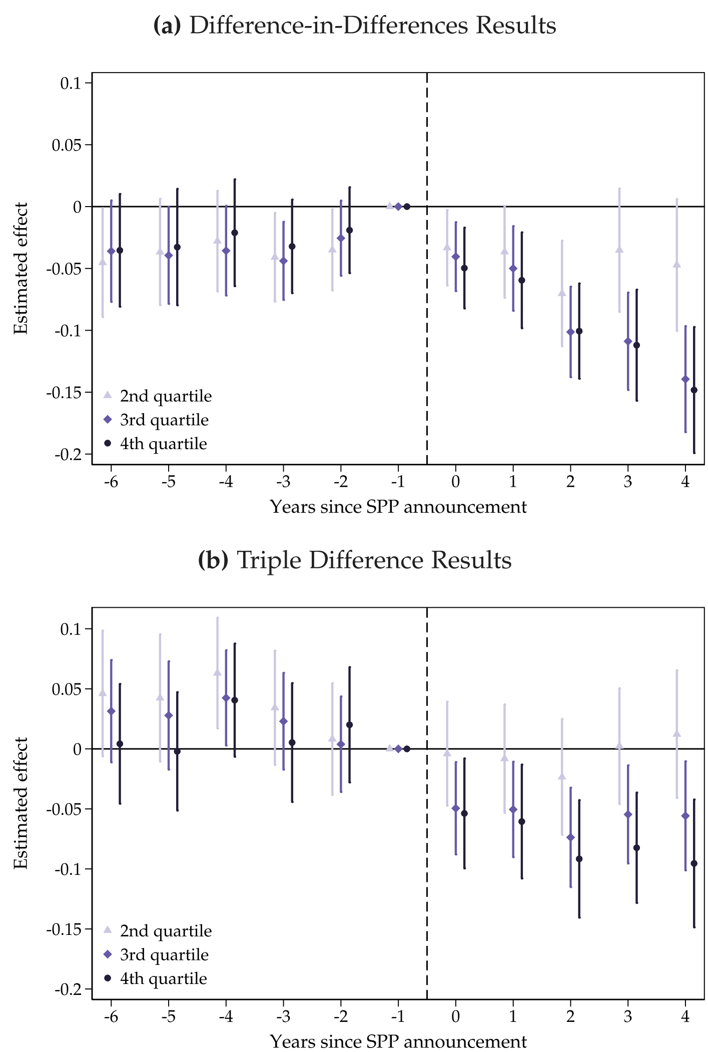


Fig. A.8. Event Study Estimates by Quartile of Initial SPP Eligibility.
Notes: This figure plots the difference-in-differences estimates of $\exp(\alpha_\tau) - 1$ from Eq. (5) and the triple difference event study estimates of $\exp(\beta_\tau) - 1$ from Eq. (6) using indicators for the quartile of SPP eligibility rates instead of an indicator for being above the median in SPP eligibility. All estimates use the number of births as the outcome and the population of women as the exposure variable. Standard errors are clustered at the municipality level.

