

Borradores de ECONOMÍA

Twin instrument, fertility and women's
labor force participation: evidence
from Colombian low-income families

By: Erika Raquel Badillo
Lina Cardona-Sosa
Carlos Medina
Leonardo Fabio Morales
Christian Posso

No. 1071
2019



Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá

Twin instrument, fertility and women's labor force participation: evidence from Colombian low-income families

By ERIKA RAQUEL BADILLO, LINA CARDONA-SOSA, CARLOS MEDINA,
LEONARDO FABIO MORALES, AND CHRISTIAN POSSO *

The opinions contained in this document are the sole responsibility of the authors and do not commit Banco de la República or its Board of Directors

Abstract

This paper uses the twin births instrument to estimate the causal effect of fertility on labor market outcomes for first-time mothers in Colombia. The nature of the data used allows us to overcome previous concerns regarding the validity of this instrument. We can control for observed pre-pregnancy characteristics, and by including fixed effects, we can control for time-invariant unobserved heterogeneity. Results coming from the traditional instrumental variables approach suggest a reduction in female labor supply close to 5.2 percentage points (18%) due to a marginal increase in fertility. Additionally, after following Farbmacher, Guber and Vikström (2018) to eliminate the bias of twins resulting from mothers' characteristics and in vitro fertilizations (e.g., a non-exogenous event), the obtained results are of the same magnitude but less significant than those of the traditional instrument. By making additional assumptions on the correlation of the twin instrument and the error of the structural equation following Nevo and Rosen (2012), we find an upper bound estimate for the reduction in female labor supply close to -5 ppt.

JEL Code: J13, J22, C26.

Keywords: fertility, female labor supply, twin births, instrumental variables.

* Badillo: *Universidad Autónoma Latinoamericana* (erika.badilloen@unaula.edu.co), Cardona: *Institute for Fiscal Studies* (lina_s@ifs.org.uk), Medina: *Banco de la República* (cmedindu@banrep.gov.co), Morales: *Banco de la República* (lmoralzu@banrep.gov.co), Posso: *Banco de la República* (Corresponding author. Medellín, Calle 50 N 50-21. E-mail: cpososu@banrep.gov.co). We are grateful to the DNP for providing us access to the SISBEN datasets. We thank Salome Arango, Carolina Echeverri, Arlen Guarín, Sara Londoño and Laura Prada for their excellent research assistance at different points of the project. Erika Badillo acknowledges financial support from the *Universidad Autónoma Latinoamericana* Research Fund, 26-000004. The opinions expressed herein belong to the authors and do not necessarily reflect the views of *Banco de la República* or its Board of Directors.

Usando nacimientos múltiples para identificar el efecto de la fecundidad en la oferta laboral de mujeres de bajos ingresos en Colombia.

By ERIKA RAQUEL BADILLO, LINA CARDONA-SOSA, CARLOS MEDINA,
LEONARDO FABIO MORALES, AND CHRISTIAN POSSO

Las opiniones contenidas en el presente documento son responsabilidad exclusiva de los autores y no comprometen al Banco de la República ni a su Junta Directiva.

Resumen

En este estudio usamos nacimientos múltiples como instrumento para estimar el efecto causal de la fecundidad en diferentes resultados de mercado laboral para madres primerizas en Colombia. La naturaleza de los datos nos permite superar problemas, sobre la validez de este instrumento, que han sido ampliamente documentados en la literatura. Esto principalmente porque podemos controlar por características observadas de la madre antes del parto, más aún, al incluir efectos fijos por madre podemos controlar por cualquier característica no observada que sea invariante en el tiempo. Nuestras estimaciones de regresiones tradicionales tipo variables instrumentales, sugieren una reducción de 5,2 puntos porcentuales (18%) como resultado de un incremento marginal en la fecundidad. Adicionalmente, después de seguir la metodología de Farbmacher, Guber and Vikström (2018) con el fin de eliminar el sesgo que resulta de diferencias inobservadas de las madres antes del parto y de posibles fertilizaciones in-vitro, las cuales no son un fenómeno exógeno, los resultados obtenidos se mantienen en magnitud. Finalmente, asumiendo la dirección de la correlación entre el instrumento y el error de la ecuación estructural, usando una metodología tipo Nevo and Rosen (2012), encontramos una cota superior del efecto de la fecundidad en la participación laboral femenina de -5 puntos porcentuales.

JEL Code: J13, J22, C26.

Keywords: Fecundidad, oferta laboral femenina, nacimientos múltiples, variables instrumentales.

1. Introduction

The relationship between fertility and female labor supply has been a subject of wide interest among researchers (Aguero & Marks, 2008; J. D. Angrist & Evans, 1998; Bronars & Grogger, 1994; Cáceres-Delpiano, 2012; Jacobsen, Wishart, & Rosenbloom, 1999). Nevertheless, the identification of the causal effect of fertility on female labor supply is challenging for two reasons. The first challenge is due to the simultaneity bias inherent in the choices of whether or not to participate in the labor market and whether or not to have a child. The second reason is related to the difficulty of finding truly valid instruments (Browning, 1992; Clarke, 2018; Goldin, 1990; Nakamura & Nakamura, 1992; Shultz, 1981).

A standard approach widely used in the literature is using twin births as an instrumental variable that deals with the endogeneity of the fertility decision. This is because a twin birth is an event that is considered to be exogenous and unexpected in most cases (Rosenzweig and Wolpin, 2000). Thus, twin births are used as an instrument for the number of children in the household (or family size). Most of the previous studies following this approach have made use of cross-section data. One of the first studies is by Rosenzweig and Wolpin (1980). Since then, this approach has been employed in several contexts, for instance, to examine the effect of fertility on children's outcomes (Angrist & Lavy, 2010; Black, 2005; Cáceres-Delpiano, 2012; Clarke, 2016; Fitzsimons & Malde, 2014; Li, Zhang, & Zhu, 2008; Marteleto & de Souza, 2012; Ponczek & Portela, 2011; Sanhueza, 2009) and on parental outcomes (Aguero & Marks, 2008; J. D. Angrist & Evans, 1998; Cáceres-Delpiano, 2008; Jacobsen et al., 1999). On average, the results suggest a negative effect of fertility on labor market outcomes.

Other studies have used alternative sources of exogenous variation beyond multiple births. For instance, Angrist and Evans (1998) use both mixed gender sibling composition and twin births as instruments for woman's fertility. Hotz and Sanders (2005) use an indicator for women suffering miscarriage during their first pregnancy to estimate the effect of delaying childbearing on annual hours of work and earnings among teenage mothers. Similarly, Cristia (2008) uses a sample of childless women receiving fertility treatments. In all cases, these instruments are shown to be important drivers of fertility, although in each case the exclusion restriction can be placed into doubt (Clarke, 2018).

Despite all such efforts to disentangle the causal effect of fertility on labor supply, the twin instrument has been shown to be not independent from mothers' observable characteristics, which might also affect their labor market outcomes. The literature on twin births suggests that the presence of twins, especially dizygotic twins, is positively related to mothers' characteristics such as age, marital status, race, ethnicity and health (Bulmer, 1970; Farbmacher et al., 2018; Hall, 2003). In fact, healthier mothers are more likely to have twins (Bhalotra & Clarke, 2016). One of the reasons for this is that a twin pregnancy can be more demanding, and it is related to more health risks for mothers and newborns than is single birth, leading to the selective miscarriage of twins. In such case, instrumental variables (IV) estimates tend to underestimate the effect of fertility on women's labor supply. However, if these characteristics are observable, they can be accounted for.

