

Intra-household child penalty in Brazil: Formal employment and entrepreneurship gaps among vulnerable families

Soraya Roman* Janaína Feijó Valdemar Pinho Neto Cecilia Machado

Center for Empirical Studies in Economics - FGV
April 2024

Abstract

What are the gender differences in employment outcomes of couples from developing countries after the birth of their first child? Are women disproportionately negatively affected as in developed countries? Using public administrative records, we follow couples from low-income families in Brazil to estimate the changes in the intra-household gender gaps of being in the formal labor market and opening a small enterprise three years before and four years after the birth of their first child. Consistent with the literature, our analysis reveals a stark gender penalty in formal employment for women that accounts for 45% of women's participation before childbirth. The formal employment penalty happens regardless of women's relative educational advantage and is larger for poorer families. However, we also find that, rather than exiting the labor market completely, some women opt to open a small enterprise after the birth of their child: The probability of opening a small enterprise is 5% higher for women than men. Not all women opt for this alternative; the relative rise in female small enterprises comes from families in urban areas with a basic needs index below the median. Furthermore, this entrepreneur gap is higher among women with more than primary education, whose partners are also equally or more educated, or with jobs in the formal labor market before the birth of the first child.

Keywords: gender gap, child penalty, self-employment, developing countries

JEL Classification: J13, J16

1 Introduction

Nowadays, parenthood is one of the critical determinants of gender gaps in the labor market, both in developed and developing countries. Extant empirical studies have shown that after the birth of the first child, women's labor participation and earnings decrease sharply, while men's remain unaltered (e.g. Aguilar-Gomez, Arceo-Gomez, & la Cruz Toledo, 2019; Berniell, Berniell, de la Mata, Edo, & Marchionni, 2021, 2023; Berniell, Berniell, Mata, et al., 2021; de Quinto, Hospido, & Sanz, 2021; Kleven, Landais, & Sjøgaard, 2019; Rabaté & Rellstab, 2022). The gender differences in the labor market that surge after the first child are known as *the child penalty*. One explanation for this

*Graduate School of Economics, Catholic University of Brasilia

phenomenon is that women are still seen as the main caregivers and responsible for most domestic work. Thus, they look for alternative employment arrangements, more flexible and family-friendly, i.e., arrangements compatible with child-rearing (Almond, Cheng, & Machado, 2023; Cortés & Pan, n.d.; Goldin, 2014).

The manifestation of the child penalty differs in developing and developed countries. Recent studies have shown that in Latin America the child penalty in employment is smaller than this penalty in developed countries. One of the reasons is the informal sector which offers a more flexible yet more vulnerable job alternative to unemployment for women after childbirth (Berniell, Berniell, de la Mata, et al., 2021; Berniell et al., 2023). On the other hand, conservative views about women’s role within the family - gender norms - increase the child penalty and also explain the low female labor participation in the region (Berniell et al., 2023; Molina, 2022).

In this paper, we exploit Brazilian public administrative records from different sources to identify the child penalty of a sample of low-income families in Brazil. We follow couples from CadUnico (Cadastro Unico), the registry of social programs beneficiaries in Brazil, over eight years, three years before and four years after the birth of their firstborn. Combining CadUnico with RAIS (Relação Anual de Informações Sociais) and MEI (Microempreendedor Individual) databases, we are able to identify if the couples are in the formal labor market or if any of them opened a micro-enterprise.

We estimate the intra-household gender gap using Kleven et al. (2019)’s event study model joint with couple fixed effects. By doing so, we control for observed and unobserved attributes of the woman and her partner. Under a collective labor supply model, the decision to have children comes at a cost, expressed in future income loss and a subsequent adjustment of the time and effort in home production. How the losses will be distributed would likely depend on the bargaining power and comparative advantages of each parent in the labor market (Chiappori, 1992). Our estimate of the intra-household gender gap allows us to see if the losses after the birth of the first child are unequally split. Then, we conduct several split-sample analyses to understand whether (1) the family’s initial socioeconomic conditions, and (2) the relative advantage of one partner in terms of education and employment background (formal or non-formal) affect the child penalty.

We find significant child penalties for women relative to men in our sample. There is a drastic drop in women’s participation relative to men in the formal labor market after the child is born. The year the child is born, the penalty is 29%, widening further the next year and stabilizing around 45% four years later. However, the situation is the opposite for the creation of small enterprises, where we find no penalty but an increase in women’s participation relative to men’s. Four years after the birth of the firstborn, the probability of opening a small enterprise is 5% higher for women relative to men.

The intra-household gender differences in the formal labor market and entrepreneurship activities vary depending on the household’s socioeconomic conditions and the couple’s background. The formal employment penalty is larger for poorer families, with relatively worse living conditions, but at the same time is among these families where the participation of women in small enterprises increases relative to men. The educational levels are particularly relevant to observe an enterprise gender gap in favor of women. Families where both partners have primary education at most present larger penalties for women in the formal labor market with no compensation in the creation of small enterprises. The enterprise gender gap favors women in families where both partners have more

than primary education.

We find no evidence suggesting that couples decide to participate in the labor market based on their comparative advantages. Even for families where the woman is more educated than the man or where the man had no formal job before the birth of the child, the employment penalty against women in the formal labor market still occurs. However, we find that having a partner with a job in the formal labor market before the birth of the first child seems to be advantageous for women as it increases their chances of having a formal job or opening a small enterprise.

This paper contributes to the child penalty literature in developing countries, which until recently have focused on the analysis of formal labor markets (Aguilar-Gomez et al., 2019; Medeiros, 2022; Querejeta & Bucheli, 2023). Findings outside the formal labor market are scarcer, the exception being Berniell et al. (2023) and Berniell, Berniell, de la Mata, et al. (2021)'s work on Latin American markets. Our results are consistent with their findings that show that women opt out of the formal labor market in search of more flexible jobs, usually in the informal sector. However, contrary to previous studies, we focus on intra-household gender gaps. Although estimates of intra-household exist for developed countries (Almond et al., 2023; Angelov, Johansson, & Lindahl, 2016; Rabaté & Rellstab, 2022), we find no similar estimates in developing countries. Our findings suggest couples do not choose their participation in the labor market based on comparative advantages. Some evidence of talent misallocation from developed countries supports this finding (Almond et al., 2023; Berniell, Berniell, Mata, et al., 2021). However, there is also some evidence supporting collective decision-making and the use of comparative advantages within a household (Angelov et al., 2016; Rabaté & Rellstab, 2022).

The paper is organized as follows. The current section contains the introduction. Section 2 describes the data and section 3 the event study methodology. Section 4 presents the results. Finally, section 5 describes the conclusions.

2 Data

We use a sample of 1,054,488 couples from CadÚnico (Cadastro Único) that we follow for eight years, three before the birth of their first child, and four years after¹. Their firstborns belong to one of three cohorts: 2013, 2014, or 2015, thus, the span of our sample is from 2010 to 2019. We focus on two labor market outcomes: Whether they enter or exit the formal labor market and whether they create or not a small enterprise. The information about their status in the formal labor market comes from the RAIS database, whereas the source of the small enterprises' creation is the MEI database². The period of analysis was determined by the restrictions of the data sources; mainly, the earliest CadÚnico database of good quality available to us is from 2012, and the creation of small enterprises using the MEI system expanded after 2010. Further details regarding the sample selection are in Section A of the appendix.

¹Cadastro Único is a registry from the Ministry of Citizenship that serves to identify and characterize low-income families and gives support to the implementation of social programs in Brazil, being *Bolsa Família* one of the most important

²The RAIS database is an annual registry from the Ministry of Employment and Labor that collects information of workers and firms ruled by the labor legislation. The MEI database is composed of fiscal administrative records from micro-enterprises ruled by the simplified taxing regime, detailing the date these businesses were created.

Table 1 shows the average characteristics of the couples in the sample. The mean age of the couples at the birth of their first child is 25, but women are younger than men. There are no significant differences between men’s and women’s races. Most of the couples are brown-skinned (56%), and nearly a third are white (27%). The median educational level, complete primary school, is below national averages (9.8 years of schooling in 2012 and 11.6 in 2019), showing the social vulnerability of the sample. However, women are more educated than men: Only 35% of men completed secondary education, whereas that percentage rises to 41% for women.

Couples in our sample were either current or past beneficiaries of social programs in Brazil. These programs are designed to support families rather than individuals, with most social aid directed to the family’s mother. Therefore, it is less common to have information about the couple’s employment status before the birth of their child, as they usually register for these programs near the time of birth (further details in Section A of the appendix). However, by combining CadUnico with RAIS and MEI data, we were able to determine the employment status of 70% of the couples in both formal and informal labor markets before the birth of their first child.

Table 1: Individual’s average characteristics

	All	Women	Men
Age at child’s birth	25.14	23.09	27.18
Race:			
- White	27.77	28.81	26.72
- Black	6.05	6.00	6.09
- Yellow	0.47	0.51	0.43
- Brown-skinned	55.98	56.36	55.61
- Indigenous	0.79	0.86	0.71
- Missing	8.95	7.46	10.43
Education:			
- No education	2.33	1.55	3.10
- Primary incomplete	26.71	21.41	32.02
- Primary complete	17.26	17.51	17.01
- Secondary incomplete	10.44	11.94	8.95
- Secondary complete	37.94	41.33	34.55
- Higher education	2.53	3.26	1.80
- Missing	2.78	3.00	2.55
Employment status before first child:			
- Formal	37.96	26.23	49.69
- Small entrepreneur	8.75	7.80	9.71
- Informal	8.04	4.46	11.63
- Unemployed	14.37	24.51	4.23
- Other jobs	0.42	0.54	0.30
- Missing	30.46	36.46	24.45
Observations	2108976	1054488	1054488

Notes: Race, education and employment status before first child are expressed in percentages. The “Missing” category in race and education indicates that no information was reported over the period of study. For the employment variable, “Missing” indicates that no information was reported before the birth of the first child.

Before the birth of their child, 50% of the men had a formal job, i.e., a job with a labor con-

tract, 12% had an informal job (self-employed, without contract or temporary), 10% were small entrepreneurs, i.e., they had opened a micro-entrepreneur identifier (MEI) previously, and 4% were unemployed. Women’s participation in the formal labor market before the birth of their first child was half of men’s; 25% of them were unemployed, and only 12% were small entrepreneurs or in the informal market.

In addition to the couple’s characteristics, our dataset has information about the families’ living conditions for 93% of the couples (See Table 2). Families are primarily nuclear; the number of extended relatives living with them is nearly zero. Three-quarters of the families live in urban areas, and close to 70% are located in the Northeast and Southeast regions. Access to electricity, potable water, and basic sanitation are three of the least satisfied basic needs: 40% of the families have no basic sanitation, 30% of them have no potable water, and 20% no access to electricity. Using these variables in addition to floor quality and garbage collection, we created a satisfied basic needs index that estimates the average percentage of satisfied basic needs. The index states that, on average, families’ living conditions satisfy 75% of the five basic needs available in our dataset.