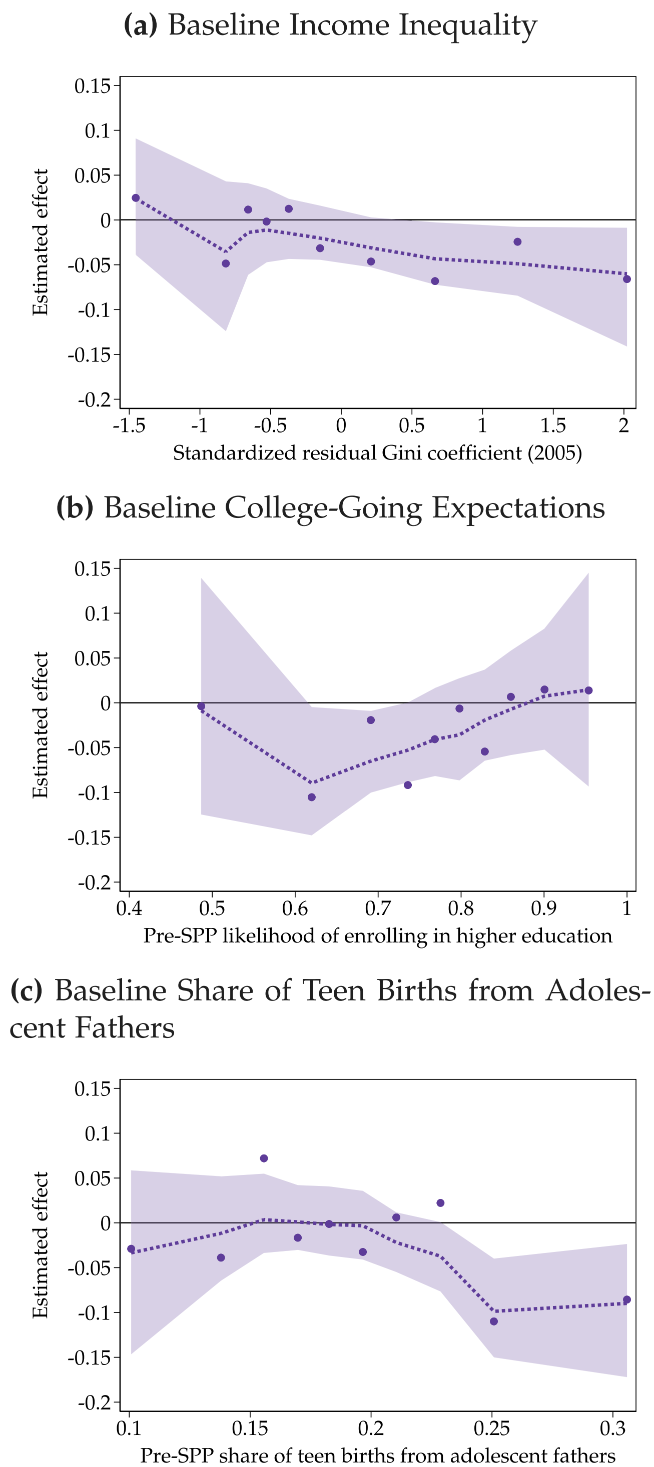


Fig. A.9. Triple Difference Estimates by Municipality Characteristics (Non-parametric Analysis).
Notes: This figure presents triple difference estimates by municipality characteristics indicated in each panel. To do this, we first divide municipalities in deciles according to the level of each characteristic. We then use Eq. (8) to estimate the triple difference proportional reduction in fertility rates for each decile. Finally, we plot these estimates (dots) together with a lowess-smoothed line of the triple difference coefficients (dashed line) on the average level of the characteristic for the deciles. The shaded area represents 95% confidence interval for the line based on 10,000 bootstrap replications of this procedure.

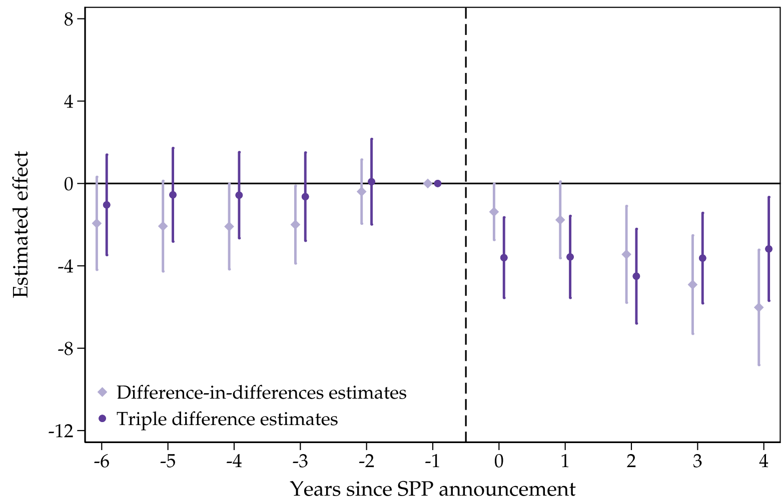


Fig. A.10. Event Study Estimates of the Effect of *Ser Pilo Paga* on Teen Fertility Rates (Linear Model).
Notes: This figure plots the difference-in-differences event study estimates from a linear model version of Eq. (5) and the triple difference event study estimates from a linear model version of Eq. (6). 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.

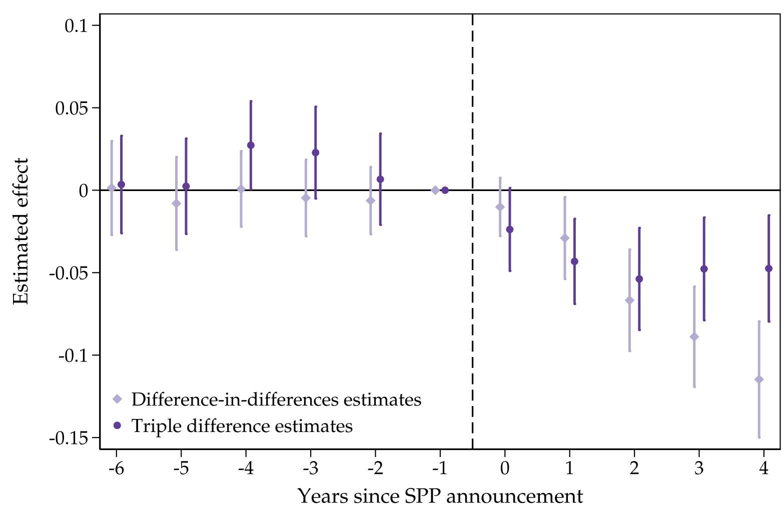


Fig. A.11. Event Study Estimates of the Effect of *Ser Pilo Paga* on Teen Fertility Rates Using Recipients in 2014.
Notes: This figure plots the difference-in-differences estimates of $\exp(\alpha_\tau) - 1$ from Eq. (5) and the triple difference event study estimates of $\exp(\beta_\tau) - 1$ from Eq. (6). Here treatment municipalities are above the median in overall recipient rates for SPP in 2014, while comparison municipalities are below the median. The dots and diamonds represent the estimated proportional effects and the vertical lines represent 95 percent confidence intervals. All estimates use the number of births as the outcome and the population of women as the exposure variable. Standard errors are clustered at the municipality level.

half-day shifts (morning and afternoon) to full-school days to extend the duration and quality of instruction. This policy is called *Jornada Única* or "Full School Day."[34] 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 to explain the sharp decline in teen fertility observed right after the introduction of SPP in 2014. Column 4 of Table C.1 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 (−5.5 percent).