Likewise, it has been well documented that fertility treatments are associated with a higher probability of having multiple births, in particular an increase in dizygotic twins (Braakmann & Wildman, 2014a; Cáceres-Delpiano & Simonsen, 2012; Fauser, Devroey, & Macklon, 2005; Kulkarni et al., 2013; Vitthala, Gelbaya, Brison, Fitzgerald, & Nardo, 2018). Such selectivity has been impossible to control for due to the lack of data able to differentiate between monozygotic and dizygotic twins. Thus, most of these concerns apply specifically to dizygotic twins since monozygotic twins are considered to be a random event (Farbmacher et al., 2018).

Several alternatives to correct for the bias of the twin instrument have arisen. One of them exploits the fact that monozygotic twins are always same-sex twins, while dizygotic twins (highly related to in vitro fertilizations (IVF), higher income families and older women) can be either same-sex or opposite-sex. By comparing the estimates using the usual twin instrument with the estimates using only same-sex twins it is possible to measure the amount of selectivity carried out in the studies using the traditional twin instrument (Black, Devereux, & Salvanes, 2007; Figlio, Guryan, Karbownik, & Roth, 2014). Farbmacher et al. (2018) propose a new instrument using the information of opposite-sex twins to correct for the remaining endogeneity in the same-sex twin instrument. Another alternative is to allow for a certain degree of correlation between the instrument and the error term. This implies the adoption of a partial identification approach like the one proposed by Nevo and Rosen (2012). This allows an inference to be made by estimating bounds around the effect of interest through their imperfect instrumental variables method.

Little is known about the causal relationship of women's fertility on female labor market outcomes in Colombia. The few studies that have documented such a relationship have used indicators for the age of children in the household as explanatory variables. That is, the case in the work of Amador, Bernal and Peña (2013), who find that having children younger than five years is related to a reduction of 2 percentage points on women's labor market participation probability. Similarly, few studies have considered the marginal effect that an additional child in the household has on female labor supply. Arango, Posada and Charry (2003) estimate the determinants of women's labor supply in Colombia using OLS. The authors find a positive and non-significant effect of the number of children younger than five years on women's participation decision. Such a result could be biased due to the endogeneity of the number of children in the household.

This study estimates the effect of fertility on labor market outcomes for first time mothers in Colombia. It contributes to the literature on the causal relationship between family size and labor market indicators using twin births as an expected fertility shock in a developing country. By restricting the analysis to low-income families, we are less likely to face a highly selected sample of twins' mothers that would be the case when considering high-income families (i.e., families with different individual preferences for delaying childbearing, more labor market experience and higher probability of going through IVF). Moreover, we follow Farbmacher et al. (2018) to correct for the remaining bias present in the usual twin instrument, and we also add to the literature of imperfect instruments developed by Nevo and Rosen (2012) by using data for a developing country.

Our results suggest that an additional child reduces the labor force participation rate of first time mothers by approximately 5 percentage points. Similarly, by comparing the IV results using the traditional twin instrument and the same-sex instrument, we do not observe significant differences, suggesting the absence of selection on unobservables in our data. When using the new instrument proposed by Farbmacher et al. (2018) (i.e., the one that isolates the bias coming from dizygotic twins), we find a similar reduction in labor force participation but significance at the 11% level. We also follow the Nevo and Rosen (2012) approach for imperfect instruments, which supports our previous results. Finally, once we assume different degrees for the correlation between the twin instrument and the error term, the results reinforce the negative relationship between family size and mothers' labor force participation rate. We also find

negative effects on employment probability (between -2 ppt to -5 ppt) and on the probability of being formal (-2 ppt).

Our results reinforce the advantage of our sample. On the one hand, we estimate the effect for low-income mothers from a developing country. They have a lower probability of going through IVF and experience different economic preferences in the labor market; on the other hand, we can control for observable and time-invariant unobservable women's characteristics before pregnancy. This is done by including pre-pregnancy characteristics and individuals' fixed effects.

This paper is organized as follows. In the next section we describe the data used for the analysis. In Section 3 we present the estimation strategy using multiple instruments and report the estimates of the effect of fertility on labor market outcomes using the standard twin instrument. Section 4 presents additional estimates correcting for the selection bias of the twin instruments relaxing of the exogeneity assumptions, and it provides evidence for the effect when using alternative instruments. Section 5 concludes.

2. Dataset

We use data from SISBEN, a Colombian information system containing registry data for low-income households. The information is collected with the aim of identifying potential beneficiaries of government programs such as subsidized healthcare and conditional cash transfer programs.¹ The data identifies all individuals within a family using identification numbers, names and dates of birth, which we then use to identify twin births. It also contains socio-demographic information for all household members such as age, gender, education, marital status, labor force participation, perceived health status, and house amenities, among other data. The data has been collected three times in the last 20 years. We use information from the last two periods, 2005-2007 and 2009-2010, mainly due to the quality of the data and the information available. Our sample consists of women ages 10 to 49 years old without children in 2005-2007 who were first time mothers by 2009-2010. Even though the data does not aim to

¹ The database is used to classify potential beneficiaries of social programs and it assigns a score between 0 and 100 depending on the household's socio-economic characteristics. One of the data versions used in this study classifies households on six different levels depending on the level of economic disadvantage. Thus, level one comprises the most disadvantaged households and level six the least disadvantaged ones. Usually, individuals in the first two levels of SISBEN are the ones eligible for government benefits. In the most recent version, also used in this study, eligibility is determined by a cut-off of the SISBEN score.

be a panel, we are able to merge the information of future mothers who are childless between 2005-2007 with next period's information. The initial period, when these women are childless, is defined as our baseline period while the follow up period corresponds to the next period we have data for (2009-2010). Since fertility might affect women's choices during the first few years after the child is born, our main analysis is conducted for a sample of first time mothers whose children are younger than five years old in the follow up.

The matching rate between baseline and follow up was close to 75%, which provided information for 230,852 first-time mothers across the whole country. One of the reasons why we may not observe some of our initial sample of childless women the next period is due to socio-economic improvement by these women. To account for this, we focus on childless women who belong to the levels 1 and 2 in SISBEN, i.e., those belonging to highly disadvantaged households with a low probability of socio economic improvement during the study window. This leave us with a sample of 210,563 first time mothers from 2009-2010 (with children younger than five years old). Thus, we do not think that the sample of mothers from levels 1 and 2 of SISBEN have unobserved characteristics that make them stay in the database and that could bias our results. The estimate of the fertility effect on labor market outcomes among first time mothers results from comparing labor market indicators between mothers of twins and mothers of singletons.

The main advantage of these data is the possibility of overcoming one of the most recent concerns raised by researchers regarding the use of twins as an instrumental variable. That is, the increase in the correlation between mothers' socio-economic characteristics and the probability of giving birth to twins. As Farbmacher et al. (2018) suggest, twin births are related not only to mothers' health status and age but also to mothers' economic preferences in delaying childbearing and the use of IVF. Our sample overcomes this in several ways: first of all, by restricting the analysis to low-income families, we are reducing the probability that twin births are the result of IVF as it is less widespread in Colombian than in a high-income country². IVF is also an expensive procedure (by 2010, the cost of IVF was equivalent to 20 times the national minimum wage) and was not covered by the subsidized health system during the period of analysis. Second, we use a homogeneous sample of low-educated women who are less likely to

² In general, assisted reproduction is higher in higher-income countries. While in Colombia there are 25 cycles per million people, in the US there are 325 cycles.

delay child bearing and less likely to have strong career preferences. Third, by using the longitudinal component of the data, we are able to control for time-invariant unobserved characteristics such as women's preferences and health conditions. Similarly, by using first time mothers we are able to estimate how two children affect the labor force participation of mothers in contrast to having only one child. By doing so we avoid the endogeneity that could exist in second time mothers. Finally, we are able to control for pre-pregnancy socio-economic characteristics and labor market outcomes before motherhood.

