DOCUMENTOS DE TRABAJO

The Long-Run Effects of Conditional Cash Transfers: the Case of Bolsa Familia in Brazil

Luis Laguinge, Leonardo Gasparini y Guido Neidhöfer

Documento de Trabajo Nro. 328
Abril, 2024
ISSN 1853-0168
www.cedlas.econo.unlp.edu.ar

Cita sugerida: Laguinge, L., L. Gasparini y G. Neidhöfer (2024). The Long-Run Effects of Conditional Cash Transfers: the Case of Bolsa Familia in Brazil. Documentos de Trabajo del CEDLAS Nº 328, Abril, 2024, CEDLAS-Universidad Nacional de La Plata.

The Long-Run Effects of Conditional Cash Transfers: the Case of *Bolsa Familia* in Brazil *

Luis Laguinge† Leonardo Gasparini Guido Neidhöfer

March, 2024

Abstract

Conditional Cash Transfers (CCTs) have become a key antipoverty policy in Latin America in the last 25 years. The ultimate goal of this kind of programs is to break the intergenerational transmission of poverty through the promotion of human capital accumulation of children in vulnerable households. In this paper, we explore this issue by estimating the long-run effects of the largest CCT in Latin America: the Brazilian *Bolsa Familia*. Through a combination of the two-stage-two-sample method and a difference-in-differences approach, we find evidence consistent with a positive long-run impact of *Bolsa Familia* among former beneficiaries. In particular, we find a significant positive effect on education and labor income, and a negative effect on the likelihood of being a current beneficiary of this social transfer.

JEL Classification: D04, I38, J24.

Keywords: Conditional cash transfers, long term effects, human capital formation, *Bolsa Familia*, Brazil, Latin America.

*We thank Matías Ciaschi, Inés Berniell, Carlo Lombardo, participants at CEDLAS internal seminars and members of the UNLP Doctoral Committee for their useful comments and suggestions. Guido Neidhöfer thanks the Jacobs Foundation (JF Research Fellowship) for funding. The usual disclaimer applies.

†Luis Laguinge, CEDLAS-IIE-UNLP & CONICET; luislaguinge4@gmail.com. Leonardo Gasparini, CEDLAS-IIE-UNLP & CONICET; leonardo.gasparini@econo.unlp.edu.ar. Guido Neidhöfer, ZEW Mannheim & Turkish-German University guido.neidhoefer@zew.de.

1 Introduction

Conditional Cash Transfers (CCT) have become a fundamental tool for alleviating poverty in Latin America, both in the short and long term (Millán et al., 2019). While the first goal is addressed through a cash subsidy with an immediate effect on real incomes, the second (and more ambitious) one relies on conditionality on schooling to promote the human capital accumulation of children and, consequently, contribute to breaking the cycle of intergenerational transmission of poverty (Garcia and Saavedra, 2023). If CCTs effectively assist poor households in overcoming barriers that hinder their access to education and human capital formation, the next generation would be less likely to be poor and dependent on government assistance.

There is a large literature that has analyzed the short-term effects of CCT programs on several outcomes such as household consumption, income poverty, labor market participation, informality, school enrollment and access to preventive health services (Fiszbein and Schady, 2009). However, the long-term effects of these policies have been studied less so far. This is understandable, since CCTs are relatively new policies, and the data requirements to study long-term effects are more demanding. This paper contributes to the literature by analyzing the impact of the Brazilian *Bolsa Familia* CCT on human capital and labor market outcomes. *Bolsa Familia* (BF) was launched by the federal government in October 2003 and rapidly reached one quarter of the Brazilian population. It has the typical features of a CCT: a cash subsidy to poor households with children under 18 years old, which is conditional on children's school attendance, as well as the compliance with an immunization schedule and medical check-ups. BF's predecessor, *Bolsa Escola*, is considered a pioneer of this family of transfers along with PROGRESA in Mexico and the Female Secondary School Stipend Program in Bangladesh.