[34] See Hincapie (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.

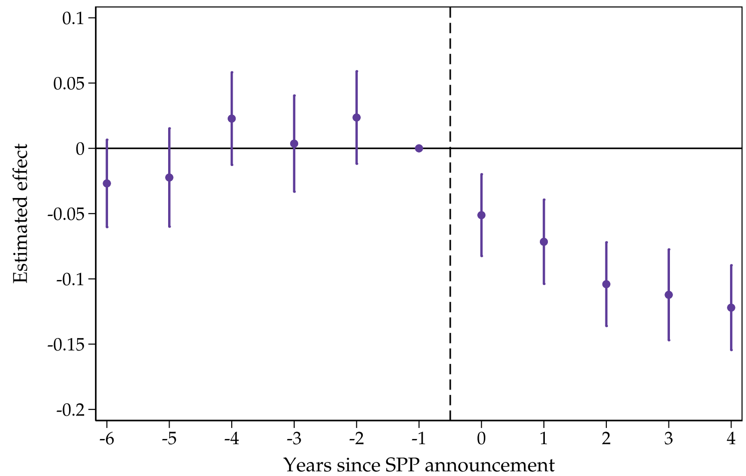


Fig. A.12. Event Study Estimates Exploiting Distance to Nearest SPP-Eligible Institution.
Notes: This figure plots the triple difference event study estimates from Eq. (6), where SPP^*_m is replaced with an indicator for whether a municipality is below the median distance to the nearest SPP-eligible institution. The dots represent the estimated proportional effects and the vertical lines represent 95 percent confidence intervals. The estimates use the number of births as the outcome and the population of women as the exposure variable. Standard errors are clustered at the municipality level.

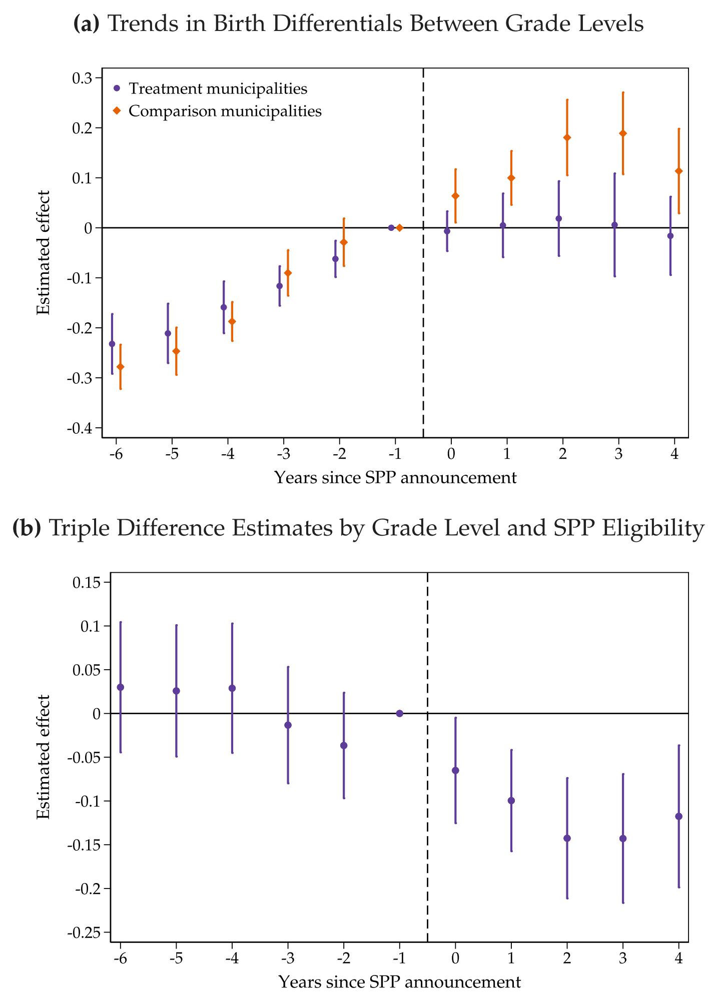


Fig. A.13. Event Study Estimates Exploiting Mothers' Grade Level.
Notes: This figure plots event study estimates from Eq. (6), where $Teen_a$ is replaced with an indicator for mothers who have completed sixth grade or higher. Mothers aged 15–19 with less than a sixth grade education is used as the comparison group instead of women aged 25–29. We use the number of births as the outcome instead of a birth rate because we cannot reliably calculate the number of young women in each municipality with above or below a sixth grade level of education over time. 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 proportional effects and the vertical lines represent 95 percent confidence intervals. Standard errors are clustered at the municipality level.