Figure 1 shows the percentage of twins among individuals younger than 15 years old in our data. The observed rate is approximately 1.3%, with a few exceptions experiencing a twin rate above the average. From the figure it is also possible to infer that twin births are distributed across the country, and they are neither systematic nor associated with a region. The proportion of twins' mothers is shown at the bottom of Table 1. For both the whole sample and the subsample of mothers in levels 1 and 2 of SISBEN, the percentage of mothers giving birth to twins is less than 1% (0.6%), and most of them are same-sex twins.

[Insert Figure 1 around here]

Table 1 also reports some pre-pregnancy socio-economic characteristics for the sample of first-time mothers with children younger than five years old used in the study (we do the same for mothers of children younger than three years old in Appendix A). The first two columns present the whole sample of mothers, while the next two columns report the descriptive information for the most disadvantaged sample (i.e., those classified in levels 1 and 2 of SISBEN). When looking at the whole sample we observe that, on average, first time mothers in SISBEN are around 20 years old. This figure and its standard deviation of 5 years are evidence of the early stage at which low-income women start to have children in Colombia; this figure also contrasts with the age of first pregnancies found in countries such as Sweden (26 to 29). Around 60% of the first time mothers in our data obtained secondary education and 3% are heads of their households. Between 7% and 8% are married, and half of them do not have access to any health insurance.

From the set of household characteristics and household amenities, it is possible to see that our data contains very disadvantage households. On average the household size is five, and most

of these households belong to the first three strata.³ A bit more than half of them have sewerage, and one third have access to gas. Similarly, less than half have refrigerators in their homes, and 10% enjoy a washing machine. In terms of the SISBEN score assigned to the households to which the mothers belong, more than half scored 11 or less, which is the traditional cut-off to be a beneficiary of social programs under version II of SISBEN (i.e., the one administered between 2005-2007). In terms of labor market status, future mothers had a labor force participation rate close to 30% before pregnancy. 12% were employed and 4% were working in the formal sector. When restricting the descriptive statistics to mothers in levels 1 and 2 of SISBEN, we observe a set that is slightly more disadvantaged than the whole sample. They are slightly less educated, living a lower strata, and few of them have access to normal household amenities. In terms of labor market outcomes, future mothers have a low participation in the labor market (29%), 10% are employed and only 2% have a formal job.

[Insert Table 1 around here]

It is possible to check whether or not mothers of twins are different in their observable characteristics from mothers of singletons. The unexpected fertility shock that results from twin birth is defined as the treatment. We regress each of the observable characteristics, previously described, on a binary variable that takes the value of 1 for mothers of twins and 0 for mothers of singletons. The coefficients of the treatment variable are reported in Table 2. When using the sample of mothers from all the levels of SISBEN, we see that mothers of twins are older, more likely to be married and more likely to contribute to the health system.⁴ Additionally, mothers of twins are located in the higher levels of house strata, have larger houses (a larger number of rooms) and are more likely to have a washing machine. This suggests that mothers of twins are not only a non-random sample of the population but are better off than mothers of singletons at least at the low end of the income distribution in Colombia. Such a finding is in line with what is found in previous literature: twins are positively related to mothers' age and economic characteristics.

³ Houses in Colombia are classified in six different strata depending on the household amenities and the public goods the household have access to. This classification is used to discriminate public utilities pricing. In particular, prices for households belonging to Stratas 1, 2, and 3 are subsidized by what is charged to households in Stratas 5 and 6.

⁴ Contributory health insurance is related either to employment in the formal sector or being the beneficiary of a formal employer. In both cases, it is better than being uninsured.

Nevertheless, when we restrict the analysis to mothers in levels 1 and 2 of SISBEN, the observed differences between mothers of twins and mothers of singletons are reduced. The only characteristics that remain significantly different are mothers' age and number of rooms. This result supports the idea that our data allows us to reach a less selective sample of twin mothers. In the following sections we check the magnitude of the remaining selectivity bias in the traditional twin instrument framework.

[Insert Table 2 around here]

3. Estimations using the traditional twin instrument

To identify the effect of fertility on female labor market outcomes, we follow the instrumental variables standard approach of using twin births as an exogenous source of variation for family size. In the next section, we also introduce multiple instruments based on the concept of weak instruments (see Nevo & Rosen, 2012).

As previously shown, mothers of twins in the restricted sample of levels 1 and 2 of SISBEN were not found to be systematically different from mothers of singletons during the pre-pregnancy period (with the exception of mother's age and number of rooms, which we can control for). This suggests that the twin shock is uncorrelated with women's initial characteristics. Nevertheless, recent studies, such as the one by Bhalotra and Clarke (2016), find that mothers of twins have higher levels of education and better health and have better medical care. We can account for this by including women's education and type of health insurance at the baseline. Even though we do not observe women's health, we exploit the panel structure of the data to control for health status and any other unobserved variables that are expected to be constant within the time lapse between the baseline and the follow up.⁵

The Instrumental Variables estimation of the effect of family size on labor market outcomes is conducted through 2SLS. The first stage follows the structure of equation (1)

$$FSize_i = \theta Twin_i + X_i' \pi + u_i \quad (1)$$

⁵ Previous literature has shown that health status is highly state-dependent, in the sense that good health practices and health markers are highly determined by individuals' past occurrences (see Morales (2016) for example on obesity). Therefore, health status among young women, like the ones used in the present study, is highly dependent on past occurrences and is not expected to vary substantially in a short period of time as in our study.

where $FSize_i$ is a continuous variable for the number of household members for women i . $Twin_i$ is an indicator variable for having a twin birth versus having a singleton birth. X_i is a vector containing women i 's characteristics, measured before the first child is born (baseline) and reported in Table 2. We include mothers' age to account for her life cycle and to control for the fact that multiple births are related with women's age. u_i is the error term. The parameter θ recovers the effect of having twin births on women's family size. The estimates of the effect of fertility on labor supply follow the specification shown in equation (2) which corresponds to the second stage of a 2SLS procedure using the estimates of family size obtained from equation (1):

$$Y_i = \beta \widehat{Size}_i + X_i' \gamma + \varepsilon_i \quad (2)$$

The β coefficient in equation (2) recovers the effect of an additional child on labor supply for mothers who in the absence of the twin birth (shock) would not have a second child. In other words, it recovers the Local Average Treatment Effect (LATE). ε_i accounts for the disturbance term⁶, while vector X_i contains all the observable pre-pregnancy characteristics available at the baseline.

The 2SLS described previously is equivalent to a Wald estimator when the Instrumental Variable is binary, which is precisely our case. Therefore, the estimator of β in equation (2) converge in probability to

$$Plim(\beta) = \frac{E[y_i | X_i, z_i = twin] - E[y_i | X_i, z_i = no\ twin]}{E[Size_i | X_i, z_i = twin] - E[Size_i | X_i, z_i = no\ twin]} \quad (3)$$

The numerator and denominator in equation (3) can be estimated by ordinary least squares. In the case of the numerator, we estimate a regression of labor market outcomes on the twin instrument and all observable characteristics available at baseline (X), including an indicator for having health insurance and mothers' age. Analogously, in the case of the denominator, we

⁶ Jacobsen et al., (1999) directly estimate the effect of the exogenous variation of the twin shock on labor outcomes using a simple linear regression as shown in (3)

$$Y_i = \gamma Twin_i + X_i' \pi + v_i \quad (3)$$

where $Twin$ is the twin shock, X is a set of control variables and v_i is a disturbance term. Due to the independence between the $Twin$ variable and observable and unobservable variables (except for age), we may expect the coefficient γ to recover the impact of twins during the first birth. Since giving birth to twin does not increase household's size by exactly one person, estimates of γ do not measure the effect that an additional child has on labor market outcomes.

estimate a regression of family size on the twin instrument and all the pre-pregnancy characteristics.