Ideally, studying the long-term effects would require longitudinal data; specifically, information on whether adults accessed (or not) CCTs when they were children. However, this type of information is usually not available in administrative data and national household surveys. Hence, the evidence has to be obtained from alternative, indirect sources. Our identification strategy follows a two-step procedure, which builds upon the one adopted by Neidhöfer and Niño-Zarazúa (2019) to estimate the long-run effects of the *Chile Solidario* program.

We implement a two-stage-two-sample method (TSTS) in order to predict the probability that adult individuals received the transfer during childhood. First, we estimate the likelihood of households with certain characteristics to be BF beneficiaries using data from the 2006 *Pesquisa Nacional por Amostra de Domicílios* (PNAD), the official national household survey of Brazil. Then, we take advantage of the rich module of retrospective questions of the PNAD 2014 and compute for each individual the likelihood that their household of origin was beneficiary of the program in the past. Finally, we exploit two sources of exogenous variation to estimate the impact of the program on education and labor market outcomes through a difference-in-differences approach: First, the difference in the estimated likelihood of having been a beneficiary of BF in childhood allows us to divide the sample into a treatment and a control group. Second, we exploit the age restriction imposed by the program by adopting an age-cohort approach.

Our estimates suggest a significant and positive effect of *Bolsa Familia* on years of schooling (around 0.8 years), and on the probability of having completed primary and secondary education among former beneficiaries. We also find a significant positive effect on monthly labor income (around US$250), and a negative effect on the probability of being a current beneficiary of this social transfer. It is worth noting that the related literature has reached a certain consensus regarding the positive effect of CCTs on variables such as years of schooling and primary and secondary school completion (Neidhöfer and Niño-Zarazúa, 2019 for *Chile Solidario* in Chile; Behrman et al., 2011 and Parker and Vogl, 2023 for PROGRESA/*Oportunidades* in Mexico; Araujo et al., 2019 for *Bono de Desarrollo Humano* (BDH) in Ecuador; Baez and Camacho, 2011, Duque et al., 2019 and Attanasio et al., 2021 for *Familias en Acción* in Colombia; Barham et al., 2017 for *Red de Protección Social* (RPS) in Nicaragua, Ham and Michelson, 2018 and Millán et al., 2020 for *Programa de Asignación Familiar-II* (PRAF-II) in Honduras; Gaentzsch, 2020 for *Juntos* in Peru, Sanchez Chico et al., 2018 for *Comunidades Solidarias Rurales* in El Salvador; Alam et al., 2011 for the Punjab Female School Stipend Program (FSTP) in Pakistan; Filmer and Schady, 2014 for the CESSP Scholarship Program in Cambodia and Baird et al., 2019 for a program that targeted adolescent girls in Malawi). However, the evidence regarding labor market outcome is so far ambiguous. While some studies find positive effects on labor participation, employment and income (Barham et al., 2017; Barham et al., 2018; Behrman et al., 2011; Neidhöfer and Niño-Zarazúa, 2019 and Parker and Vogl, 2023), others find no significant impact (Araujo et al., 2019; Baird et al., 2019; Millán et al., 2020; Filmer and Schady, 2014). One interesting and relevant reference is Bailey et al. (2023), where the long-term effects of the U.S. Food Stamps program are evaluated. The authors find positive effects on income, labor participation, and college graduation, and negative effects on adult poverty and receipt of public benefits.

Regarding the positive effects on labor outcomes, the literature explores mechanisms such as migration from semi-rural areas to urban areas, the fall in average reproductive age and fertility of women, and sectoral reallocation. The latter is a channel that we also evaluate in this paper and for which we find evidence consistent with the transition from low productivity to higher productivity jobs. Hence, our findings contribute to the literature on the effects of CCTs by providing first evidence on the long-term effects of one of the pioneering programs worldwide and shedding light on the mechanisms that connect the accumulation of human capital with the performance of individuals in the labor market.