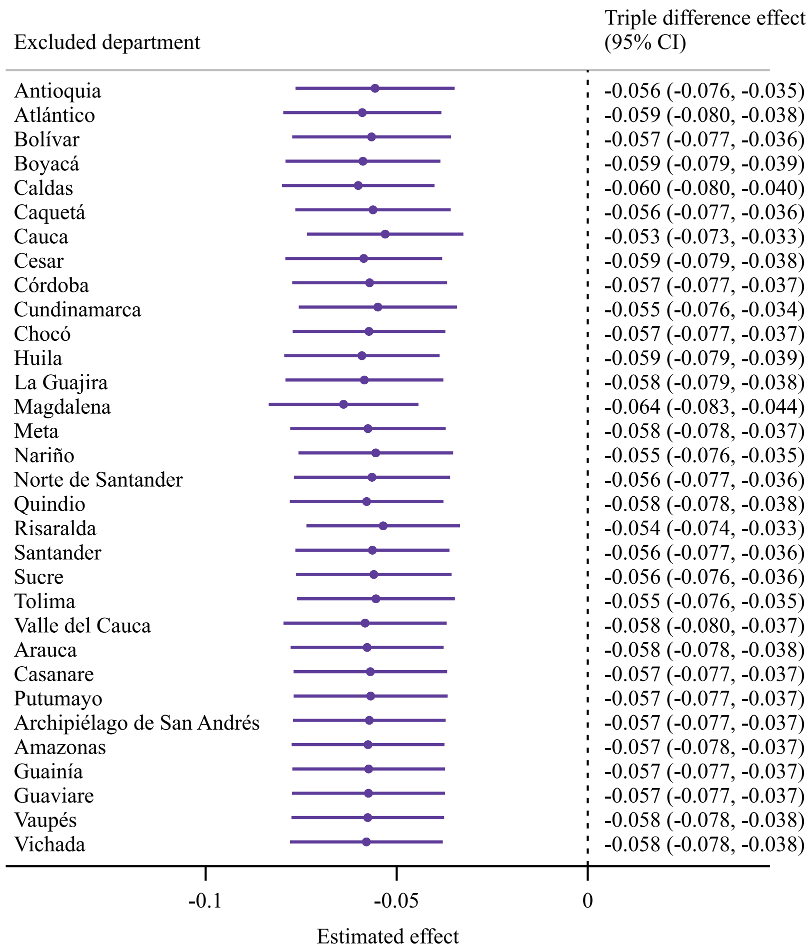


Fig. A.14. Robustness to Excluding All Municipalities in a Given Department.
Notes: This figure reports triple difference estimates of $\exp(\beta_\tau) - 1$ from Eq. (8) where a single department is excluded in each regression. Dots represent estimated proportional effects and horizontal lines represent 95 percent confidence intervals. All estimates use the number of births as the outcome and the population of women as the exposure variable. Standard errors are clustered at the municipality level.

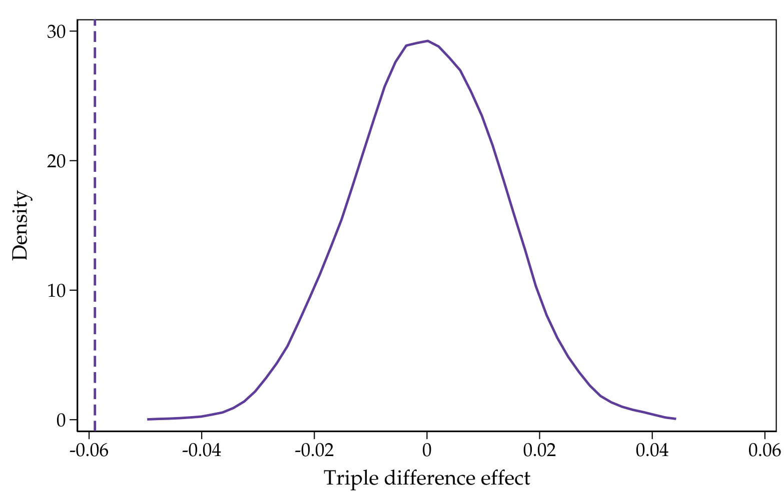


Fig. A.15. Randomization Inference (RI): Distribution of Placebo Treatment Effects.
Notes: This figure presents the distribution of placebo treatment effects after 5,000 random permutations of the treatment assignment (*i.e.*, we randomize municipalities to be treatment or comparison municipalities). We run the regressions using our summary specification in Eq. (8). The vertical dashed line represents the original estimated proportional effect. RI-based p-value = 0.000. The procedure was implemented using the routine by Heß (2017).