Using the Wald estimator, we can easily illustrate changes in estimations when we control for individual fixed effects. This is important because some variables such as women's health status are unobserved. We exploit the panel structure of the data under the assumption that health and other unobserved factors are time invariant. Then, we compute a Wald estimator controlling for an individual's fixed effects in the numerator and denominator equation. To the best of our knowledge, in the large set of studies using the binary twin instrument, this is the first study that controls for the biases associated with time-invariant unobservables related to the twin birth, such as mothers' health.⁷ In the next section we go further and relax the assumption that twinning is independent of any time varying unobserved heterogeneity.

In Table 3, we present the results for the traditional Wald estimator and the Fixed Effect Wald estimator for the sample of mothers with children younger than 5 years old.⁸ The results for a sample of mothers with children younger than 3 years old are similar and they are presented in Appendix A. The panel at the left of Table 3 shows the reduced form estimates for the numerator and denominator regressions coming from equation (3). As the reader may notice, we estimate effects for three binary outcomes: labor force participation, employment probability and working in the formal sector. The right panel in Table 3 presents the Wald estimators and the traditional t-test for the difference in the estimators. The Wald estimators are computed as the ratio of reduced forms. The estimated parameter of the effect of family size on mothers' labor force participation is -0.077, which corresponds to the ratio between two reduced form coefficients. The coefficient resulting from the regression of mothers' labor force participation on the twin instrument is (-0.064) and the coefficient resulting from the regression of family size on the twin instrument is (0.832).

Our estimates are local and the estimate should be interpreted as the Local Average Treatment Effect. That is, the estimate identifies the change in mothers' labor market outcomes due to a change in the number of children among compliers. In other words, we estimate the

⁷ Recently, Aaronson et al., (2017) used instrumental variables for a wide set of countries and periods. For a subset of countries, they found that once they controlled for contemporaneous women's health measures, the results did not change significantly.

⁸ First stage estimates are shown in the first row on table 3, "household size".

effect that the number of children has on mothers' labor supply across families for which this marginal increase in household size is the result of a twin birth. As the reader may remember, our sample consists of first time mothers who have either a twin birth or a singleton birth; thus, the subpopulation of compliers are all twins' mothers. In fact, the twin birth is the only reason for which a twin's mother would increase their family size by 2 children, in contrast to the increase of 1 child experienced by a mother of a singleton. In this case, the LATE and the average treatment effect (ATE) should be similar. Their differences should be explained only by the margin to which twin births determine family size, an effect that should be close to one. This was the case for the estimates of the reduced forms presented in Table 3 (the estimated effect is between 0.93 and 0.83). Since abortion is penalized in Colombia⁹, we assume full compliance in the case of twinning.

Finally, our estimates can be interpreted as an approach to identify the effect of a marginal child after the firstborn. As has been mentioned in the literature, our methodology does not allow the identification of the marginal effect from having no children to having one; nevertheless, our estimated effects are still relevant since most women in their prime productive years have at least one child (Jacobsen et al., 1999).

The effect of an additional child (second child) is related to a reduction of 5.2 percentage points in women's labor market participation probability, which is equivalent to an 18% relative reduction. In our preferred specification, the one that controls for fixed effects, this effect is even higher and close to -7.7 percentage points (or 26%). These negative effects extend to the employment probability and the probability of being in formal employment. From our preferred specification (instrumental variables fixed effects), the results suggest that the effect of a second child reduces the probability of being an employee at least 5.7 percentage points (or 56%) and the probability of being a formal employee 1.7 percentage points. Nevertheless, the result related to formal employment vanishes once we control for fixed effects.

⁹ Abortion is permitted only in three cases: 1) Due to rape, statutory rape, or incest; 2) To preserve the physical health of the woman (and her life) and 3) Due to fetal inviability.

The comparison between the estimator with and without individual fixed effects is illustrative of the potential degree of endogeneity present in our instrument. For labor market participation and employment probabilities, the point estimate of the LATE is higher than the one estimated with fixed effects. The latter is a more robust estimate. It does not require independence for our instrument with any time-fixed unobserved factor. In the case of formal employment, the point estimates are very similar. In general, we do not observe significant differences between both estimates.

An initial result is that the negative effects of an additional child on female labor force participation and the probability of being employed are robust to the inclusion of fixed effects. It could be the case that there is still correlation between the twin instrument and the time varying un-observables. As Farbmacher et al. (2018) have shown, dizygotic twinning is correlated with maternal characteristics. To tackle this possible source of bias, we relax the assumptions associated with the twin instrument in the next section.

[Insert Table 3 around here]

4. Additional instruments and exogeneity assumption

Recently, economists have raised several concerns regarding the use of multiple-birth or twin births as an instrument for family size. The literature on human reproduction has shown that fertilization of twins, particularly dizygotic twins, is associated with the mother's genetic history, height, body mass index, marital status, ethnicity and age (Bulmer, 1970; Farbmacher et al., 2018; Hall, 2003; Hoekstra et al., 2008). Similarly, literature in economics has shown that fertility treatments, which are usually unobservable for the analyst, are correlated with twin births (Bhalotra & Clarke, 2016; Braakmann & Wildman, 2014b; Caceres-Delpiano, 2012). Additionally, Bhalotra and Clarke (2016) show that twin births require more maternal resources than singletons, and as a consequence, if the mother's health is at risk, a miscarriage of twin births is more likely to occur. Finally, Farbmacher et al. (2018) show that twin births in Sweden are correlated with mothers' observable and unobservable pre-pregnancy characteristics.

In this section, we relax the exogeneity assumptions associated with the twin instrument. Additionally, we introduce multiple additional instruments in order to provide wider evidence on the effect that an additional child has on women's labor force participation and other labor market outcomes.

4.1 Imperfect Instrumental Variables

We start by introducing the concept of Imperfect Instrumental Variables (IIV) described by Nevo and Rosen (2012). In the framework of Instrumental Variables, Nevo and Rosen (2012) show that if instead of assuming zero correlation between the instrument and the error term, one is willing to assume that the correlation not only exists but is lower than the correlation between the endogenous variable and the error term¹⁰, then it is possible to learn about the bounds of the structural parameters of interest. To bound the structural parameter, it is also necessary to add verifiable restrictions on the observable correlation between the instrument and the endogenous variable. In particular, Nevo and Rosen (2012, Lemma 1) show that when the correlation is negative, it is possible to find the two-sided bounds. In contrast, if the correlation is positive, it is only possible to find the upper bound of the structural parameter.

These assumptions are feasible in our set up. First, there may be reasons to believe that the twin instrument is somehow correlated with the error term of the labor market equation. Second, it is very likely that such correlation is smaller than the correlation between the same error term and household size. Finally, as we show in Table 3, the correlation between the twin instrument and household size is positive and close to one. Therefore, it is possible to estimate the upper bound of the structural parameter.

The first column of Table 4 presents the upper bounds for all outcomes considered in the previous sections. In all cases, the upper bound is negative and precisely estimated as different from zero. Additionally, as we expect, the upper bound is closer to zero than the IV estimator. Thus, we interpret these bounds as the smallest negative effect that an additional child as a result of twin births has on labor market outcomes. The results from IIV are fairly consistent with the IV estimates: an additional child reduces mothers' participation in the labor market approximately 5.2 percentage points. The negative effects extend to employment probability and formality. The probability of being employed is reduced by 2.7 ppts, and the probability of being a formal employee is reduced by 2.2 ppts.

[Insert Table 4 around here]

¹⁰ Nevo and Rosen (2012) also assume that the direction of the correlation is the same in both cases.