The rest of the paper is organized as follows: Section 2 briefly describes the functioning of the Bolsa Familia program, the context of its implementation, its scope and its importance in terms of purchasing power. Section 3 details the data and the methodology used to identify beneficiaries and to estimate the causal effect of the transfer in the long run. Section 4 presents the main results and some of placebo and robustness tests. Section 5 explores some channels that could explain how human capital accumulation allows individuals to increase their labor income. Section 6 concludes.

2 A brief history of *Bolsa Familia*

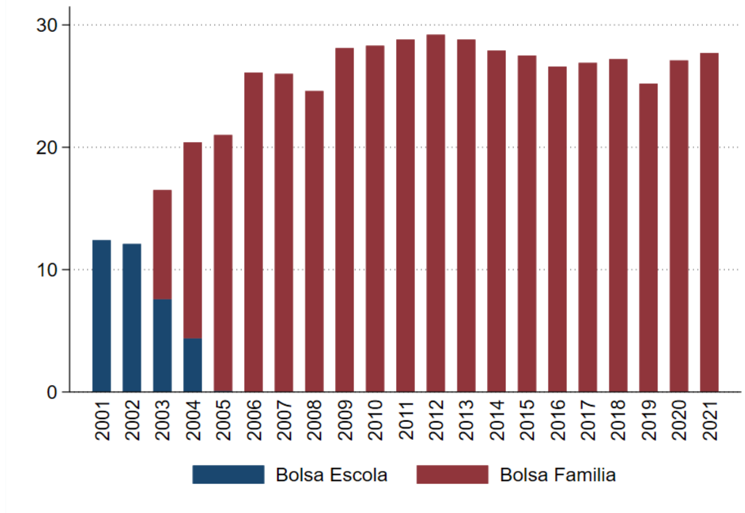
Bolsa Familia (BF) is a CCT program created by the Brazilian federal government in October 2003, which unified several initiatives aimed at the poorest households (*Bolsa Escola*, *Acesso à Alimentação*, *Bolsa Alimentação*, *Auxílio-Gás* and *Cadastramento Único do Governo Federal*). The program is managed by the Ministry of Development and Social Assistance, Family and Fight against Hunger (MDS), which coordinates the enrollment of families in the Federal Government Single Registry of Social Programs (*Cadastro Único*) with municipal governments. MDS defines BF's annual budget and municipal quotas based on micro-area poverty estimates elaborated by the national institute of statistics (IBGE). Once enrolled in the Single Registry, families must answer a questionnaire which allows municipal governments to collect information such as household composition, access to basic services, schooling of each member, labor market status and self-declared income. This last variable is crucial since it determines whether the household is poor or not according to the current thresholds and, consequently, eligible to receive the transfer.

Initially, the program consisted of two type of benefits: the Basic Benefit (a lump sum transfer aimed at extreme poor households) and the Variable Benefit (aimed at extreme/moderate poor households with children up to 15 years old). The payment of the benefits was conditional on school attendance and medical check-ups. In December 2007, the program expanded its coverage to include teenagers aged 16 and 17 through the Variable Youth Benefit and established ceilings for the amount of benefits per household (three for the VB and two for the VYB). The cap for the VB was raised to five in June 2011. It's worth noting that the *Bolsa Escola* (BE) program was in force since March 2001 and had the same target population and conditionalities as those set up by *Bolsa Familia*. Therefore, both can be considered as one and the same program in practice.

Figure 1 shows the evolution of beneficiaries of BE-BF as a percentage of the total population. The coverage of both programs increased from around 10% in 2001 to more than 25% in 2006 and remained at that level until 2021. This share represents about 50 million people, a figure that is illustrative of the size of the program.

The growing importance of BF can be seen not only in terms of its coverage but also in the evolution of the purchasing power of the transfer. Figure 2a shows the evolution of the real mean value of the transfer measured in current U.S. dollars. BF reaches a maximum of almost US$70 in 2014 (an increase of 172% over 2003) and collapses to around US$40 in 2021. The variation over the entire period is about 56%. Figure 2b shows that the mean value of BF remained stable between 2005 and 2021 at around 20% of the minimum wage (which increased 30% in real terms during this period). In summary, since its introduction, BF has experienced extraordinary growth in both its coverage and purchasing power of the benefit.