Table 2: Household characteristics

	Mean	(S.D.)
Number of extended family members	0.0424	(0.265)
Rural area	25.64	(43.66)
Regions:		
- North	10.91	(31.18)
- Northeast	37.17	(48.33)
- Southeast	32.52	(46.84)
- South	12.81	(33.42)
- Central-west	6.581	(24.80)
Family income before firstborn	117.1	(152.3)
Satisfied basic needs index	0.753	(0.267)
Electricity	79.05	(40.69)
Basic sanitation	58.30	(49.31)
Water	71.64	(45.07)
Floor quality	96.59	(18.14)
Garbage collection	77.54	(41.73)
Observations	977326	

Notes: S.D. = Standard Deviation. The statistics were estimated over all non-missing observations. There are 7.3% missing observations from the initial sample. Family income statistics are estimated with 50% of the observations due to missing information on household characteristics before the birth of the first child. Rural area, regions, electricity, basic sanitation, water, high-quality floor, and garbage collection are expressed in percentages. With the exception of family income, each variable was estimated using the most frequent value observed during the period of study.

3 Methods

Our study follows closely Kleven et al. (2019)’s event study methodology. Kleven et al. (2019) estimate two stand-alone equations for men and women that measure the changes in labor market outcomes after the birth of their first child compared to the year before the birth. Since we are focused on a sample of couples instead of two independent samples of men and women, we use a joint model where the estimate of interest is the gender gap. This specification allows us to introduce

a couple fixed effect that can control for observed and unobserved couple’s characteristics fixed in time (differentials in labor performance, preferences for childcare, values, etc.). The specification is as follows:

$$\begin{aligned}
Y_{icst} = & \sum_{j \neq -1} \alpha_j I[j = t] \times G_i + \sum_k \beta_k I[k = age_{is}] \times G_i + \sum_y \gamma_y I[y = s] \times G_i \\
& + \sum_{j \neq -1} \alpha_j^w I[j = t] + \sum_k \beta_k^w I[k = age_{is}] + \sum_y \gamma_y^w I[y = s] + \epsilon_c + v_{icst}
\end{aligned} \tag{1}$$

where Y_{icst} is a labor market outcome of individual i , from couple c , in year s and at event time t . G_i is a gender dummy that takes the value of one if the individual is a man and zero otherwise. Variables $I[j = t]$ are a set of event time dummies, with the event of interest being the birth of the first child. Event time -1 is omitted, hence, gender gaps are measured as differences with respect to their values a year before the birth of the child. $I[k = age_{is}]$ are a set of age dummies that control for life-cycle trends, and $I[y = s]$ are a set of year dummies that control for cohort effects, basically any unobserved time trend such as inflation or business cycle trends. Finally, ϵ_c is a couple fixed effect, and v_{icst} is an idiosyncratic error.

The parameters of interest are the α_j coefficients as they measure the intra-household gender differences in outcomes before and after the birth of the child. For these coefficients to have a causal interpretation, we assume that the timing of parenthood is not based on expected changes in the within-couple gender gap in the absence of the child.

Following Kleven et al. (2019), we express the intra-household gender gap at time t as a percentage of women’s outcome in the absence of a child, $\tilde{Y}_{icst}^w = \sum_k \beta_k^w I[k = age_{is}] + \sum_y \gamma_y^w I[y = s] + \epsilon_c$. Thus, the child penalty is given by:

$$P_t = -\frac{\alpha_t}{E[\tilde{Y}_{icst}^w | t]} \tag{2}$$

The gender gaps are estimated as the difference between men’s and women’s outcomes. However, to ease the interpretation of the child’s penalty, we add a minus sign at the beginning of the previous expression so we can interpret the outcome as a change of women’s outcome with respect to men’s.

4 Results

In the following sections, we describe the estimates of the intra-household *child penalties* on formal employment and the creation of small enterprises. First, we show the results for the entire sample. Next, we conduct several split-sample analyses to identify whether (1) the family’s socioeconomic conditions and (2) the relative advantage of one partner in terms of education and employment background affect the child penalty.

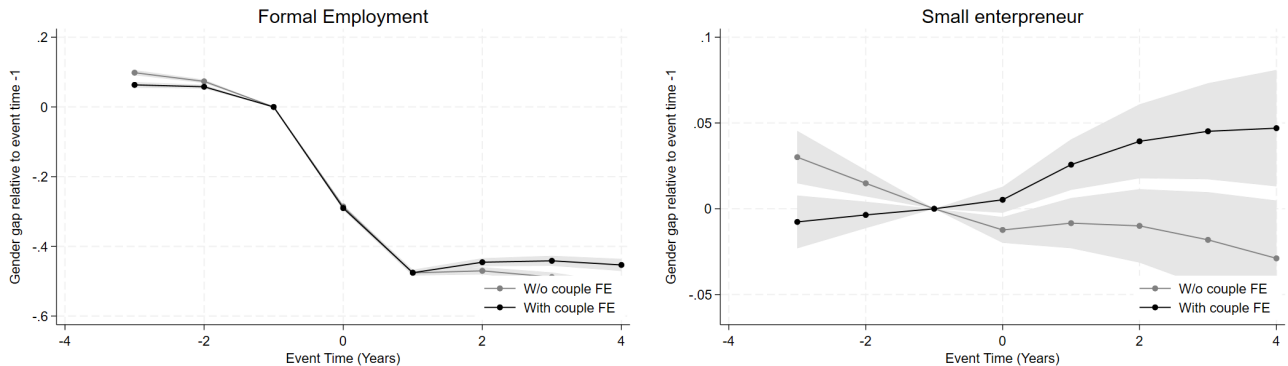
4.1 Child penalty on formal employment and the creation of small enterprises

Consistent with the literature, we find a stark formal employment penalty for women relative to men after the birth of the first child. As shown in the left side of Figure 1, the year the first child

is born, women’s participation in the formal labor market drops by 29% in relation to men’s. In the subsequent years, the penalty augments, stabilizing at 45% by year four after the child’s birth. The results are robust to the inclusion of couple FE. As can be seen, the black and gray lines of Figure 1, which present the model results with and without couple fixed effects, are very similar. This result implies that couple’s gender differences mainly come from within the household instead of systematic differences in the characteristics of men and women.

In addition, we find a statistically significant within-couple gap in the creation of small enterprises that favor women. This result appears after including couple fixed effects in the estimation. One year after the birth of the child, women’s probability of opening a small enterprise increases with respect to men. The gender gap is of 2.6% a year later and rises to 4.7% four years after the birth of the child. Similar results were found in other Latin American countries, where women migrated to more informal or independent sectors of the labor market after the birth of the first child (Berniell, Berniell, de la Mata, et al., 2021; Berniell et al., 2023).

Figure 1: Child penalty on employment



Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. Sample size: 1,054,488 families from CadUnico.

4.2 Regional effects

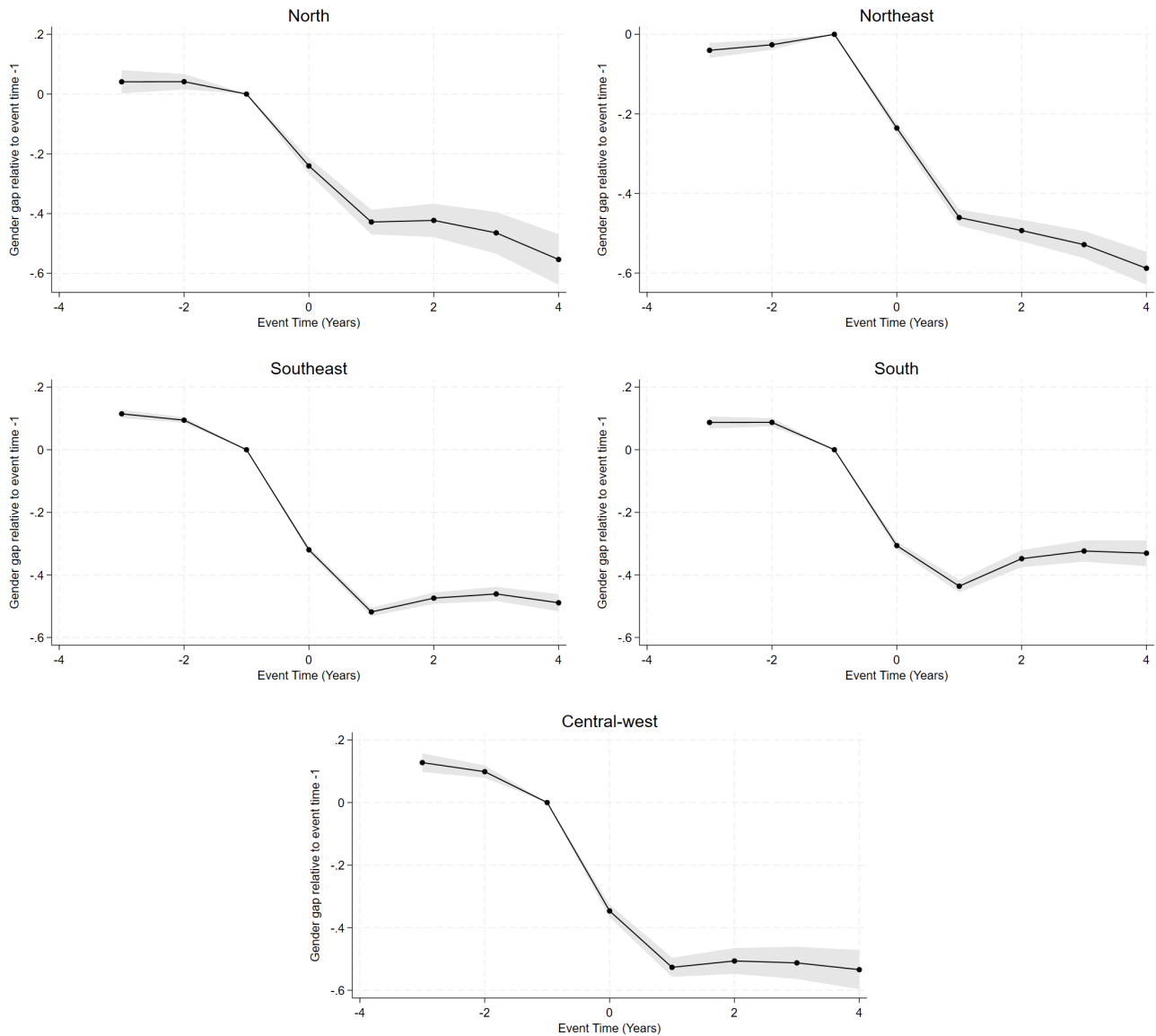
Next, we analyze if the intra-household gender gaps vary by region. Henceforth, all estimates include couple fixed effects. Figure 2 shows women’s formal employment penalty by region. As can be seen, significant penalties occur across all regions, but the largest are in the North and Northeast regions. Four years after the birth of the first child, the child penalties in the North and Northeast regions are 55 and 59%, respectively. The penalties in the remaining regions are close to the complete-sample penalty. One reason to explain these results could be related to the fact that North and Northeast regions are known to have higher incidences of domestic violence and more conservative gender norms (Perova & Reynolds, 2017). Another possibility that we explore later is that these regions are among the poorest regions in the country.