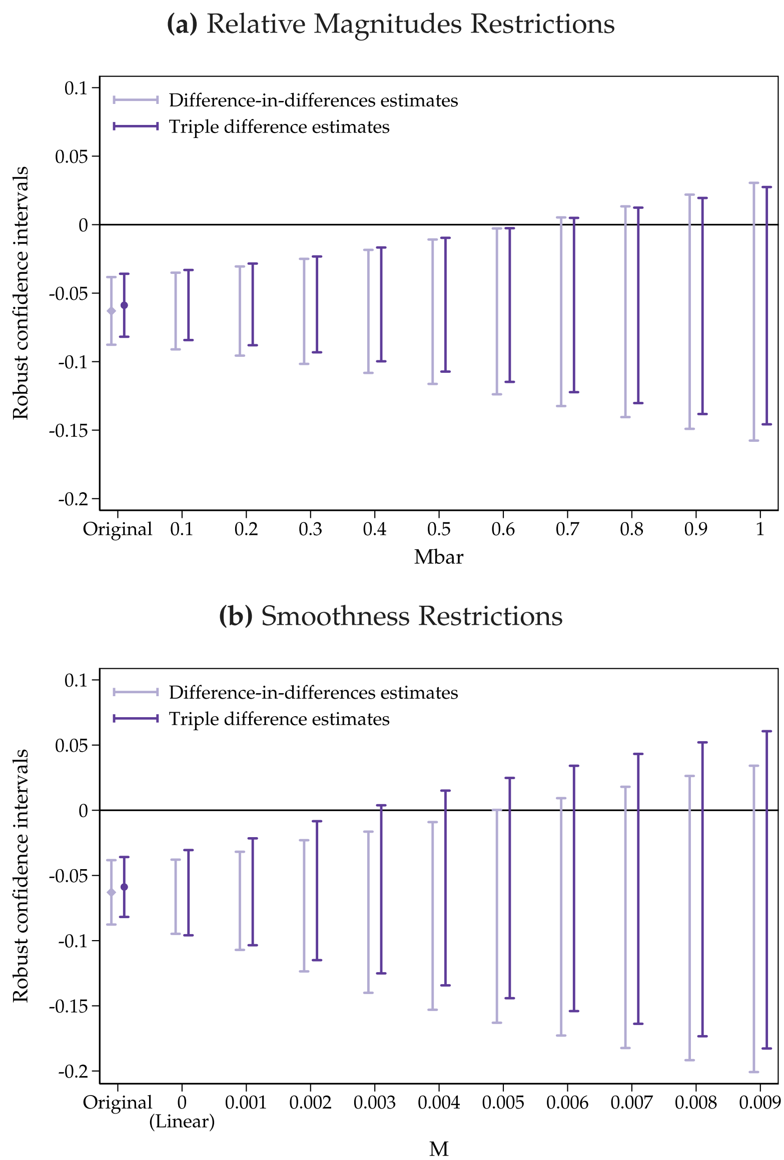


Fig. A.16. Sensitivity Analysis to Violations of Parallel Trends.
Notes: This figure presents sensitivity analyses for the average proportional treatment effect on the treated to potential violations of our parallel trends assumption using the methods developed by Rambachan and Roth (2023). The figure shows 95 percent robust confidence intervals for the $\%ATT$ under different restrictions. The $ATT\%$ is estimated using Eqs. (5) and (6), and calculated as the average of all the event study coefficients for the post-SPP periods. Panel (a) presents sensitivity to allowing the violation of parallel trends across all consecutive post-SPP years to be as large as $\overline{M}$ times the maximum violation observed in the pre-SPP period. Panel (b) presents sensitivity to allowing for a linear extrapolation of the trends observed in the pre-SPP years to the post-period and to deviations from the slope of the linear pre-trend by M across all consecutive post-SPP years.

Table B.1
Triple Difference Estimates on SABER 11 Test Scores.

	Standardized SABER 11 test score	
	(1)	(2)
$SISBEN^{1\text{-}2} \times SPP \times Post$	0.028**	
	(0.014)	
$SISBEN^{1\text{-}2} \times SPPQuartile \times Post$		
1st quartile		[Reference]
2nd quartile		0.017
		(0.024)
3rd quartile		0.037*
		(0.021)
4th quartile		0.064***
		(0.024)
Observations	1,953,546	1,953,546
Treatment municipalities	553	–
Comparison municipalities	553	–
Pre-trends test p-value	0.110	–
Pre-SPP socioeconomic achievement gap	0.713	–

Notes: The table above reports triple difference estimates of ϕ from Eq. (B.1). 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. The student-weighted mean eligibility rates (per 1,000 female students) for the 1st through 4th quartile are 0, 9.128, 20.156, and 47.550, respectively. The pre-trends p-value in column 1 comes from a test of pre-period trends obtained by estimating a dynamic version of Eq. (B.1) using only pre-SPP observations and running a joint test of these coefficients. Standard errors are clustered at the municipality level and are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