4.2 Alternative instruments

Human biology literature has shown that even though dizygotic twinning is associated with several factors, evidence on monozygotic twinning remains inconclusive (Hall, 2003). Yet, monozygotic twins are considered a random event (Farbmacher et al., 2018; Hall, 2003; MacGillivray, Samphier, & Little, 1988; Tong & Short, 1998) and are always same-sex twins. Thus, even though twin births may be an imperfect instrument, same-sex twins (Black et al., 2007; Figlio et al., 2014) and monozygotic twins (Farbmacher et al., 2018) may provide a better source of exogenous variation to identify the effect of family size on women's labor market outcomes.

a. Same-sex twins

In most twin databases it is difficult to distinguish between dizygotic and monozygotic twins. Nonetheless, sometimes it is possible to differentiate between same-sex twins and opposite-sex twins. By using only same-sex twins, it is possible to build an instrument that contains a larger fraction of monozygotic twins. Black et al. (2007) and Figlio et al. (2014) use this idea to check the exogeneity of the twin instrument. They compare the estimates using the twin instrument with the ones using the same-sex twin instrument. Following these authors, if the estimates are not significantly different, then there is evidence that the selection on unobservables is not an issue in our data. Otherwise, the validity of the twin instrument would be cast into doubt.

This test is feasible in our set up. We build the same-sex instrument and obtain the new Wald estimators. Then, we compare the estimates for the twin instrument presented in the previous section with the estimates associated with the same-sex twin instrument (see Table 5). In the case of labor force participation and employment, the Wald estimator is more negative when using the traditional twin instrument, while the point estimates for formal employment are almost the same. Nevertheless, there are not significant differences across the different sets of estimates. Thus, we conclude that selection of unobservables does not play an important role in our set up.

[Insert Table 5 around here]

b. Monozygotic twins

Farbmacher et al. (2018) propose a monozygotic twin instrument using observable information of the twin instrument (e.g., same-sex and opposite sex). Even though monozygotic twinning is usually unobserved, their approach does not require observing zygosity. The key idea of the new instrument is to correct for the remaining endogeneity in the same-sex twin instrument using the information from opposite-sex twins (which are dizygotic), which is usually observable. The monozygotic twins instrument is defined as the twin instrument minus twice the opposite-sex twins.

$$\text{Monozygotic_Twin}_i = \text{Twin}_i - 2(\text{Twin}_i - \text{Samesex_Twin}_i)$$

For the monozygotic twins' instrument to be a valid instrument, the following assumptions introduced by Farbmacher et al. (2018) are needed. First, monozygotic twinning is exogenous. Second, dizygotic twinning is equally likely to be same-sex twins as twins of opposite sex. This assumption is called Weinberg's (1901) rule. Finally, they assume that the same-sex instrument is exogenous within the group of mothers with dizygotic twins. Proposition 1, in Farbmacher et al. (2018) shows that under these assumptions, the monozygotic twin instrument is a valid instrument.

Using our data, we test the relevance of the instrument. Similar to the other instruments, we find that the instrument is strong in magnitude and statistically significant at the 5% level. Column 4 of Table 5 reports the Wald estimator. The point estimate for the labor force participation equation is -5.1 ppts with a standard error of 3.2, which means that it is significant at the 11% level. For employment and formal employment, the point estimate is negative but not statistically significant at conventional levels.

Similar to Farbmacher et al. (2018), we perform a sensitivity analysis that builds on the idea of imperfect instruments developed by Nevo and Rosen (2012). Now we assume that the monozygotic twin instrument is less endogenous than the dizygotic twinning instrument. In this case, the instrument is given by the following expression:

$$\text{Monozygotic_Twin}_i(\theta) = \text{Samesex_Twin}_i - \lambda(\theta)(\text{Twin}_i - \text{Samesex_Twin}_i)$$

where $\lambda(\theta)$ is the correction factor for the *Samesex_Twin_i* instrument, and θ represents the correlation between the monozygotic twin instrument and the structural error term.¹¹ The correction factor is given by $\lambda(\theta)=1 - \theta \left(1 - \frac{\pi_{ss}}{\pi_{ms}}\right)$, where π_{ss} is the proportion of same-sex twins in the data and π_{ms} is the proportion of opposite-sex twins. Since π_{ss} is larger than π_{ms} in our data, negative values of θ imply an amplification bias, while positive values of θ imply an attenuation bias of the effect.

Since the degree of endogeneity is not observable, and without further assumptions it is not possible to estimate, Figure 2 shows the Wald estimator for a grid of values of θ between -1 and -1. The figure shows twin birth, same-sex twin and monozygotic twin instruments as special cases.¹² In general, for all negative correlations, there is a negative and significant effect on labor force participation, while for positive correlations there are not significant effects. Similarly, we observe significant effects for employment. Nonetheless, for most values of θ , there are not significant effects on formal employment.

[Insert Figure 2 around here]

5. Conclusions

The identification of the causal effect of fertility on female labor market outcomes has been a topic widely studied in the literature. This is mainly because the number of children in a family is an endogenous decision. Socio-economic characteristics and individual preferences affecting the decision might have also an effect on labor market indicators.

Several strategies have been used to isolate the exogenous variation of the fertility variable. One of them is the use of twin births as an instrument for family size (or number of children in the household). Nevertheless, a set of recent studies have shown a high dependence of twin births on mothers' characteristics such as age, weight, height, income and health. Similarly, with delays in childbearing, women with particular labor market preferences were also more likely to

¹¹ Note that if $\theta = 0$, then $Monozygotic_Twin_i(\theta = 0) = Twin_i - 2(Twin_i - Samesex_Twin_i)$

¹² Consider the following cases:

If $\theta = 1 / \left(1 - \frac{\pi_{ss}}{\pi_{ms}}\right)$, then $Monozygotic_Twin_i(\theta) = Twin_i$

If $\theta = 2 / \left(1 - \frac{\pi_{ss}}{\pi_{ms}}\right)$, then $Monozygotic_Twin_i(\theta) = Samesex_Twin_i$

go through in vitro fertilization, increasing their probability of having multiple births. In such cases, the exogeneity assumption is invalidated.

More recently, several studies have attempted to correct the existent bias in the twin instrument by making use of the twins' sex. This is done by making use of the fact that monozygotic twins are always of the same sex and are thought to be random. This contrasts with dizygotic twins, which can be the result of IVF and can be either same-sex or opposite sex.

In this paper, we overcome those challenges by making use of a unique dataset for a developing country. The data is a census of low-income families in Colombia and allows us to control not only for women's pre-pregnancy characteristics but for time-invariant unobserved variables that have been found to be highly correlated with twin births such as health. Another advantage of the data is that twin mothers from low income families are a less selective sample than what can be observed in the developed world. They are women with lower probabilities of going through IVF, which in turn makes a sample with more random events of twins.

We present the results using the standard twin instrument and then using several of the corrections suggested by the literature. Using the usual twin instrument, we find a negative effect of fertility on female labor force participation of 5.2 percentage points (or 18%). Once we conduct the analysis using same-sex twins, which are more likely to be monozygotic twins, the results do not change. This suggests the absence of selection on unobservables in our sample. When we use the monozygotic twin instrument proposed by Farbmacher et al. (2018), the point estimate for the labor force participation equation is -5.1 ppts, with a p-value for the significance test very near to 10%. Furthermore, from sensitivity analysis based on the idea of imperfect instruments, we find that for all negative correlations between the monozygotic twin instrument and the structural error term, there is a negative and significant effect on labor force participation and employment probability. All this evidence shows the lower degree of selectivity in our sample.

We find similarities with Farbmacher et al. (2018) for an early cohort in Sweden. They do not observe differences between the usual twin instrument and the new one due in part to the absence of IVF in those years. Our sample, which is a census of low-income families, has been shown to be a less selective sample that allows us to reach robust estimates of the causal relationship between fertility and female labor market outcomes. Thus, we conclude that the

twin instrument in our data is not very biased and leads to real estimates of the causal effect under analysis when we control for observable and time-invariant unobservable characteristics. Our results also show that data from developing countries and from low income families should be less contaminated by the use of IVF.