Figure 3 details the intra-household gender differences in the creation of small enterprises. Contrary to the gender gaps in formal labor markets where the child penalty is observed in every region, the gap in the creation of small businesses is increasing and statistically significant only in the North and Southeast regions. In the Southeast region, the gap is twice the complete-sample level (9.4%

gap). A possible explanation for observing wider gaps in the Southeast region is the characteristics of its labor markets, as they harbor several industries and services and are more urban than other regions. This may be a suitable context to open businesses.

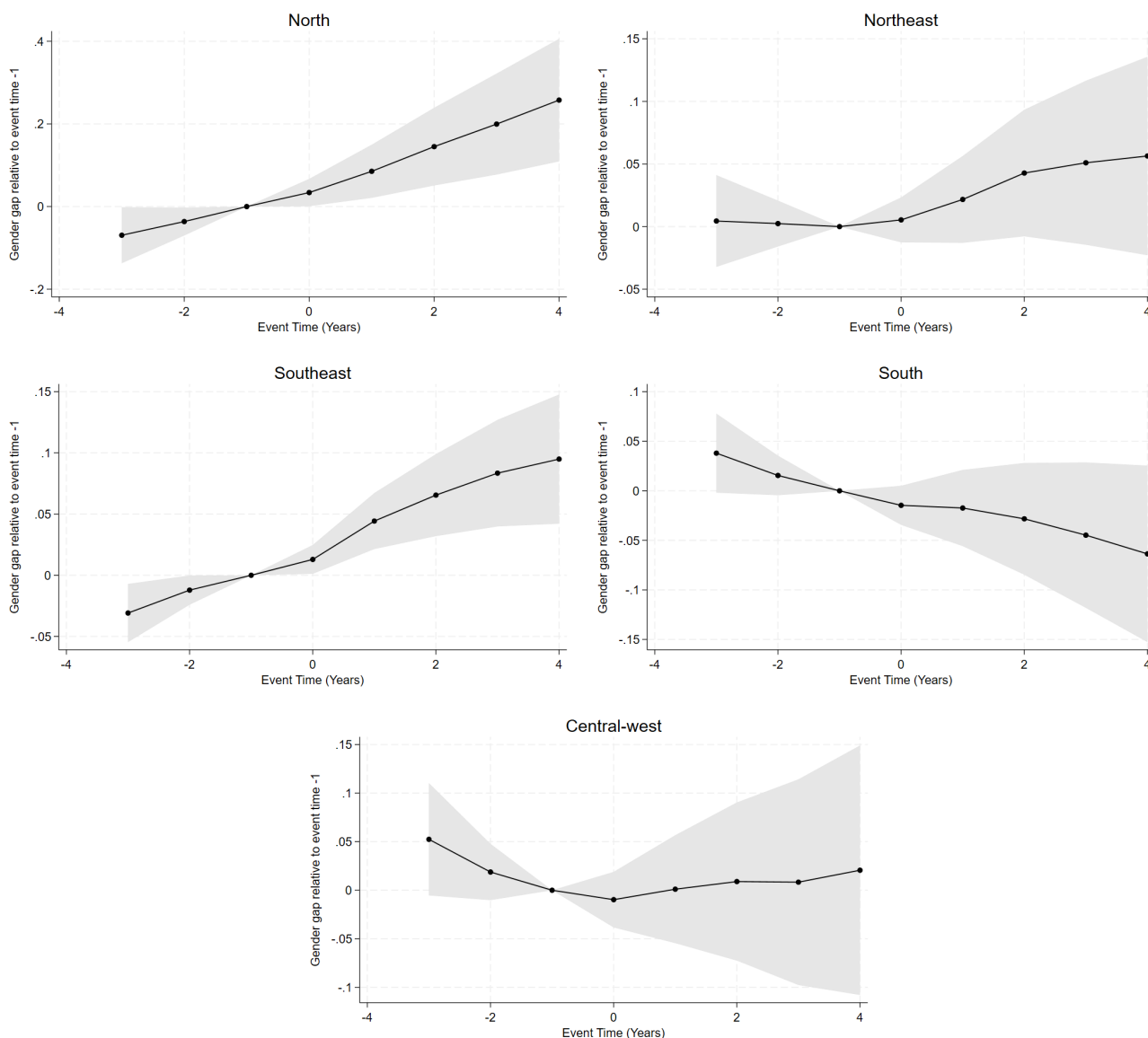
In the North region, the intra-household gender gap is more than four times higher than the complete-sample level (25% gap). Having a high gender gap in this region cannot be correlated with urbanization or the availability of industries and services because the North region is a low-density, more rural area. However, the reason for these results could be correlated with the poverty of the North area, as we explore below.

Figure 2: Child penalty on formal employment by region



Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 977,326 families from CadUnico.

Figure 3: Intra-household gender gaps on the creation of small enterprises by region



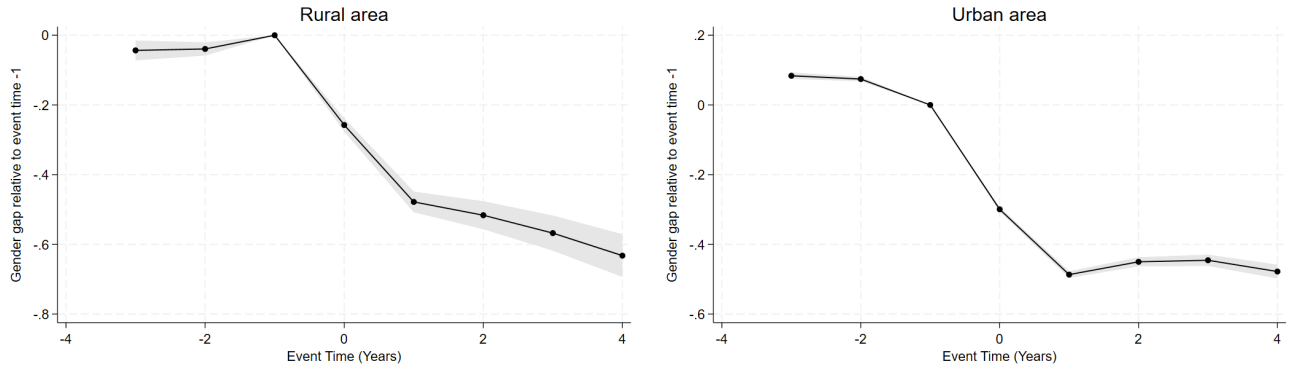
Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 977,326 families from CadUnico.

4.3 Differentiated gender gaps by living conditions

Figure 4 shows that formal employment penalty against women after the birth of their firstborn exists both in urban and rural areas. However, penalties are larger in rural areas. In these areas, the child penalty is 25% the year of childbirth and continuously increases to 63% four years later. Female formal employment participation rate in rural areas would be 10% in the absence of a child, thus, having a 63% penalty implies a reduction of 6.3 percentage points in the participation rate. On the other hand, penalties in urban areas mimic the complete-sample results, reaching a penalty of 47% four years after the birth of the child. Participation rates in formal labor markets in urban areas are higher than in rural areas; 26% of women would be employed in formal jobs in the absence

of a child. Under these circumstances, the penalty implies a reduction in participation rate of 12.2 percentage points.

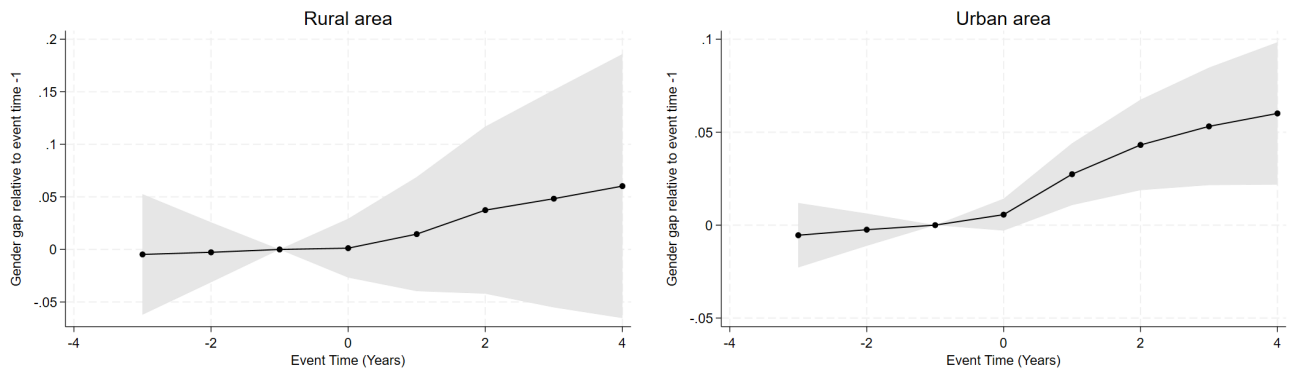
Figure 4: Child penalty on formal employment by area of residence



Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 977,326 families from CadUnico.

The increase in women’s participation in the creation of small enterprises in relation to men after the birth of their firstborn is coming from urban areas. Although both areas present a positive trend in the intra-household gender gap, as shown in Figure 5, the estimates are only significant for urban areas. In both areas, the gender gap is close to the complete-sample results of 5% four years after the birth of the child.

Figure 5: Intra-household gender gaps on the creation of small enterprises by area of residence

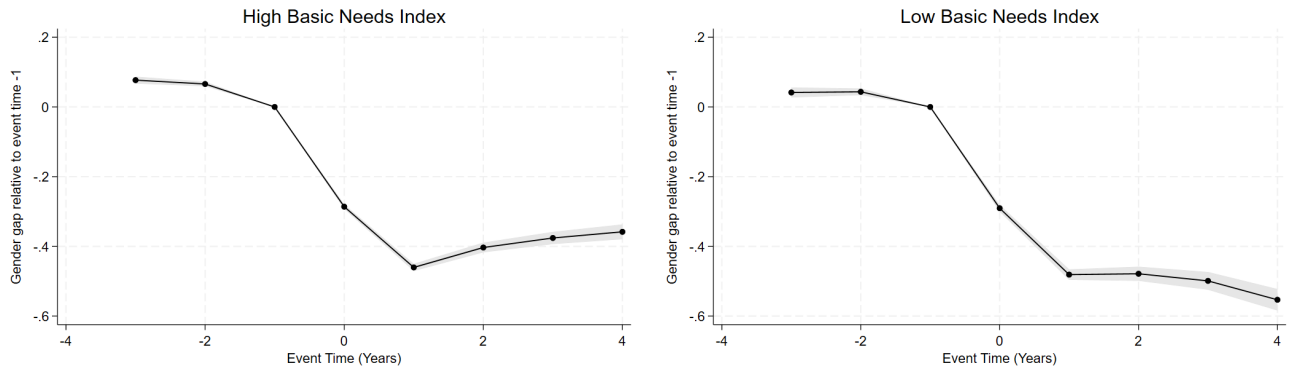


Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 977,326 families from CadUnico.