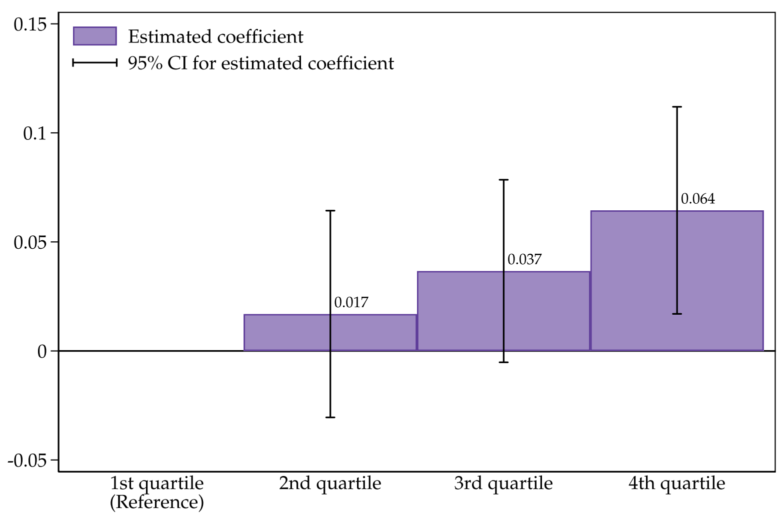


Fig. B.1. Triple Difference Estimates on SABER 11 Test Scores by Quartile of SPP Eligibility.
Notes: This figure plots summary triple difference estimates of ϕ from Eq. (B.1) using indicators for the quartile of SPP eligibility rates instead of an indicator for being above the median in SPP eligibility. The student-weighted mean eligibility rates (per 1,000 female students) for the 1st through 4th quartile are 0, 9.128, 20.156, and 47.550, respectively. Standard errors are clustered at the municipality level.

Table C.1
Robustness to possible confounding events.

	Fertility rate			
Municipalities:	All	No FARC	Low Zika incidence	No *Jornada Única*
	(1)	(2)	(3)	(4)
$Teen \times SPP \times Post$	−0.057***	−0.056***	−0.085***	−0.055***
	(0.010)	(0.011)	(0.017)	(0.015)
Observations	24,332	21,428	11,990	15,246
Treatment municipalities	553	498	263	314
Comparison municipalities	553	476	282	379
Pre-trends testing p-value	0.988	0.935	0.844	0.714

Notes: Column 1 reproduces the preferred triple difference results. Column 2 reports results from our summary specification in Eq. (8) 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 Eq. (8) for municipalities below the median incidence of Zika in 2016. The mean incidence of Zika in these municipalities was 6.3 cases (including both confirmed and probable cases) per 100,000 inhabitants, versus 313.3 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 use the number of births as the outcome and the population of women as the exposure variable. Standard errors are clustered at the municipality level and are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

References

Abadía, Luz Karime, Bernal, Gloria, 2017. A widening gap? A gender-based analysis of performance on the Colombian high school exit examination. Revista de Econ. del Rosario 20 (1), 5–31.

Acevedo, Karina Margarita, Bornacelly, Iván David, 2014. Panel Municipal del CEDE. Documentos CEDE 26, Universidad de los Andes, Facultad de Economía, CEDE.

Aizer, Anna, Devereux, Paul, Salvanes, Kjell, 2022. Grandparents, moms, or dads? Why children of teen mothers do worse in life. J. Hum. Resour. 57 (6), 2012–2047.

Alzúa, María Laura, Velázquez, Cecilia, 2017. The effect of education on teenage fertility: Causal evidence for Argentina. IZA J. Dev. Migr. 7 (1), 1–23.

Attanasio, Orazio, Sosa, Lina Cardona, Medina, Carlos, Meghir, Costas, Posso-Suárez, Christian Manuel, 2021. Long term effects of cash transfer programs in Colombia. Working Paper Series 29056, National Bureau of Economic Research.

Azevedo, Joao Pedro, Favara, Marta, Haddock, Sarah E., López-Calva, Luis F., Muller, Miriam, Perova, Elizaveta, 2012. Teenage Pregnancy and Opportunities in Latin America and the Caribbean: On Teenage Fertility Decisions, Poverty and Economic Achievement. World Bank, Washington, DC.

Bandiera, Oriana, Buehren, Niklas, Burgess, Robin, Goldstein, Markus, Gulesci, Selim, Rasul, Imran, Sulaiman, Munshi, 2020. Women's empowerment in action: Evidence from a randomized control trial in Africa. Am. Econ. J.: Appl. Econ. 12 (1), 210–259.

Becker, Gary S., 1960. An economic analysis of fertility. In: Becker, Gary S. (Ed.), Demographic and Economic Change in Developed Countries. Princeton University Press, Princeton, NJ, pp. 209–231.

Bernal, Gloria L., Penney, Jeffrey, 2019. Scholarships and student effort: Evidence from Colombia's Ser Pilo Paga program. Econ. Educ. Rev. 72, 121–130.