References

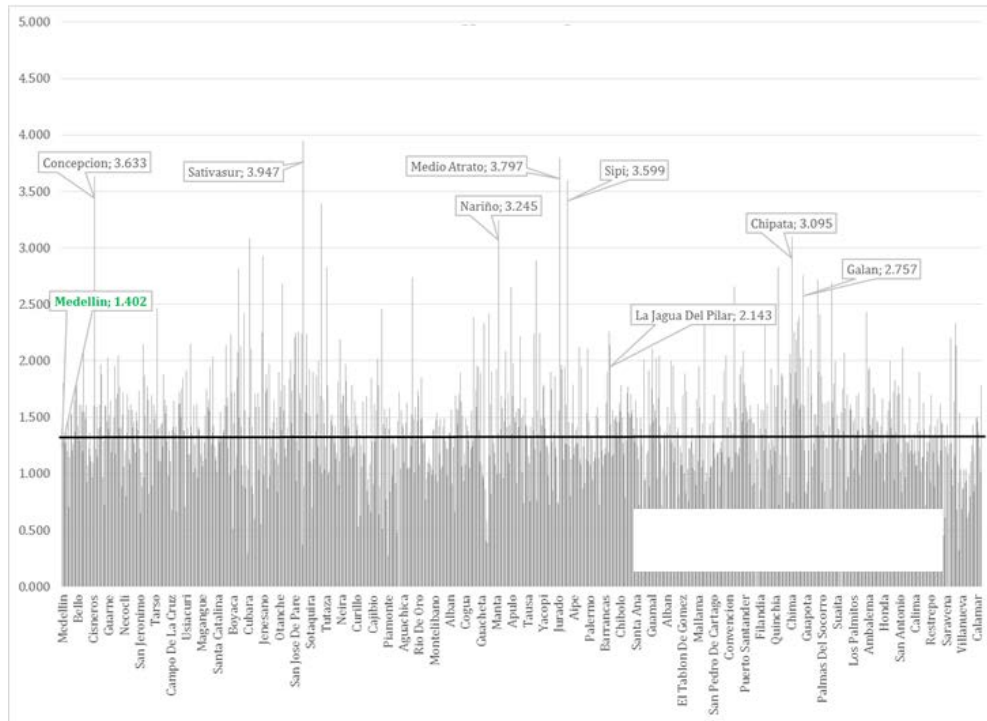
- Aguero, J. M., & Marks, M. S. (2008). Motherhood and Female Labor Force Participation : Evidence from Infertility Shocks. *The American Economic Review*, 98(2), 500–504.
- Amador, D., Bernal, R., & Peña, X. (2013). The rise in female participation in Colombia: Fertility, marital status or education? *Ensayos Sobre Política Económica*, 31, 54–63.
- Angrist, J., & Evans, W. (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *The American Economic Review*, 88(3), 450–477.
- Angrist, J., & Lavy, V. (2010). Multiple Experiments for the Causal Link between the Quantity and Quality of Children. *Journal of Labor Economics*, 28(4), 773–824.
- Arango, L. E., Posada, C., & Charry, A. (2003). *La participación laboral en Colombia según la nueva encuesta : ¿cambian sus determinantes?* (Borradores de economía No. 250).
- Bhalotra, S., & Clarke, D. (2016). *The Twin Instrument* (IZA DP No. 10405).
- Black, S. (2005). The more the merrier? The effect of family composition on children's education. *Quarterly Journal of Economics*, 120(2), 669–700.
- Black, S., Devereux, P., & Salvanes, K. (2007). From the cradle to the labor market ? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122(1), 409–439.
- Braakmann, N., & Wildman, J. (2014a). *Fertility treatments and the use of twin births as an instrument for fertility* (MPRA No. 54106).
- Braakmann, N., & Wildman, J. (2014b). *Reconsidering the impact of family size on labour supply: The twin-problems of the twin-birth instrument* (Working Paper Series in Economics No. 316).
- Bronars, S. G., & Grogger, J. (1994). The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment. *The American Economic Review*, 84(5), 1141–1156.
- Browning, M. (1992). Children and Household economic behavior. *Journal of Economic Literature*, 30(3), 1434–1475.
- Bulmer, M. . (1970). *The Biology of Twinning in Man*. Oxford: Clarendon Press.

- Caceres-Delpiano, J. (2012). Impacts of Family Size on the Family as a Whole : Evidence from the Developing World. *The B.E. Journal of Economic Analysis & Policy*, 12(1).
- Cáceres-delpiano, J. (2008). *Keeping the Best for Last . Impact of Fertility on Mother ' s Employment . Evidence from Developing Countries* (Economic Series No. 32).
- Cáceres-Delpiano, J., & Simonsen, M. (2012). The toll of fertility on mothers ' wellbeing. *Journal of Health Economics*, 31(5), 752–766.
- Clarke, D. (2016). *Fertility and Causality* (CSAE Working Paper No. 2016-32).
- Cristia, J. P. (2008). The Effect of a First Child on Female Labor Supply Evidence from Women Seeking Fertility Services. *The Journal of Human Resources*, 43(3), 487–510.
- Farbmacher, H., Guber, R., & Vikström, J. (2018). Increasing the credibility of the twin birth instrument. *Journal of Applied Econometrics*, 33(3), 457–472.
- Fausser, B. C., Devroey, P., & Macklon, N. S. (2005). Multiple birth resulting from ovarian stimulation for subfertility treatment. *The Lancet*, 365(9473), 1807–1816.
- Figlio, D., Guryan, J., Karbownik, K., & Roth, J. (2014). The Effects of Poor Neonatal Health on Children ' s Cognitive Development. *American Economic Review*, 104(12), 3921–3955.
- Fitzsimons, E., & Malde, B. (2014). Empirically probing the quantity – quality model. *Journal of Population Economics*, 27(1), 33–68.
- Goldin, C. (1990). *Understanding the Gender Gap*. New York: Oxford University Press.
- Hall, J. G. (2003). Twinning. *The Lancet*, 362, 735–743.
- Hoekstra, C., Zhao, Z. Z., Lambalk, C. B., Willemsen, G., Martin, N. G., Boomsma, D. I., & Montgomery, G. W. (2008). Dizygotic twinning. *Human Reproduction Update*, 14(1), 37–47.
- Hotz, V. J., & Sanders, S. G. (2005). Teenage Childbearing and Its Life Cycle Consequences Exploiting a Natural Experiment. *The Journal of Human Resources*, 40(3), 684–715.
- Jacobsen, J. P., Wishart, J., & Rosenbloom, J. L. (1999). The Effects of Childbearing on Married Women's Labor Supply and Earnings: Using Twin Births as a Natural Experiment. *The Journal of Human Resources*, 34(3), 449–474.
- Kulkarni, A. D., Jamieson, D. J., Jones, H. W., Kissin, D. M., Gallo, M. F., Macaluso, M., & Adashi, E. Y. (2013). Fertility Treatments and Multiple Births in the United States. *New England Journal of Medicine*, 369(23), 2218–2225.
- Li, H., Zhang, J., & Zhu, Y. (2008). The quantity-quality trade-off of children in a developing country : identification using chinese twins. *Demography*, 45(1), 223–243.
- MacGillivray, I., Samphier, M., & Little, J. (1988). Factors affecting Twinning. In *Twinning and*