Figure 1: Coverage of *Bolsa Escola* and *Bolsa Familia* - % of total population (2001-2021)



Own elaboration based on CEPALSTAT

Figure 2: Mean value of *Bolsa Familia* (2003-2021)

(a) Current US$

(b) % of minimum wage

Own elaboration based on open data from Brazilian Federal Government, Brazilian Central Bank and *Instituto de Pesquisa Econômica Aplicada* (IPEA). Note: values correspond to December of each year.

3 Identification Strategy and Data

In the absence of longitudinal data or direct information on program participation in childhood of adult individuals, our methodology consists of two steps. In the first step, we estimate the likelihood of adults to have been program beneficiaries during childhood, whereas in the second step we estimate the long-run effects of the program. The first step is an application of the so called two-stage-two-sample method that has been widely used by the literature on intergenerational income mobility (Björklund and Jäntti, 1997).[1]

3.1 First stage: estimating the probability of participation in the program

In the first stage of our methodology, we use data from the *Pesquisa Nacional por Amostra de Domicílios* (PNAD) conducted in the year 2006 and representative of the whole population of Brazil, in order to estimate the characteristics associated with the probability of being eligible for *Bolsa Familia*. Table 1 shows descriptive statistics of households with children comparing beneficiaries and non-beneficiaries. Beneficiaries of *Bolsa Familia* have a substantially lower education than non-beneficiaries; fathers are over-represented among informal salaried workers and self-employed, while almost one fifth of mothers work in own-production or as non-salaried workers; more than half reside in the northeast region and around one third in rural areas.

Table 1: Characteristics of households with children (%)
Beneficiaries (B) and non-beneficiaries (NB) of *Bolsa Familia*

	NB	B	Total		NB	B	Total
Father and mother present	83.3	81.6	82.9				
Parents' educational level				**Mother's occupational category**			
None	3.2	11.4	5.2	Formal salaried worker	20.9	7.8	17.6
Incomplete primary	30.0	59.9	37.4	Informal salaried worker	13.9	17.6	14.8
Complete primary	11.3	9.4	10.8	Public salaried worker/Serviceman	7.0	2.0	5.7
Incomplete secondary	7.2	5.9	6.9	Self-employed	10.9	11.4	11.0
Complete secondary	29.9	12.4	25.6	Employer	2.6	0.2	2.0
Incomplete tertiary	6.2	0.8	4.8	Other	6.4	19.2	9.6
Complete tertiary	12.3	0.3	9.3	Unemployed/Inactive	38.4	41.8	39.2
Father's occupational category				**Region**			
Formal salaried worker	40.0	23.4	36.0	North	7.8	9.4	8.2
Informal salaried worker	13.2	23.8	15.8	Northeast	19.7	51.8	27.6
Public salaried worker/Serviceman	6.5	1.9	5.4	Southeast	46.7	25.3	41.4
Self-employed	23.3	36.5	26.5	South	17.5	8.8	15.3
Employer	7.9	2.1	6.5	Central-West	8.3	4.7	7.4
Other	1.0	3.1	1.5				
Unemployed/Inactive	8.1	9.2	8.3	**Urban location**	87.7	66.8	82.5
Population (in millions)	20.8	6.9	27.7	**Population (in millions)**	20.8	6.9	27.7