We additionally explore whether a family’s broader living conditions, beyond the areas of residence, affect the intra-household gender gaps. Specifically, we test whether having differentiated access access to basic needs, such as water and electricity, affects these gaps. To do so, we divided our sample between families with a satisfied basic needs index below the median (0.81) and families with such an index above it. As Figures 6 and 7, both the child penalty on formal employment and the intra-household gender gap on the creation of small enterprises are more prominent in families

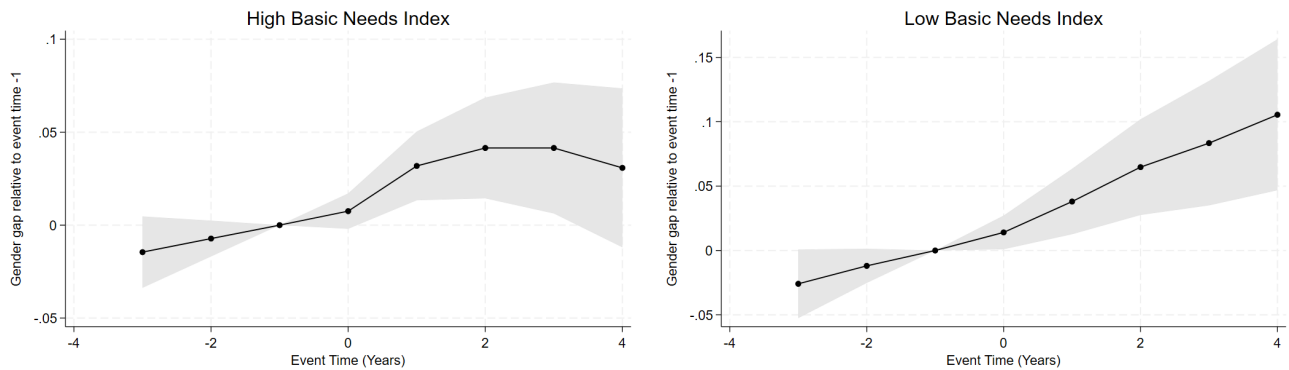
with low basic needs index, i.e., poorer families. Four years after the birth of the first child, the formal employment penalty is 55% for low-basic-needs families, while it is 36% for high-basic-needs families. Likewise, the intra-household gender gap in small enterprise creation reaches 10.5% for low-basic-needs families four years after the birth of the child, doubling the value of high-basic-needs families.

Figure 6: Child penalty on formal employment by satisfied basic needs



Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 977,326 families from CadUnico. The sample is split into a High and Low group using the median of the satisfied basic needs index (0.81).

Figure 7: Intra-household gender gap on the creation of small enterprises by satisfied basic needs



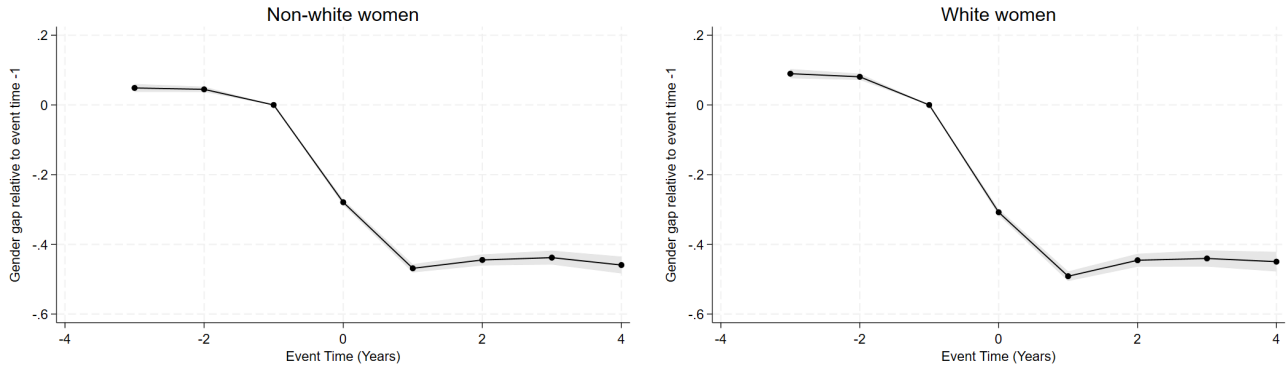
Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 977,326 families from CadUnico. The sample is split into a High and Low group using the median of the satisfied basic needs index (0.81).

4.4 Differentiated gender gaps by women’s race

We estimate the intra-household gender gaps by dividing the sample between couples with non-white and white women. Figures 8 and 9 present the results. As can be seen, there are no significant differences in formal employment and small enterprise creation gender gaps between women’s races.

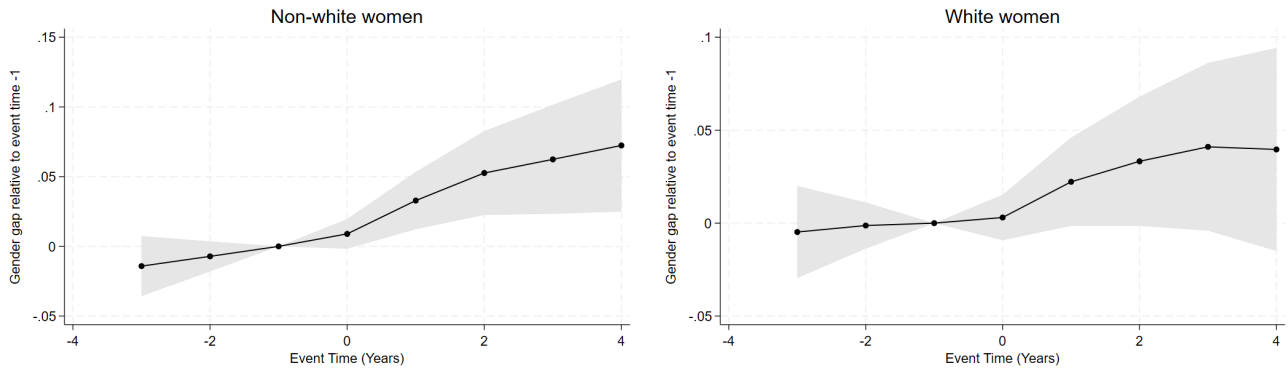
For both groups, we observe a child penalty in formal labor markets and an increase in the small enterprise creation gap in favor of women that mimic complete-sample results.

Figure 8: Child penalty on formal employment by women’s race



Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 977,326 families from CadUnico. The white category includes both white and yellow races whereas the non-white category includes indigenous, brown-skinned, and black races.

Figure 9: Intra-household gender gap on the creation of small enterprises by women’s race



Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 977,326 families from CadUnico. The white category includes both white and yellow races, whereas the non-white category includes indigenous, brown-skinned, and black races.

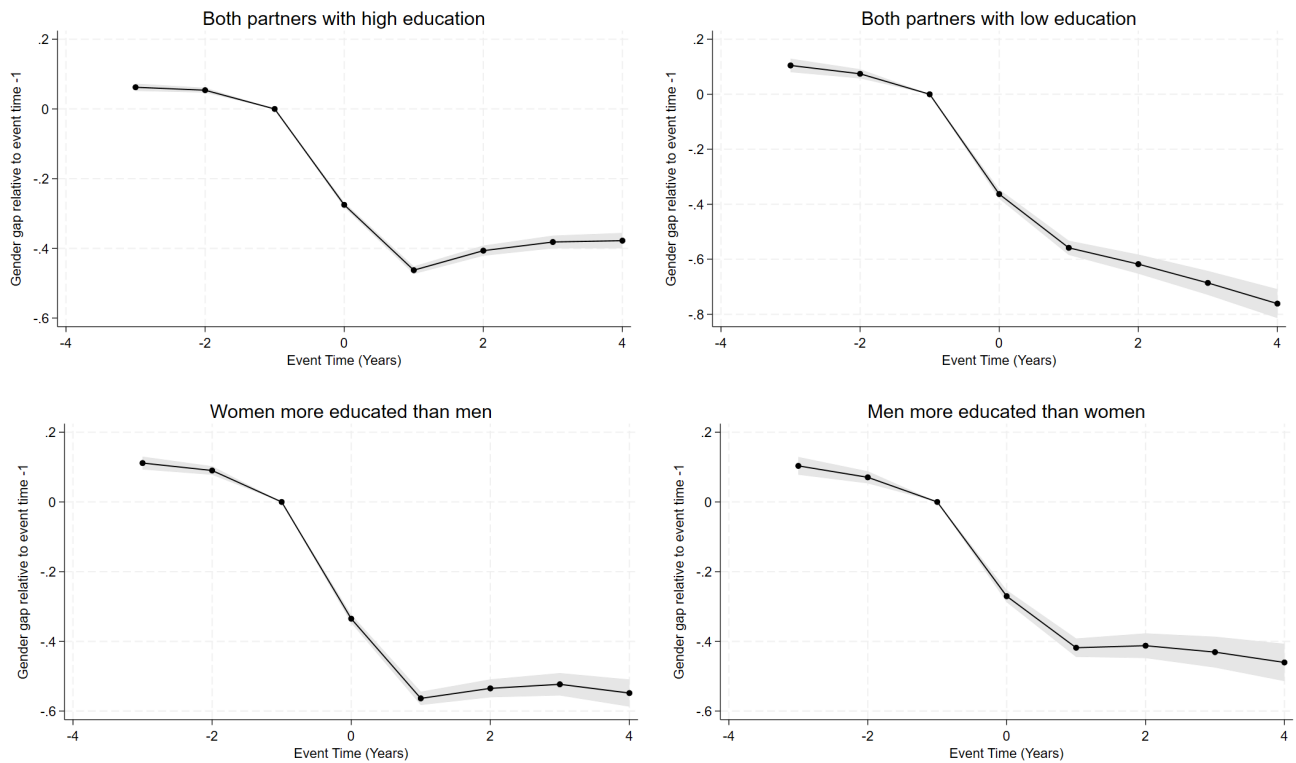
4.5 Differentiated gender gaps by couple’s education

We explore the role of a couple’s education in gender gaps after the birth of their first child. To do this, we divide the sample of men and women into two educational groups: The high-education group contains individuals with more than complete primary education. In contrast, the low-education group contains individuals with, at most, that educational level. From there, we will create four groups based on the educational groups of the man and woman in each couple. Thus, we have high-high couples, low-low couples, couples when the woman is in the high-education group and the man in the low-education group, and couples when the man is in the high-education group and the

woman in the low-education group. By categorizing couples in this way, we can determine whether more highly educated couples experience lower penalties after having children and whether couples adjust their decision to work in the formal labor market or to open a small enterprise based on their partner's relative educational advantage.