Bernal, Carolina, Prem, Mounu, Vargas, Juan F., Ortiz, Mónica, 2024. Peaceful entry: Entrepreneurship dynamics during Colombia's peace agreement. J. Dev. Econ. 166, 103119.

Berthelon, Matias E., Kruger, Diana I., 2011. Risky behavior among youth: Incapacitation effects of school on adolescent motherhood and crime in Chile. J. Public Econ. 95 (1), 41–53.

Bettinger, Eric P., Long, Bridget Terry, Oreopoulos, Philip, Sanbonmatsu, Lisa, 2012. The role of application assistance and information in college decisions: Results from the H&R block FAFSA experiment. Q. J. Econ. 127 (3), 1205–1242.

Black, Sandra E., Devereux, Paul J., Salvanes, Kjell G., 2008. Staying in the classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births. Econ. J. 118 (530), 1025–1054.

Cáceres-Delpiano, Julio, Giolito, Eugenio, Castillo, Sebastián, 2018. Early impacts of college aid. Econ. Educ. Rev. 63, 154–166.

Chen, Jiafeng, Roth, Jonathan, 2023. Logs with zeros? Some problems and solutions. Q. J. Econ. 139 (2), 891–936.

Correia, Sergio, Guimarães, Paulo, Zylkin, Tom, 2020. Fast Poisson estimation with high-dimensional fixed effects. Stata J. 20 (1), 95–115.

Cowan, Benjamin W., 2011. Forward-thinking teens: The effects of college costs on adolescent risky behavior. Econ. Educ. Rev. 30 (5), 813–825, Special Issue on Education and Health.

Dettling, Lisa J., Kearney, Melissa Schettini, 2023. The cyclicality of births and babies' health, revisited: Evidence from unemployment insurance. Working Paper Series 30937, National Bureau of Economic Research.

Doleac, Jennifer L., Gibbs, Chloe R., 2016. A promising alternative: How making college free affects teens' risky behaviors. Unpublished manuscript.

Duflo, Esther, Dupas, Pascaline, Kremer, Michael, 2015. Education, HIV, and early fertility: Experimental evidence from Kenya. Amer. Econ. Rev. 105 (9), 2757–2797.

Duflo, Esther, Dupas, Pascaline, Kremer, Michael, 2021. The impact of free secondary education: Experimental evidence from Ghana. NBER Working Paper 28937.

Dynarski, Susan, Libassi, CJ, Michelmore, Katherine, Owen, Stephanie, 2021. Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. Amer. Econ. Rev. 111 (6), 1721–1756.

Farin, Sherajum Monira, Hoehn-Velasco, Lauren, Pesko, Michael F., 2023. The Impact of Legal Abortion on Maternal Mortality. Am. Econ. J.: Econ. Policy (Forthcoming).

Ferreyra, María Marta, Avitabile, Ciro, Botero Álvarez, Javier, Haimovich Paz, Francisco, Urzúa, Sergio, 2017. At a Crossroads—Higher Education in Latin America and the Caribbean. Report, Washington, DC: World Bank Group.

Fischer, Stefanie, Royer, Heather, White, Corey, 2018. The impacts of reduced access to abortion and family planning services on abortions, births, and contraceptive purchases. J. Public Econ. 167, 43–68.

Flórez, Carmen Elisa, Soto, Victoria Eugenia, 2019. Tendencias del Embarazo en la Adolescencia y sus Determinantes Próximos en Colombia. In: Vargas Trujillo, Elvia, Flórez, Carmen Elisa, Cortés, Darwin, Ibarra, Marta Carolina (Eds.), Embarazo Temprano: Evidencias de la Investigación en Colombia. Universidad de los Andes, Ediciones Uniandes-Universidad del Rosario, pp. 59–98.

Gamboa, Luis Fernando, Rodríguez-Lesmes, Paul, 2019. The fertility-inhibiting effect of mosquitoes: Socio-economic differences in response to the Zika crisis in Colombia. Econ. Hum. Biol. 35, 63–72.

Gaviria, Alejandro, 2010. Cambio Social en Colombia Durante la Segunda Mitad del Siglo XX. Documentos CEDE 30, Universidad de los Andes, Facultad de Economía, CEDE.

Giacobino, Hélène, Huillery, Elise, Michel, Bastien, Sage, Mathilde, 2023. Schoolgirls, not brides: Education as a shield against child marriage. Am. Econ. J.: Appl. Econ. (Forthcoming).

Guerra-Cújar, María Elvira, Prem, Mounu, Rodríguez-Lesmes, Paul, Vargas, Juan F., 2023. A peace baby boom? Evidence from Colombia's peace agreement. J. Hum. Resour. (Forthcoming).

Heß, Simon, 2017. Randomization inference with stata: A guide and software. Stata J. 17 (3), 630–651.

Hincapie, Diana, 2016. Do longer school days improve student achievement? Evidence from Colombia. IDB Working Paper Series IDB-WP-679, Inter-American Development Bank (IDB), Washington, DC.