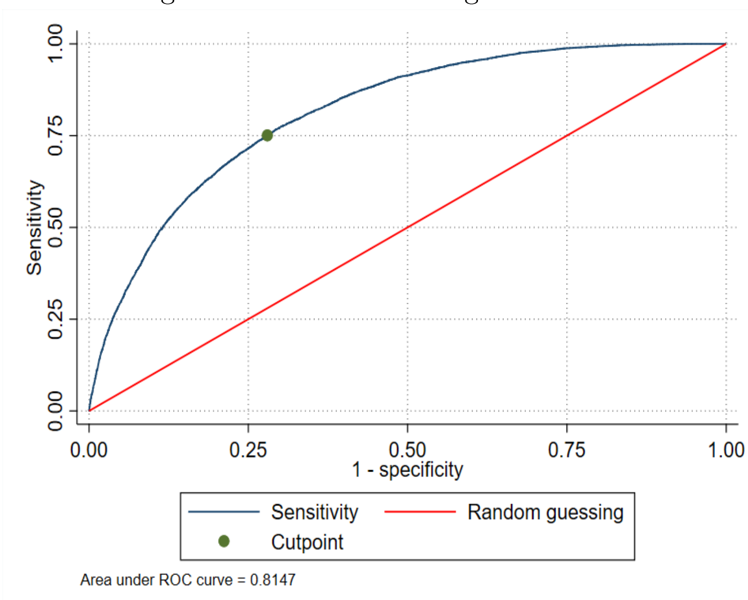
Own elaboration based on PNAD-2006

[1] For a review of this literature for Latin America, see Brunori et al., 2023.

Since in the second step of the procedure we use information on both parents, we restrict the sample to households with children where both parents are present. Within that sample, we estimate a Logit model where the dependent variable is a dummy that indicates whether a household received the benefit or not. The group of covariates includes the maximum level of education attained by parents (considering the one with the highest educational level), the occupational category of fathers and mothers, the state of residency and the location area (rural or urban). The results of this estimation are shown in Table A.2.

In order to asses the predictive performance of the model, we use the Nearest to (0,1) method that finds the cutpoint on the Receiver Operating Characteristic (ROC) curve closest to (0,1) (Liu, 2012). This curve plots the true-positive as a fraction of actual beneficiaries (sensitivity) against the false-positive as a fraction of non-beneficiaries (1-specificity). Therefore, the point (0,1) is the one associated with perfect prediction. Following this procedure, we find that with the included characteristics 75% of actual beneficiaries are correctly identified when using a cutpoint equal to 0.257 while the false-positive rate reaches a percentage of 28%. Figure 3 shows the ROC curve and the cutpoint.

Figure 3: ROC curve - Logit estimates



Given the importance of identifying potential beneficiaries as accurately as possible for the credibility of the identification strategy, we perform a number of robustness exercises based on the application of three machine learning algorithms in order to predict the probability of participation in the program. We consider LASSO, Ridge and Elastic net (each of them estimating a linear and a Logit model).[2] Table 2 shows the cutpoint, the area under the ROC curve, the True Positive Rate (TPR) and the False Positive Rate (FPR) for each algorithm. We can note that there are no major differences in terms of performance. The AUC-ROC is almost identical for all the algorithms reflecting a similar ability to distinguish between classes. Regarding TPR and FPR, the maximum difference between models does not exceed 3 percentage points. Lastly, the cutpoints chosen by the algorithms do not differ much either. While the Logit models suggest a value close to 0.25, the linear ones lean towards a threshold around 0.3. All in all, we believe that this evidence strengthens the hypothesis that the threshold is robust to the use of different classification algorithms.

Table 2: Cutpoint and performance metrics for alternative algorithms

	Cutpoint	AUC-ROC	TPR	FPR
Logit	0.2574	0.74	0.75	0.28
LASSO (linear)	0.2972	0.73	0.73	0.26
LASSO (Logit)	0.2592	0.74	0.73	0.27
Ridge (linear)	0.2960	0.73	0.73	0.27
Ridge (Logit)	0.2479	0.74	0.76	0.29
Elastic net (linear)	0.2966	0.73	0.73	0.27
Elastic net (Logit)	0.2449	0.74	0.76	0.29

Note: AUC-ROC: area ROC curve at cutpoint. TPR: True positive rate. FPR: False Positive rate.