The child penalty estimates on formal employment by educational groups are displayed in Figure 10. We find that regardless of their own and their partners' educational level, women experience a penalty on formal employment after having their first child. However, penalties are more extensive for low-educated couples for whom the penalties increase over time, reaching almost 80% four years after the birth of the first child. Further, we find no evidence that couples adjust their decision to participate in the formal labor market based on the relative educational advantage of their partners. We expected that when the woman is more educated than the man, the child penalty would be lower. However, the opposite is observed: When the woman is more educated than the man, the child penalty is even greater than when the man is more educated than the woman.

Figure 10: Child penalty on formal employment by couple's education

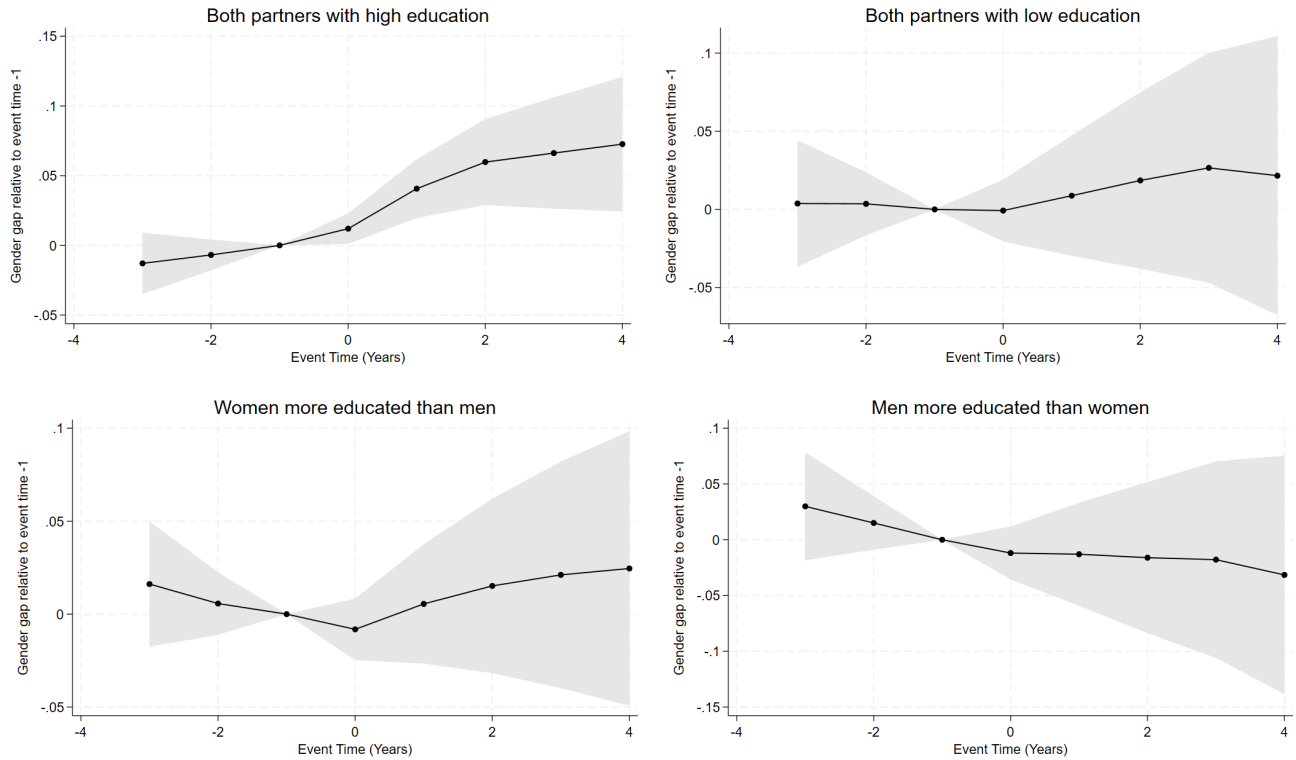


Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 1,001,764 families from CadUnico. The high-education group has secondary education or more. The low-education group at most completed primary education. In the top-left figure, both, men and women, are in the high-education group. In the top-right, both are in the low-education group. The bottom figures show the cases where men are in the high-education group and women in the low-education group and vice versa.

Figure 11 presents the estimates of the intra-household gender gap on the creation of small enterprises by educational groups. The only group where women's probability of opening a small enterprise increases more than men's after childbirth is the high-high educational group. For couples

when one or two members belong to the low-education group, i.e., have at most completed primary education, we find no significant changes in the gender gap after the birth of the first child.

Figure 11: Intra-household gender gap on the creation of small enterprises by couple’s education



Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 1,001,764 families from CadUnico. The high-education group has secondary education or more. The low-education group at most completed primary education. In the top-left figure, both, men and women, are in the high-education group. In the top-right, both are in the low-education group. The bottom figures show the cases where men are in the high-education group and women in the low-education group and vice versa.

4.6 Differentiated gender gaps by couple’s comparative employment background advantages

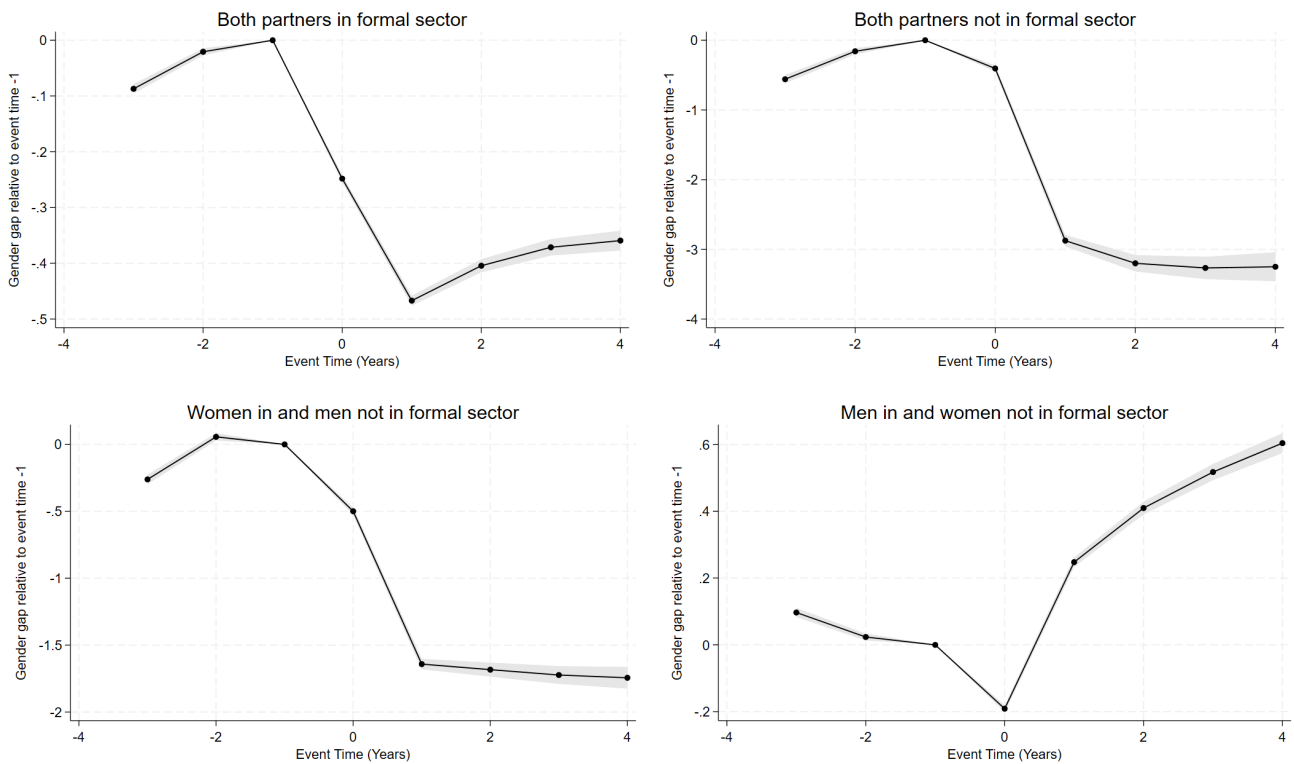
We perform a final analysis focusing on a sub-sample of our dataset that contains the employment status of each partner of a couple before the birth of their first child. This sample contains 56% of the observations. As with the educational variables, we define four groups depending on the combination of employment status of the couples. The groups are (1) partners in the formal sector, (2) partners not in the formal sector, (3) women in the formal sector and men outside the formal sector, and (4) men in the formal sector and women outside the formal sector.

The estimates of the intra-household gender gaps are detailed in Figures 12 and 13. Women’s formal employment penalty after the birth of the child is more prominent if the woman was previously in the formal labor market, regardless of the status of their partners. For women initially outside the formal labor market, the penalty is smaller if their partners are not in the formal sector.

In this case, the penalty is close to 30% four years after the birth of the child, a lower value compared to 40% and 150%, which are the results for couples with both partners in the formal sector and the man not in the formal sector, respectively.