- twins* (I. MacGill). New York: NY:Wiley.
- Marteleteo, L., & de Souza, L. (2012). The impact of family size on children's educational attainment over time: Assessing the exogenous variation in fertility using twins in Brazil. *Demography*, *49*(4), 1453–1477.
- Nakamura, A., & Nakamura, M. (1992). The econometrics of Female Labor Supply and Children. *Econometric Reviews*, *11*(1), 1–71.
- Nevo, A., & Rosen, A. M. (2012). Identification With Imperfect Instruments. *Review of Economics and Statistics*, *94*(3), 659–671.
- Ponczek, V., & Portela, A. (2011). *New Evidence of the Causal Effect of Family Size on Child Quality in a Developing Country* (Working paper No. 4).
- Rosenzweig, M., & Wolpink, K. (1980). Testing the Quantity-Quality Fertility Model : The Use of Twins as a Natural Experiment. *Econometrica*, *48*(1), 227–240.
- Sanhueza, C. (2009). *Family Size and Birth Order in Chile: Using Twins as a Natural Experiment* (ILADES-Georgetown University Working Papers No. 234).
- Shultz, T. (1981). *Economics of population*. Mass: Addison-Wesley Publishers.
- Tong, S., & Short, R. V. (1998). Dizygotic twinning as a measure of human fertility Dizygotic twinning as a measure of human fertility. *Human Reproduction*, *13*(1), 95–98.
- Vitthala, S., Gelbaya, T. A., Brison, D. R., Fitzgerald, C. T., & Nardo, L. G. (2018). The risk of monozygotic twins after assisted reproductive technology : a systematic review and meta-analysis. *Human Reproduction*, *15*(1), 45–55.
- Weinberg, W. (1901). Beiträge zur physiologie und pathologie der mehrlingsgeburten beim menschen. *European Journal of Social Psychology*, *88*(6-8), 346–430.

Figures

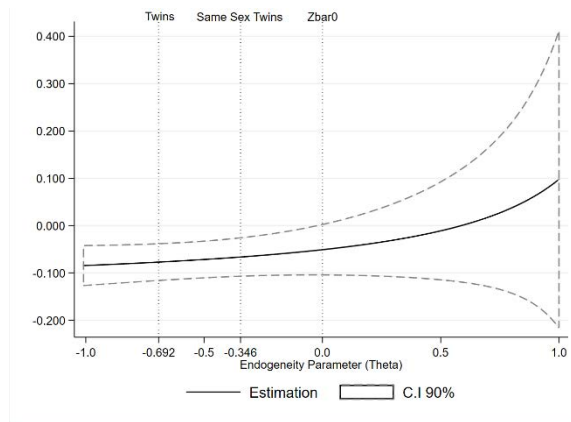
Figure 1. Distribution of the proportion of twins among individuals younger than 15 years old in Colombia



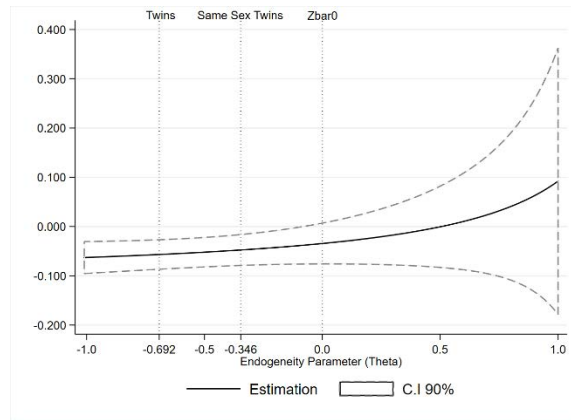
Source: Authors' calculation using data from the 2009-2010 SISBEN

Figure 2. Estimates of the fertility effect on labour market outcomes assuming different degrees of correlation between the instrument and the error of the structural equation

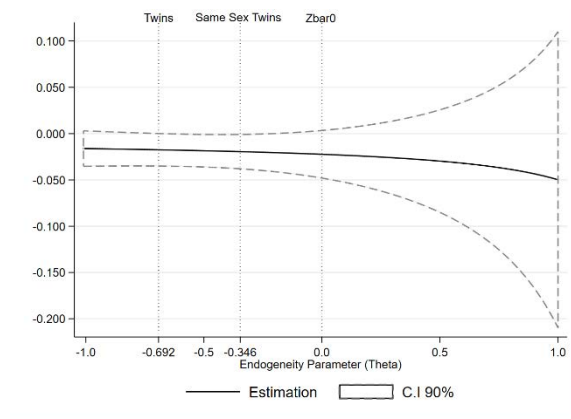
a. Labour force participation



b. Employment probability



c. Working in the formal sector



Tables

Table 1. Prepregnancy characteristics (2003-2007) for first time mothers by 2009-2010. (Children younger than five years old)

	All SISBEN' mothers		Mothers with SISBEN levels' 1 & 2	
	Mean	Standard Error	Mean	Standard Error
Age in Sisben II	20.711	5.149	20.493	5.012
High school	0.600	0.490	0.592	0.492
Head of household	0.033	0.180	0.031	0.174
Married or free marriage	0.078	0.268	0.074	0.262
Health insurance: subsidized	0.360	0.480	0.385	0.487
Health insurance: contributions	0.149	0.356	0.114	0.318
Health insurance: Uninsured	0.491	0.500	0.501	0.500
Own house	0.302	0.459	0.298	0.457
Stratum 0 y 1	0.520	0.500	0.567	0.495
Stratum 2	0.418	0.493	0.430	0.495
Stratum 3-6	0.063	0.242	0.003	0.052
Household size	4.869	2.959	4.973	3.003
Urban zone	0.714	0.452	0.696	0.460
Sisben score less or equal to 11	0.574	0.494	0.630	0.483
Number of rooms	2.813	1.397	2.746	1.365
Sewerage	0.558	0.497	0.524	0.499
Gas	0.302	0.459	0.283	0.450
Aqueduct service	1.270	0.444	1.292	0.455
Electric power	0.930	0.256	0.923	0.267
Fridge	0.461	0.498	0.427	0.495
Washing machine	0.090	0.286	0.062	0.241
Risk of flooding or landslide	0.140	0.347	0.150	0.357
Labor Force Participation	0.308	0.462	0.298	0.457
Working	0.120	0.325	0.106	0.308
Working in the formal sector	0.041	0.199	0.027	0.163
<u>Twins proportion</u>				
Twins	0.006	0.077	0.006	0.077
Same Sex Twins	0.005	0.069	0.005	0.068
Mix Sex Twins	0.001	0.036	0.001	0.035
	230,852		210,563	

Notes: Author's own calculation using data from SISBEN 2005-2007 (II) and SISBEN 2009-2010 (III)

**Table 2. Covariate imbalance' test between mothers of twins and mothers of singletons.
(Children younger than five years old)**

	Mothers from all SISBEN levels		Mothers in SISBEN levels 1 & 2	
	Control means	Coefficient	Control means	Coefficient
Age in SisbenII	20.703 (5.144)	1.394 (0.153)*** [0.000]	20.485 (5.008)	1.257 (0.155)*** [0.000]
High school	0.600 (0.490)	0.017 (0.013) [0.873]	0.591 (0.492)	0.021 (0.014) [0.906]
Head of household	0.033 (0.180)	0.003 (0.005) [0.994]	0.031 (0.174)	0.001 (0.005) [0.995]
Married or free marriage	0.078 (0.268)	0.021 (0.008)*** [0.998]	0.074 (0.262)	0.011 (0.008) [0.906]
Medical care availability: Subsidized	0.360 (0.480)	-0.006 (0.013) [0.989]	0.385 (0.487)	-0.002 (0.014) [0.998]
Medical care availability: Contributory	0.149 (0.356)	0.023 (0.010)** [0.979]	0.114 (0.318)	0.014 (0.009) [0.906]
Uninsured	0.491 (0.500)	-0.017 (0.013) [0.998]	0.501 (0.500)	-0.012 (0.014) [0.993]
Own house	0.302 (0.459)	-0.009 (0.012) [0.989]	0.298 (0.457)	-0.009 (0.013) [0.993]
Stratum 0 y 1	0.520 (0.500)	-0.020 (0.013) [0.966]	0.567 (0.495)	-0.011 (0.014) [0.993]
Stratum 2	0.418 (0.493)	0.002 (0.013) [0.979]	0.430 (0.495)	0.009 (0.014) [0.993]
Stratum 3-6	0.062 (0.242)	0.018 (0.007)** [0.784]	0.003 (0.052)	0.002 (0.002) [0.916]
Household size	4.869 (2.959)	-0.024 (0.082) [0.998]	4.973 (3.002)	0.011 (0.088) [0.998]
Urban zone	0.714 (0.452)	0.013 (0.012)	0.696 (0.460)	0.012 (0.013)