3.2 Second stage: predicting the probability of having participated in *Bolsa Familia* and estimating long run effects

In the second stage, we use data from the PNAD in 2014 and compute the probability of having participated in *Bolsa Familia* taking advantage of the fact that a random sub-sample of this survey answered to retrospective questions related to the covariates included in the first stage (specifically, respondents were asked to provide information about their circumstances when they were 15 years old and living with their parents). Then, we set a threshold of 0.257 for the predicted probability to divide the sample into treated and control groups (according to the optimality criteria described in the previous section) and restrict the sample to individuals aged 25 to 40. Furthermore, we exploit the fact that *Bolsa Escola* (and, later, the variable benefit of *Bolsa Familia*) was aimed to children and teenagers under 17 years old. Since the former was created in March 2001, only individuals born in 1985 or later were eligible to receive the benefit.

These two sources of exogenous variation allow us to estimate the long-run effect of *Bolsa Familia* through a difference-in-differences strategy. The econometric specification is given by the following equation:

[2] In all cases, we use a grid of 100 regularization parameters and choose one by 10-fold cross validation.

$$y_{it} = \alpha + \beta Treated_i + \gamma Post_t + \delta Treated_i \cdot Post_t + X_{it}^{'}\theta + \epsilon_{it} \tag{1}$$

where y is the outcome of interest for individual i belonging to cohort t, $Treated_i$ is a dummy equal to one when this individual belongs to the treatment group (i.e. has a predicted probability of program participation in childhood higher than 0.257), $Post_t$ is a dummy that equals one if this individual was born after 1984 and is, hence, eligible for the program, X includes control variables such as sex, age, age squared, household size, region and location area (rural or urban), and ϵ_{it} is the error term. Then, the estimate of the δ parameter is the difference-in-differences coefficient (DD) and, under the parallel trends assumption, captures the causal effect of the program in the long term. This assumption postulates that, in the absence of the implementation of BF, the two groups would have followed a similar trend in schooling, labor income and other outcomes of interest. Since municipalities are responsible for collecting the information that determines a household's eligibility to receive the transfer, we cluster standard errors at this geographic level taking into account the potential correlation within these units.

4 Results

4.1 Baseline estimates

Figure 4 shows the average value of each outcome for eligible and non-eligible cohorts. Years of schooling evolved at a similar pace in the treatment and control groups before the creation of the program. After that, years of education remained stagnant in the control group, while they increased slowly but steadily in the treatment group. As for labor income, both groups show a similar trend, mostly driven by age effects, and the average of the control group was about twice that of the treatment group before the intervention (beyond some noise in the estimates for the 1977 and 1981 cohorts). After the implementation of *Bolsa Familia*, this gap was substantially reduced.

Table 3 shows the estimates for the whole sample. We find a positive and significant overall effect for both years of schooling (around 0.8) and labor income (about US$250).

Table 3: Diff-in-diff estimates - Years of schooling and labor income

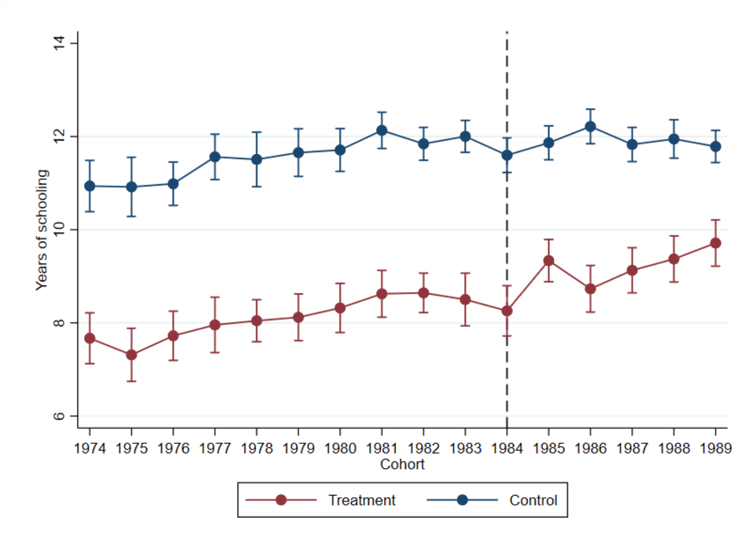
	Years of schooling		Labor income	
	Unconditional	Conditional	Unconditional	Conditional
DD	0.803***	0.784***	250.2***	247.6***
	(0.171)	(0.166)	(44.39)	(45.40)
Mean of the dependent variable	8.1	8.1	581.5	581.5
Observations	10,722	10,722	8,189	8,189