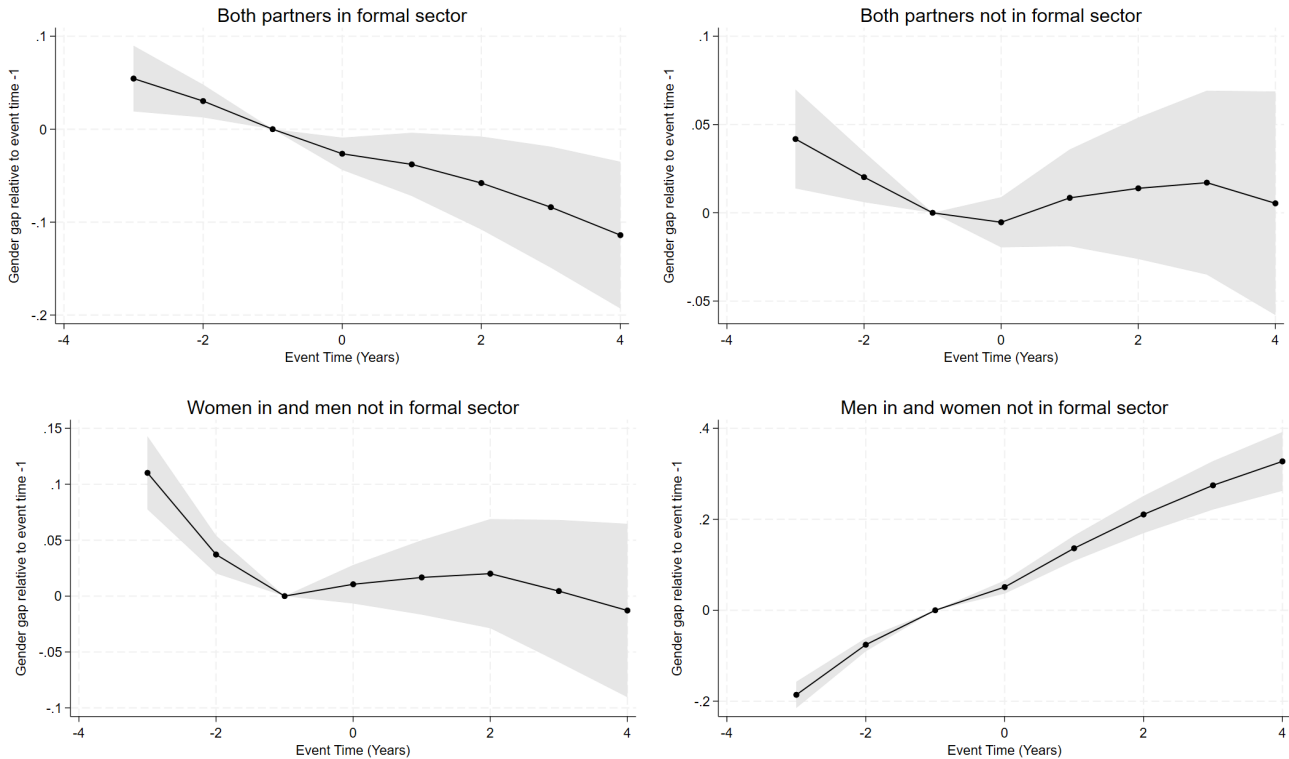
Results also suggest that women are less likely to pursue entrepreneurship if they have a formal job before giving birth. When their partners are not in the formal sector before the birth of their child, women’s chances of opening a small business remain below men’s. Further, if both partners had formal jobs before having a child, the chances revert, meaning men tend to have an advantage over women in starting a small business. However, when women are outside the formal sector, but men are in this sector before the birth of their first child, women’s chances of opening a small enterprise or even getting a formal job increase after the birth of their child (See bottom-right panel of Figures 12 and 13). This result suggests that for women outside the formal market before the birth of their child, having a partner in a formal job helps in their (re)insertion into the labor market further on.

Figure 12: Child penalty on formal employment by couple’s employment status before the child’s birth



Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 590,513 families from CadUnico (those where both members of the couple have an employment background before the child’s birth). Belong to the formal employment group those who had a labor contract, were public servants, or were in the army before the birth of their first child. The non-formal sector includes the economically active population that has no labor contract. Among them are unemployed, unpaid workers, small entrepreneurs, and informal workers. In the top-left figure, both, men and women, had formal employment before the birth of the child. In the top-right, both had no formal employment. The bottom figures show the cases where men had formal employment and women had no formal employment and vice versa.

Figure 13: Intra-household gender gap on the creation of small enterprises by couple’s employment status before the child’s birth



Note: Figure plots equation (2) at each event time, which is the gender gap before and after the birth of the first child. Gray areas represent 90% confidence intervals. All specifications include couple fixed effects. Sample size: 590,513 families from CadUnico (those where both members of the couple have an employment background before the child’s birth). Belong to the formal employment group those who had a labor contract, were public servants, or were in the army before the birth of their first child. The non-formal sector includes the economically active population that has no labor contract. Among them are unemployed, unpaid workers, small entrepreneurs, and informal workers. In the top-left figure, both, men and women, had formal employment before the birth of the child. In the top-right, both had no formal employment. The bottom figures show the cases where men had formal employment and women had no formal employment and vice versa.

5 Conclusions

The study examines the impact of having a first child on the employment outcomes of couples from low-income families in Brazil. Specifically, the research focuses on intra-household gender differences in employment and the probability of opening a small enterprise. The study utilizes public administrative records, namely CadUnico, RAIS and MEI database, to compare the changes in formal labor market participation and small enterprise ownership of couples three years before and four years after the birth of their first child.

The findings suggest that women experience a stark gender penalty in formal employment, resulting in a significant decline in their participation in the labor market by 45% four years after the birth of the child. This penalty is not related to women’s educational advantage and is more pronounced among poorer families. However, some women opt for opening small enterprises instead of leaving the labor market altogether. The probability of women opening small enterprises is 5%

higher than that of men after the birth of their child.

Women's increased participation as small entrepreneurs is higher among poorer households. However, the study also reveals that not all women can opt for small enterprise ownership due to constraints related to their education and employment background. Women are more likely to open a small enterprise in relation to men if they have more than primary education, have equally or more educated partners, and their partners are in the formal labor market before the birth of their first child.

References

- Aguilar-Gomez, S., Arceo-Gomez, E., & la Cruz Toledo, E. D. (2019). Inside the black box of child penalties. *SSRN Electronic Journal*. Retrieved from <https://www.ssrn.com/abstract=3497089> doi: 10.2139/ssrn.3497089
- Almond, D., Cheng, Y., & Machado, C. (2023, 7). Large motherhood penalties in us administrative microdata. *Proceedings of the National Academy of Sciences*, 120. Retrieved from <https://pnas.org/doi/10.1073/pnas.2209740120> doi: 10.1073/pnas.2209740120
- Angelov, N., Johansson, P., & Lindahl, E. (2016, 7). Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34, 545-579. Retrieved from <https://www.journals.uchicago.edu/doi/10.1086/684851> doi: 10.1086/684851
- Berniell, I., Berniell, L., de la Mata, D., Edo, M., & Marchionni, M. (2021, 5). Gender gaps in labor informality: The motherhood effect. *Journal of Development Economics*, 150. doi: 10.1016/j.jdeveco.2020.102599
- Berniell, I., Berniell, L., de la Mata, D., Edo, M., & Marchionni, M. (2023, 7). Motherhood and flexible jobs: Evidence from latin american countries. *World Development*, 167. doi: 10.1016/j.worlddev.2023.106225
- Berniell, I., Berniell, L., Mata, D. D. L., Edo, M., Fawaz, Y., Machado, M. P., & Marchionni, M. (2021, 6). *Motherhood and the allocation of talent*. Retrieved from <https://docs.iza.org/dp14491.pdf>
- Chiappori, P.-A. (1992, 6). Collective labor supply and welfare. *Journal of Political Economy*, 100, 437-467. Retrieved from <https://www.journals.uchicago.edu/doi/10.1086/261825> doi: 10.1086/261825
- Cortés, P., & Pan, J. (n.d.). Children and the remaining gender gaps in the labor market. *Journal of Economic Literature*, (Forthcoming). Retrieved from <https://www.aeaweb.org/articles?id=10.1257/jel.20221549>
- de Quinto, A., Hospido, L., & Sanz, C. (2021, 12). The child penalty: evidence from spain. *SERIEs*, 12, 585-606. Retrieved from <https://link.springer.com/10.1007/s13209-021-00241-9> doi: 10.1007/s13209-021-00241-9
- Goldin, C. (2014, 4). A grand gender convergence: Its last chapter. *American Economic Review*, 104, 1091-1119. Retrieved from <https://pubs.aeaweb.org/doi/10.1257/aer.104.4.1091> doi: 10.1257/aer.104.4.1091
- Kleven, H., Landais, C., & Sogaard, J. E. (2019, 10). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics*, 11, 181-209. Retrieved from <https://pubs.aeaweb.org/doi/10.1257/app.20180010> doi: 10.1257/app.20180010
- Medeiros, M. O. (2022). *Motherhood penalty in labor market: Evidence from brazil*. Retrieved from <https://www.maxwell.vrac.puc-rio.br/59164/59164.PDF>

- Molina, J. A. (2022). *Mothers in the labor market* (J. A. Molina, Ed.). Springer International Publishing. Retrieved from <https://link.springer.com/10.1007/978-3-030-99780-9> doi: 10.1007/978-3-030-99780-9
- Perova, E., & Reynolds, S. A. (2017, 2). Women's police stations and intimate partner violence: Evidence from brazil. *Social Science & Medicine*, 174, 188-196. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0277953616306773> doi: 10.1016/j.socscimed.2016.12.008
- Querejeta, M., & Bucheli, M. (2023, 2). The effect of childbirth on women's formal labour market trajectories: Evidence from uruguayan administrative data. *The Journal of Development Studies*, 59, 209-223. Retrieved from <https://www.tandfonline.com/doi/full/10.1080/00220388.2022.2128777> doi: 10.1080/00220388.2022.2128777
- Rabaté, S., & Rellstab, S. (2022, 5). What determines the child penalty in the netherlands? the role of policy and norms. *De Economist*, 170, 195-229. Retrieved from <https://link.springer.com/10.1007/s10645-022-09403-x> doi: 10.1007/s10645-022-09403-x