		[0.916]		[0.993]
Sisben score less or equal to 11	0.574 (0.494)	-0.031 (0.013)**	0.630 (0.483)	-0.021 (0.014)
		[0.500]		[0.906]
Number of rooms	2.812 (1.398)	0.130 (0.033)***	2.746 (1.366)	0.109 (0.034)***
		[0.263]		[0.094]
Sewerage	0.558 (0.497)	-0.002 (0.013)	0.524 (0.499)	-0.008 (0.014)
		[0.998]		[0.998]
Gas	0.302 (0.459)	0.012 (0.012)	0.283 (0.450)	0.007 (0.013)
		[0.998]		[0.998]
Aqueduct service	1.270 (0.444)	-0.009 (0.012)	1.292 (0.455)	-0.008 (0.013)
		[0.998]		[0.993]
Electric power	0.930 (0.256)	0.001 (0.007)	0.923 (0.267)	0.000 (0.008)
		[0.998]		[0.998]
Fridge	0.461 (0.498)	0.034 (0.013)**	0.427 (0.495)	0.024 (0.014)*
		[0.870]		[0.831]
Washing machine	0.089 (0.285)	0.016 (0.008)*	0.062 (0.241)	0.009 (0.007)
		[0.998]		[0.921]
Risk of flooding or landslide	0.140 (0.347)	-0.005 (0.009)	0.150 (0.357)	-0.003 (0.010)
		[0.998]		[0.998]

Number of Observations

230,852

210,563

Notes: The reported coefficients are the result of regressing each covariate on a twin birth indicator. Robust standard errors are reported in parenthesis while the p-value of the multiple testing hypothesis (M-T) of R&W are reported in brackets. M-T allows making inferences about the hypothesis that the treatment effect is zero for all the result variables included in the baseline. *, **, and *** refers to the 10%, 5% and 1% level of significance, respectively

Table 3. IV estimates of the effect of family size on labour market outcomes. Mothers in levels 1 & 2 of SISBEN with children younger than five years old.

	Reduced Form				Wald Estimator (LATE)		
	Control's mean (1)	w/o. Fixed Effects (2)	w. Fixed Effects (3)	Test difference (3) - (2) (0,069)	w/o. Fixed Effects (4)	w. Fixed Effects (5)	Test difference (5) - (4) (0,150)
Household Size	4,973 -3,002	0,928 (0,058)**	0,832 (0,105)***	-0,096 (0,069)	-	-	-
Labor Force Participation	0,298 -0,457	-0,048 (0,013)***	-0,064 (0,018)***	-0,016 (0,016) [0,203]	-0,052 (0,014)*** [0,000]	-0,077 (0,024)*** [0,001]	0,025 (0,015) [0,150]
Working	0,106 -0,308	-0,025 (0,012)** [0,033]	-0,047 (0,014)*** [0,002]	-0,022 (0,014) [0,017]	-0,022 (0,013)** [0,036]	-0,057 (0,018)** [0,002]	0,03 (0,017) [0,017]
Working in the formal sector	0,027 -0,163	-0,021 (0,008)*** [0,015]	-0,015 (0,009)* [0,088]	0,006 (0,009) [0,188]	-0,022 (0,008)*** [0,008]	-0,017 -0,011 [0,103]	-0,005 (0,012) [0,412]
Number of Observations	210563						

Notes: Robust standard errors in parenthesis. Prepregnancy characteristics shown in Table 1 were included. *, **, *** corresponds to the 10%, 5% and 1% level of significance respectively.

Table 4. IIV estimates (Nevo & Rosen, 2012) of the effect of family size on labour market outcomes. Mothers in levels 1 & 2 of SISBEN with children younger than five years old

Mothers with Children ages <=5				
SISBEN LEVELS I & II	N&R Bound	Twins	Z Bar	Same Sex Twins
	(1)	(2)	(3)	(4)
Labor Force Participation	-0,052 [-0,019]	-0,077 (0,024)*** [0,001] {10,603}	-0,051 (0,032) [0,117] {2,463}	-0,066 (0,025)*** [0,007] {7,193}
Working	-0,027 [-0,007]	-0,057 (0,018)*** [0,002] {9,663}	-0,034 (0,025) [0,172] {1,863}	-0,048 (0,019)** [0,013] {6,183}
Working in the formal sector	-0,022 [-0,003]	-0,017 (0,011) [0,103] {2,653}	-0,022 (0,016) [0,155] {2,033}	-0,019 (0,011)* [0,086] {2,953}
Number of Observations	210563			

Notes: Standard errors in parenthesis. *, **, *** correspond to 10%, 5% and 1% level of significance, respectively.

() Standard error; [] p-value; { } Upper Bound CI.

Table 5. Estimates comparison among the different set of instruments, including same-sex twins

Mothers with Children ages <=5							
SISBEN LEVELS I & II	N&R Bound	Twins	Same Sex Twins	Z Bar	Testing differences		
	(1)	(2)	(3)	(4)	(2)-(3)	(2)-(4)	(4)-(3)
Labor Force Participation	-0,052 [-0,019]	-0,077 (0,024)*** [0,001] {10,603}	-0,066 (0,025)*** [0,007] {7,193}	-0,051 (0,032) [0,117] {2,463}	-0,011 [0,315]	-0,026 [0,301]	-0,015 [0,294]
Working	-0,027 [-0,007]	-0,057 (0,018)*** [0,002] {9,663}	-0,048 (0,019)** [0,013] {6,183}	-0,034 (0,025) [0,172] {1,863}	-0,009 [0,269]	-0,022 [0,258]	-0,013 [0,253]
Working in the formal sector	-0,022 [-0,003]	-0,017 (0,011) [0,103] {2,653}	-0,019 (0,011)* [0,086] {2,953}	-0,022 (0,016) [0,155] {2,033}	0,002 [0,704]	0,005 [0,706]	0,003 [0,709]
Number of Observations	210563						

Notes: Standard errors in parenthesis. *, **, *** correspond to 10%, 5% and 1% level of significance, respectively. () Standard error; [] p-value; { } Upper Bound CI.

Appendix A: The results for a sample of mothers with children younger than 3 years

Father's with Children ages <=3				
All Data	N&R Bound (1)	Twins (2)	Z Bar (3)	Same Sex Twin (4)
Labor Force Participation	-0,021 0,016]	0,008 (0,029) [0,781] {0,083}	0,010 (0,032) [0,764] {0,093}	0,010 (0,032) [0,764] {0,093}
Working	-0,032 0,011]	-0,028 (0,032) [0,379] {0,773}	-0,030 (0,035) [0,395] {0,723}	-0,030 (0,035) [0,395] {0,723}
Working in the formal sector	-0,004 -0,002]	-0,019 (0,013) [0,135] {2,233}	-0,013 (0,014) [0,363] {0,833}	-0,013 (0,014) [0,363] {0,833}
Number of Observations	93865			