Jensen, Robert, 2012. Do labor market opportunities affect Young women's work and family decisions? Experimental evidence from India. Q. J. Econ. 127 (2), 753–792.

Kahn-Lang, Ariella, Lang, Kevin, 2020. The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications. J. Bus. Econom. Statist. 38 (3), 613–620.

Kearney, Melissa S., Levine, Phillip B., 2012. Why is the teen birth rate in the United States so high and why does it matter? J. Econ. Perspect. 26 (2), 141–166.

Kearney, Melissa S., Levine, Phillip B., 2014. Income inequality and early nonmarital childbearing. J. Hum. Resour. 49 (1), 1–31.

Kearney, Melissa S., Levine, Phillip B., 2015. Media influences on social outcomes: The impact of MTV's 16 and pregnant on teen childbearing. Amer. Econ. Rev. 105 (12), 3597–3632.

Kelly, Andrea, Lindo, Jason M., Packham, Analisa, 2020. The power of the IUD: Effects of expanding access to contraception through title X clinics. J. Publ. Econ. 192, 104288.

Koohi, Shiva, 2017. College prospects and risky behavior among Mexican immigrant youth: The effects of in-state tuition policies on schooling and childbearing. Econ. Educ. Rev. 58, 162–174.

Laajaj, Rachid, Moya, Andrés, Sánchez, Fabio, 2022. Equality of opportunity and human capital accumulation: Motivational effect of a nationwide scholarship in Colombia. J. Dev. Econ. 154, 102754.

Lindo, Jason M., Packham, Analisa, 2017. How much can expanding access to long-acting reversible contraceptives reduce teen birth rates? American Economic Journal: Economic Policy 9 (3), 348–376.

Londoño-Vélez, Juliana, Rodríguez, Catherine, Sánchez, Fabio, 2020a. Upstream and Downstream Impacts of College Merit-Based Financial Aid for Low-Income Students: Ser Pilo Paga in Colombia. Replication Package, American Economic Association, Available from https://www.aeaweb.org/journals/dataset?id=10.1257/pol.20180131.

Londoño-Vélez, Juliana, Rodríguez, Catherine, Sánchez, Fabio, 2020b. Upstream and downstream impacts of college merit-based financial aid for low-income students: Ser Pilo Paga in Colombia. Am. Econ. J.: Econ. Policy 12 (2), 193–227.

McCrary, Justin, Royer, Heather, 2011. The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. Amer. Econ. Rev. 101, 158–195.

Muralidharan, Karthik, Prakash, Nishith, 2017. Cycling to school: Increasing secondary school enrollment for girls in India. Am. Econ. J.: Appl. Econ. 9 (3), 321–350.

OECD and World Bank, 2012. Reviews of National Policies for Education—Tertiary Education in Colombia. Report, OECD Publishing, Paris.

Olden, Andreas, Møen, Jarle, 2022. The triple difference estimator. Econom. J. 25 (3), 531–553.

Prem, Mounu, Saavedra, Santiago, Vargas, Juan F., 2020. End-of-conflict deforestation: Evidence from Colombia's peace agreement. World Dev. 129, 104852.

Prem, Mounu, Vargas, Juan F., Namen, Olga, 2023a. The Human Capital Peace Dividend. Replication Package, Inter-university Consortium for Political and Social Research, Ann Arbor, MI, Available from http://dx.doi.org/10.3886/E135341V1.

Prem, Mounu, Vargas, Juan F., Namen, Olga, 2023b. The human capital peace dividend. J. Hum. Resour. 58 (3), 962–1002.

Rambachan, Ashesh, Roth, Jonathan, 2023. A more credible approach to parallel trends. Rev. Econ. Stud. 90 (5), 2555–2591.

Roth, Jonathan, 2022. Pretest with caution: Event-study estimates after testing for parallel trends. Am. Econ. Rev.: Insights 4 (3), 305–322.

UNFPA, 2020. Socioeconomic Consequences of Adolescent Pregnancy in Six Latin American Countries. Implementation of the MILENA Methodology in Argentina, Colombia, Ecuador, Guatemala, Mexico and Paraguay. Report, United Nations Population Fund - Latin America and the Caribbean Regional Office.

Urdinola, B. Piedad, Ospino, Carlos, 2015. Long-term consequences of adolescent fertility: The Colombian case. Demogr. Res. 32 (55), 1487–1518.

Vargas Trujillo, Elvia, Flórez, Carmen Elisa, Cortés, Darwin, Ibarra, Marta Carolina (Eds.), 2019. Embarazo temprano: Evidencias de la investigación en Colombia. In: Embarazo Temprano: Evidencias de la Investigación en Colombia. Universidad de los Andes, Ediciones Uniandes-Universidad del Rosario.

Willis, Robert J., 1973. A new approach to the economic theory of fertility behavior. J. Polit. Econ. 81 (2), S14–S64.

Wooldridge, Jeffrey M., 2023. Simple approaches to nonlinear difference-in-differences with panel data. Econom. J. 26 (3), C31–C66.