Appendices

A Assembling of CadUnico-RAIS-MEI database

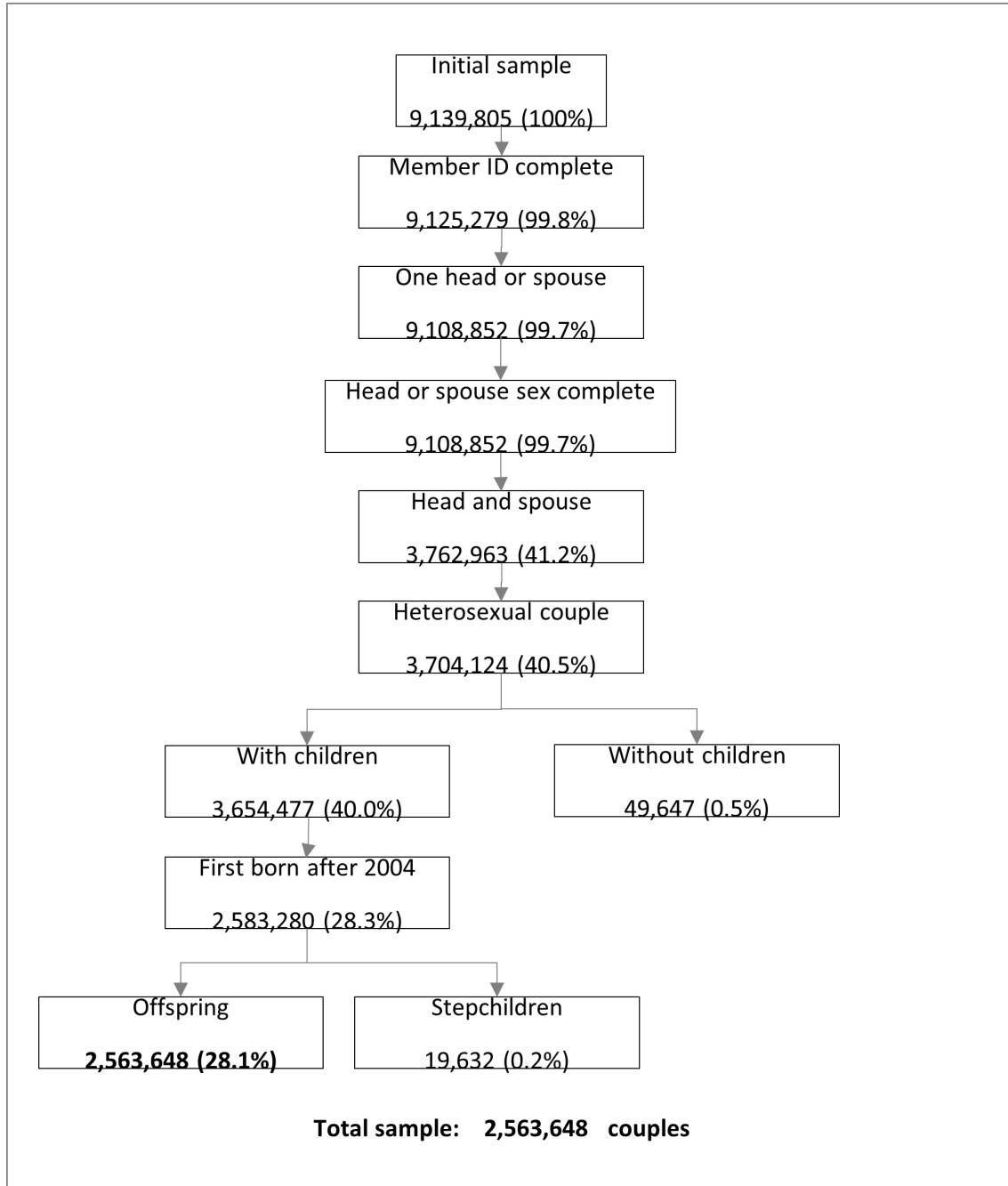
The CadUnico-RAIS-MEI database is created in four stages. First, we identify heterosexual couples from CadUnico whose first child was born in 2004 or later. Second, using the CPF number (fiscal identifier), we include the date of any individual among CadUnico couples who opened a small enterprise from fiscal administrative records of small enterprises (MEI). Third, we match CadUnico couples with the RAIS database using the NIS (social security) number. Finally, using the family code, the NIS number, and the year of information update, we add labor market and education information from CadUnico. This way, we add control variables and obtain (at least partial) labor market information from people absent from RAIS and MEI databases.

A.1 Couple master database

In this stage, we create a database containing heterosexual couples from CadUnico whose first child was born in 2004 or after. We start with a sample of families that have at least one kid with 5 years of age or less. We used CadUnico databases from 2012 to 2019. For each sample, we apply the sequence of filters described in Figure A.1, which shows 2016's selected observations. We exclude families with incomplete information about the relationship of its members, i.e., who is the family head, spouse, offspring, other relatives, and non-relatives. In addition, we eliminate families with more than one household head or spouse, and those with missing information of the sex of the head or spouse. These initial filters only reduce the database by 0.3%.

After remaining with a sample where the relevant information is complete, we proceed to eliminate families of single parents or without spouses. In 2016, 41.2% of the families in our sample had a head and spouse. We identify 0.7% of homosexual couples that we also exclude from the sample. Then, we classify the families between those with children and those without children. Among families with children, we select those that had their firstborn in 2004 or later. This is because 2004 is the first year of RAIS database available to us. Finally, we restrict the sample to couples whose firstborn is not a stepchild as they come from previous marriages. In 2016, this process left us with a sample of 2,563,648 families (couples).

Figure A.1: Sample selection of 2016's CadUnico database



Note: Decision tree tha explain the filtering process to select the sample of CadUnico couples.

We repeat the process we just described for every CadUnico database available between 2012 and 2019. The selected samples contain the family code, NIS number, sex and birth date of the household head and spouse, and the year when their first child was born, birthdate, and NIS number. We append all databases, which allows us to identify the times a family code appeared in CadUnico, and the first and last years it appeared. Since the selected variables are constant in time, for any given family code, we would expect to find the same information in every year's database. In total, we identify 5,663,233 family codes of heterosexual couples that had their firstborn in 2004 or later. However, inspecting the cases closely, we also determined that 10% of them contain

inconsistent information. For example, the NIS number of one of the spouses changes between database versions, or the children’s or parents’ birth dates change between databases. Further, we estimate the mother’s age at birth and find implausible values in 1% of the cases, i.e., mothers outside the reproductive age (15-49). We remain only with cases with consistent information over time in terms of the fields and mother’s age at birth (See Table A.1). This selection leaves us with 5,006,373 couples.

Table A.1: Case consistency of CadUnico databases over time for selected variables

Case type	N	%
Inconsistent information	589,985	10.4
Implausible mother’s age	66,875	1.2
Consistent	5,006,373	88.4
Total	5,663,233	100.0

Notes: N is number of selected observations from CadUnico. % is the percentage with respect to the total number of observations. Variables inspected for consistency were the NIS number, children and parents birth dates.

Table A.2 shows the number of databases in which a couple is for the complete selected cases and the consistent cases. Nearly 60% of the couples are in three or fewer CadUnico databases whereas couples present in all databases represent only 4.3% of the sample. This implies that the sample of couples from the master file corresponds to families that entered the registry recently, likely with few updates over time. Hence, tracking information on their living conditions using CadUnico for a long period will be difficult.

If we compare the complete sample with the sample of consistent cases, we can see the latter has a larger share of couples that are in only one or two databases and fewer couples that remain the entire period. However, the differences in the couples’ distribution between these two samples are small.

Table A.2: Couples from the master sample by the number of periods in CadUnico registry

Periods	All family codes		Consistent cases	
	N	%	N	%
1	1,197,020	21.1	1,179,468	23.6
2	1,187,289	21.0	1,106,568	22.1
3	990,452	17.5	874,603	17.5
4	825,91	14.6	697,771	13.9
5	552,3	9.8	450,556	9.0
6	436,046	7.7	341,788	6.8
7	229,55	4.1	175,623	3.5
8	244,666	4.3	179,996	3.6
Total	5,663,233	100.0	5,006,373	100.0

Notes: N is number of selected observations from CadUnico. % is the percentage with respect to the total number of observations.

Finally, we add the CPF number. Using the NIS number from the couples' database, we search for this information in the raw CadUnico's databases and select the most recent entry. As CPF is a unique identifier, we would expect to be the same regardless the year of the registry and the family code (for people switching families or starting a new one). We find this is true for 97.7% of the sample. However, we find inconsistencies, i.e., different numbers being reported, in 0.3% of the cases. Additionally, we find that 2.1% of the individuals in the sample do not have a CPF number. Details are in Table A.3.

Table A.3: Consistency and availability of CPF number

	Men		Women		Total	
	Freq.	%	Freq.	%	Freq.	%
Consistent	4,843,488	96.8	4,934,053	98.6	9,777,541	97.7
Inconsistent	147,43	2.9	61,507	1.2	208,937	2.1
No number	15,455	0.3	10,813	0.2	26,268	0.3
Total	5,006,373	100.0	5,006,373	100.0	10,012,746	100.0

Notes: Freq. is the number of selected observations from CadUnico. % is the percentage with respect to the total number of observations.

A.2 Master-MEI merge

The MEI database contains micro enterprises' information from fiscal administrative records. The database registers the creation date of these small businesses, an address, and an activity code. Between 2010 and 2022, 20,401,282 people registered the creation of a small business; 18,770,262 (92%) reported only once, 1,488,445 (7.3%) reported twice, and 142,574 (0.7%) more than twice. People with more than one entry in the database have more than one creation date, address, or activity. This could be because the business closed and reopened later, they have another branch, or created another business, defined by another activity code. The 0.7% of people with more than two entries include atypical cases with more than 100 entries, which could be simply attributed to typing errors in the CPF number. To avoid introducing unnecessary noise to the estimation, we eliminated these observations, remaining with 20,258,707 (99.3%) cases. Among those with two

entries, we included the information of each entry separately, prioritizing the earliest entry. Thus, when analyzing this information in terms of business creation after having a child, we interpret our results as the effect of creating the first micro-enterprise. Table A.4 contains the number of businesses opened by year.

Table A.4: New micro-enterprises by year

Year	Frequency	Percentage
2010	678,954	3.35
2011	850,138	4.20
2012	978,936	4.83
2013	1,158,957	5.72
2014	1,235,509	6.10
2015	1,358,318	6.70
2016	1,386,972	6.85
2017	1,546,064	7.63
2018	1,750,036	8.64
2019	2,158,837	10.66
2020	2,321,938	11.46
2021	2,742,420	13.54
2022	2,091,628	10.32
Total	20,258,707	100.00

We use the CPF number to match the couples' master file with the MEI database. We find that 13.5% of the observations opened a micro-enterprise between 2010 and 2022. This percentage is smaller for women (10.9%) than for men (16.2%). See Table A.5 for details.

Table A.5: Opened a micro-enterprise

	Men		Women		Total	
	Freq.	%	Freq.	%	Freq.	%
Yes	809,581	16.2	542,952	10.9	1,352,533	13.5
No	4,196,792	83.8	4,463,421	89.1	8,660,213	86.5
Total	5,006,373	100.0	5,006,373	100.0	10,012,746	100.0

Notes: Freq. is the number of selected observations from CadUnico. % is the percentage with respect to the total number of observations.

Table A.6 shows the percentage of individuals in the CadUnico sample opening micro-enterprises by year. The distribution is similar to the MEI database's, but in the initial years, we observe a small gender gap in favor of men. This gap closes by 2014 and reverts later.

Table A.6: Opened a micro-enterprise

Year	Men		Women		Total	
	Freq.	%	Freq.	%	Freq.	%
2010	23,432	2.89	11,688	2.15	35,12	2.60
2011	32,241	3.98	18,123	3.34	50,364	3.72
2012	39,095	4.83	23,16	4.27	62,255	4.60
2013	48,54	6.00	31,526	5.81	80,066	5.92
2014	53,421	6.60	36,254	6.68	89,675	6.63
2015	57,613	7.12	40,277	7.42	97,89	7.24
2016	57,457	7.10	39,516	7.28	96,973	7.17
2017	64,507	7.97	43,179	7.95	107,686	7.96
2018	74,432	9.19	49,679	9.15	124,111	9.18
2019	91,041	11.25	61,781	11.38	152,822	11.30
2020	90,193	11.14	60,158	11.08	150,351	11.12
2021	100,68	12.44	75,071	13.83	175,751	12.99
2022	76,929	9.50	52,54	9.68	129,469	9.57
Total	809,581	100.00	542,952	100.00	1,352,533	100.00

Notes: Freq. is the number of selected observations from CadUnico. % is the percentage with respect to the total number of observations.