*** p<0.01, ** p<0.05, * p<0.1

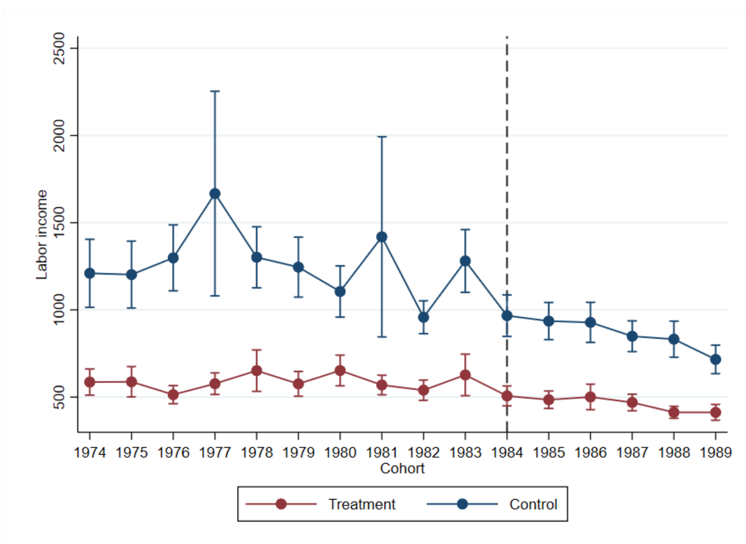
Note: Estimates were weighted by the inverse probability of selection. DD is the coefficient of the interaction term. Mean of the dependent variable corresponds to the treatment group in the pre-treatment period. Control variables included sex, age, age squared, household size, rural or urban location and region of residency. Cluster robust standard errors at municipality level in parentheses.

Figure 4: Outcome variables by cohort

(a) Years of schooling



(b) Labor income



4.2 Parallel trends

With the aim of strengthening the credibility of the causal interpretation of the results, we test the existence of parallel pre-trends between both groups. Specifically, we estimate the following model restricting the sample to 1974-1984 cohorts:

$$Y_{it} = \alpha + \beta Treated_i + \gamma Cohort_t + \delta Treated_i \cdot Cohort_t + X_{it}'\theta + \epsilon_{it} \tag{2}$$

Since $Cohort_t$ is the trend component of the model, the estimate of δ captures any difference in the pre-trends of outcome variables across groups. Table 4 shows the results.

Table 4: Pre-trends test - Years of schooling and labor income

	Years of schooling	Labor income
Treated.Cohort	0.0187	18.89
	(0.0339)	(13.30)
Observations	7,425	5,739

*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates were weighted by the inverse probability of selection. Treated.Cohort is the coefficient of the interaction term. Control variables included sex, age, age squared, household size, rural or urban location and region of residency. Cluster robust standard errors at municipality level in parentheses.

Both coefficients are small and non-significant at any conventional level. This evidence underpins the credibility of the parallel trends assumption and, consequently, the causal interpretation of our results. However, recent econometrics literature has shown that simply testing for the existence of pre-trends has several limitations.[3] Therefore, we follow the recommendations of Roth et al. (2023) in order to test the existence of parallel pre-trends and the sensitivity of the results to violations of the parallel trends assumption.