A.3 Master-RAIS merge

Next, we merge the master file of CadUnico's couples with a pre-selected and annualized set of labor market variables from RAIS. The main variable is formal employment, defined as having a labor contract at the end of the year. In addition, we included information on the hours worked per week, the tenure, and the contracted wage.

We create an annual balanced panel structure using the master file of CadUnico's couples, i.e., we expand each individual observation 16 times, which represents annual entries from 2004 to 2019. The database contains 160,203,936 entries. Using this extended database, we merge the annual information from RAIS databases. The matches in the database represent 27% of all people from CadUnico master file, however, this percentage varies depending on the year. The performance of the match in the first years, 2004 to 2008, is the worst. One explanation for this outcome is the absence of the youngest people, who enter the formal labor market later in their lives. Likewise, the percentage of women in the matched sample is only 31% for the entire period, and is even smaller in the first years, which reflects their lower participation and their delayed entrance to the formal market (See Table A.7).

Table A.7: Characteristics of the matched sample by year

Year	Complete Database		Matched Sample		
	% women	Average age	% matched	% women	Average age
2004	50	17.77	13	24	24.90
2005	50	18.77	15	25	25.27
2006	50	19.77	17	25	25.45
2007	50	20.77	20	26	25.81
2008	50	21.77	23	27	26.14
2009	50	22.77	24	28	26.62
2010	50	23.77	27	29	27.01
2011	50	24.77	32	30	27.19
2012	50	25.77	33	31	27.71
2013	50	26.77	36	32	28.21
2014	50	27.77	36	33	28.86
2015	50	28.77	34	33	29.72
2016	50	29.77	32	33	30.67
2017	50	30.77	31	34	31.57
2018	50	31.76	31	35	32.41
2019	50	32.76	31	35	33.20
Total	50	25.27	27	31	28.66

A.4 Master-CadUnico merge

We complement the labor market information and other socioeconomic characteristics using the CadUnico database. RAIS registries only possess information from the formal labor market. By using CadUnico, we can identify a wider range of socioeconomic characteristics and their participation in the informal labor market. Families in this database are beneficiaries of any social program in Brazil. It contains families' socioeconomic information of the moment they entered any program and it is updated once every three years as long as the family is a beneficiary of those programs. Thus, in a given year, we cannot know the situation of all families, only of those that updated their information that year. Using the panel data of CU couples from the previous section, we identify the percentage of individuals with updated information by year through a matching process.

The information of any family and its members should be updated once per year at most. However, in some cases, we found more than one registry per year, which represented 1.68% of the master sample. In addition, after appending all available datasets from 2012 and 2019, we found duplicated information, which is expected because the database is a registry that accumulates beneficiaries over time. Hence, even if the family information is not updated, the case will appear in every version of the database. For any duplicate, we selected the most recent case, which would probably contain the information of the highest quality because the registry improved over time.

Considering the above-mentioned criteria, we were able to match 30% of the master data cases. We matched less than 10 percent of the sample between 2004 and 2009. Since our first database is from 2012, this result indicates that only a very small share of the sample remained in the registry without any information update for more than 3 years. Further, the percentage of families with

an updated registry increased over time. From 2015 onward, more than 50% of the sample had a registry up to date (Table A.8). Therefore, for the event study, we will focus on cohorts of couples that had their firstborn in more recent years.

Additionally, Table A.8 shows the percentage of women and average age of the matched sample. Both variables are close to the values of the complete sample, particularly in the last five years.

Table A.8: Characteristics of the CadUnico matched sample by year

Year	% matched	% women	Average age
2004	1	53	15.93
2005	1	52	16.94
2006	4	52	18.36
2007	3	53	19.95
2008	4	53	20.99
2009	8	53	22.19
2010	12	53	23.65
2011	18	54	24.57
2012	38	54	25.39
2013	45	53	26.57
2014	46	53	27.51
2015	51	52	28.52
2016	54	52	29.45
2017	59	51	30.36
2018	83	51	31.47
2019	57	51	31.94
Total	30	52	28.62

A.5 Complete cases

Kleven et al. (2019) restrict their analysis to observations that have the same time span in the sample, regardless of the period when they are observed. For example, suppose we want to follow the sample three years before the birth of the first child and three years after. If the child was born in 2012, we would have to select the couple’s information from 2010 to 2015. But, if the child was born in 2015, the period of information would be from 2012 to 2018. The number of individuals (parents) by birth cohort and the available years before and after the birth of the first child are shown in Figure A.2. The lines show the time to birth in years: Zero represents the moment the child was born; negative numbers indicate the years before birth, and positive numbers indicate the years after. On the other hand, the columns detail the birth cohorts available in the database, from 2004 to 2019.

We created two sub-samples for the event study. We select 2008 to 2014 birth cohorts for the first sample, which allows us to follow couples in the labor market four years before and five years after the birth of their first child. The red cells in Figure A.2, delimited by the solid line, show the number of observations per cohort. The total number of observations is 4,955,774 per period, approximately 708,000 per cohort and period. This first sample will allow us to analyze the effects of the birth of the firstborn on formal employment. However, the results presented in this version

of the paper only focus on the second sample.

The second sample contains information from the 2013, 2014, and 2015 cohorts, restricted to three years before and four years after the birth of the first child. This sample will use information from CadUnico and MEI databases to identify the labor situation of couples outside the formal labor market. As shown in Tables A.6 and A.8, no MEI information is available before 2010, and information from CadUnico is very scarce prior to that year too. This is why we use information from cohorts that had their child after 2012 so that we can observe them at least 3 years before the birth of the child. To measure the effects of the firstborn up to four years later, we restrict the sample to observations from the 2015 cohort or before. The selected sample is highlighted in orange, delimited with dashed lines in Figure A.2. The total number of observations is 2,108,976 per year (1,054,488 couples), approximately 703,000 per cohort and period.

Figure A.2: Panel observations by event time and cohort

Time to birth	Birth Cohort																
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	
-15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	223446	
-14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	422100	223446	
-13	0	0	0	0	0	0	0	0	0	0	0	0	0	513222	422100	223446	
-12	0	0	0	0	0	0	0	0	0	0	0	0	575544	513222	422100	223446	
-11	0	0	0	0	0	0	0	0	0	0	0	666014	575544	513222	422100	223446	
-10	0	0	0	0	0	0	0	0	0	0	710844	666014	575544	513222	422100	223446	
-9	0	0	0	0	0	0	0	0	0	732118	710844	666014	575544	513222	422100	223446	
-8	0	0	0	0	0	0	0	0	730584	732118	710844	666014	575544	513222	422100	223446	
-7	0	0	0	0	0	0	0	716776	730584	732118	710844	666014	575544	513222	422100	223446	
-6	0	0	0	0	0	0	688792	716776	730584	732118	710844	666014	575544	513222	422100	223446	
-5	0	0	0	0	0	686226	688792	716776	730584	732118	710844	666014	575544	513222	422100	223446	
-4	0	0	0	0	690434	686226	688792	716776	730584	732118	710844	666014	575544	513222	422100	223446	
-3	0	0	0	679232	690434	686226	688792	716776	730584	732118	710844	666014	575544	513222	422100	223446	
-2	0	0	693228	679232	690434	686226	688792	716776	730584	732118	710844	666014	575544	513222	422100	223446	
-1	0	659724	693228	679232	690434	686226	688792	716776	730584	732118	710844	666014	575544	513222	422100	223446	
0	624462	659724	693228	679232	690434	686226	688792	716776	730584	732118	710844	666014	575544	513222	422100	223446	
1	624462	659724	693228	679232	690434	686226	688792	716776	730584	732118	710844	666014	575544	513222	422100	0	
2	624462	659724	693228	679232	690434	686226	688792	716776	730584	732118	710844	666014	575544	513222	0	0	
3	624462	659724	693228	679232	690434	686226	688792	716776	730584	732118	710844	666014	575544	0	0	0	
4	624462	659724	693228	679232	690434	686226	688792	716776	730584	732118	710844	666014	0	0	0	0	
5	624462	659724	693228	679232	690434	686226	688792	716776	730584	732118	710844	0	0	0	0	0	
6	624462	659724	693228	679232	690434	686226	688792	716776	730584	732118	0	0	0	0	0	0	
7	624462	659724	693228	679232	690434	686226	688792	716776	730584	0	0	0	0	0	0	0	
8	624462	659724	693228	679232	690434	686226	688792	716776	0	0	0	0	0	0	0	0	
9	624462	659724	693228	679232	690434	686226	688792	0	0	0	0	0	0	0	0	0	
10	624462	659724	693228	679232	690434	686226	0	0	0	0	0	0	0	0	0	0	
11	624462	659724	693228	679232	690434	0	0	0	0	0	0	0	0	0	0	0	
12	624462	659724	693228	679232	0	0	0	0	0	0	0	0	0	0	0	0	
13	624462	659724	693228	0	0	0	0	0	0	0	0	0	0	0	0	0	
14	624462	659724	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
15	624462	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Total	9991392	10555584	11091648	10867712	11046944	10979616	11020672	11468416	11689344	11713888	11373504	10656224	9208704	8211552	6753600	3575136	

Finally, as stated in the previous sub-section, socioeconomic characteristics from CadUnico are not updated yearly, thus it is not possible to use these variables as controls in the event study. However, it is possible to perform split-sample analyses to identify potential heterogeneity in the child penalty estimates. To do that, we either require variables that are stable or fixed in time, or at least measured before the birth of the child. Thus, variables in the split-sample analysis would be taken as pre-determined before the birth of the first child.

We analyzed how constant the following variables were: Education, family income, employment, family composition, municipality of residence, and dwelling characteristics. The variables were

selected based on the literature on factors that contribute to explaining gender gaps Cortés and Pan (n.d.). Using the second sample, we found that, on average, the CadUnico registry reports information from at most four of the eight periods, with the most common value being nearly 2.5 periods out of eight. Family income and employment changed more frequently than the rest of the variables, which at most changed once over time. Hence, to conduct the split-sample analysis, we assumed all variables except family income and employment remained constant over time and used the most frequent value. As employment is also the outcome of interest, we took the value from before the birth of the child for the split-sample analysis. Because the information before the birth of the child is scarcer, using this variable implied losing nearly 30% of the sample.

Table A.9: Socioeconomic information’s availability and update frequency

	Missing (%)	Times information was reported		
		Updated	Changed	Changed before first child
Education	8.76	2.45	0.98	0.38
Family income	8.71	2.46	2.35	0.50
Employment	2.99	4.30	1.46	0.82
Family composition	8.84	2.07	0.94	0.28
Municipal code	8.71	2.46	1.08	0.41
Dwelling characteristics	8.71	2.46	1.11	0.39
Observations	2108976	2108976	2108976	2108976

Notes: Column “Missing” reports observations without any information in CadUnico. Column “Updated” shows the average times any information was reported over the period of study. Column “Changed” contains the number of different values reported in CadUnico over the period of study. Column “Changed before first child” also displays the number of different values reported in CadUnico but only before the birth of the first child.