Firstly, we estimate the following dynamic specification:

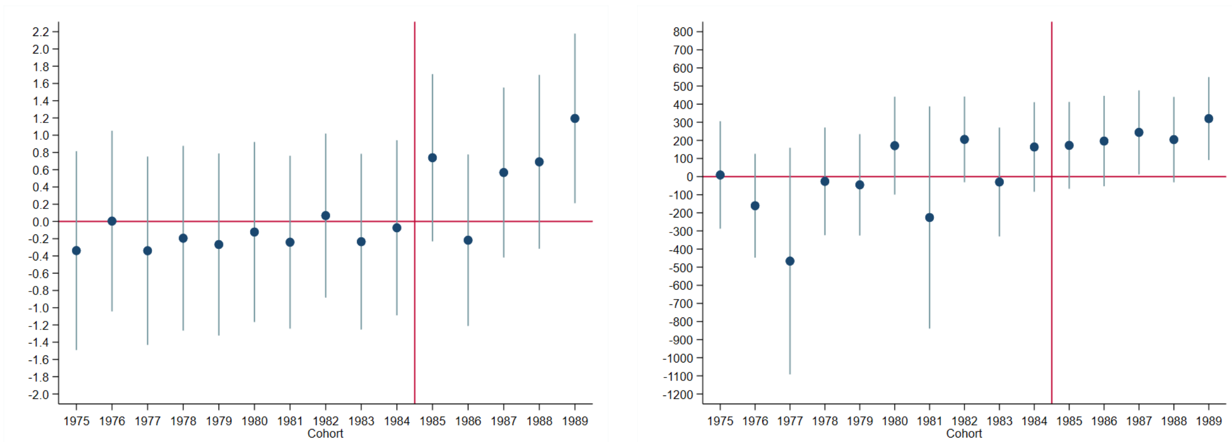
$$Y_{it} = \sum_{t=1}^{C} \beta_t Treated_i \cdot Cohort_t + \gamma Treated_i + \delta Cohort_t + \epsilon_{it} \tag{3}$$

Where C is the number of cohorts. Thus, we can obtain a different effect for each cohort. These estimates are shown in Figure 5.

[3] Roth et al. (2023) highlight four of them: the existence of parallel pre-trends does not ensure that they persist after treatment, the low power of usual tests to reject the null hypothesis of no difference in pre-trends, the existence of a selection bias when failing to reject the null hypothesis but there is a difference in pre-trends, and the absence of a clear procedure to follow when finding a difference in pre-trends.

Figure 5: Dynamic diff-in-diff estimates

(a) Years of schooling (b) Labor income



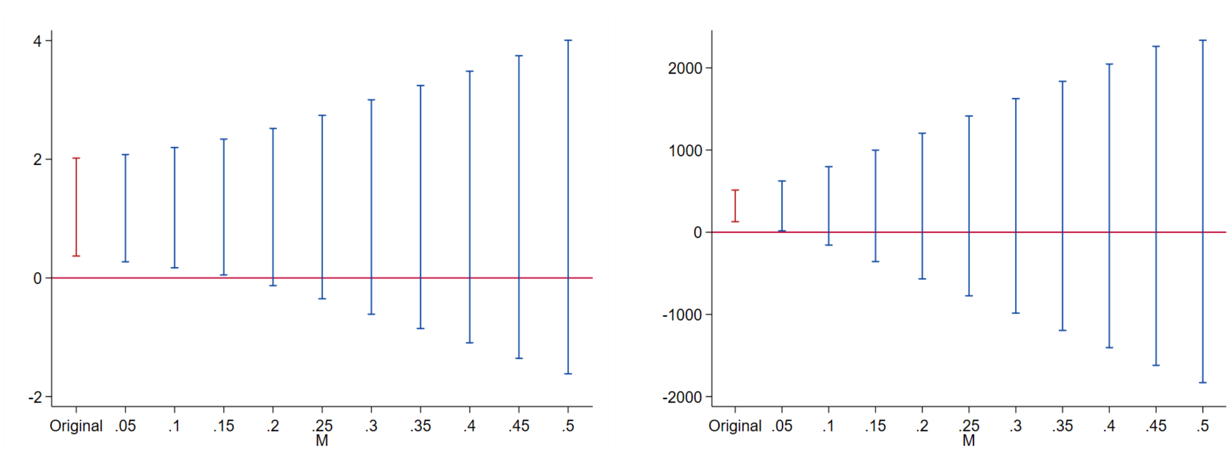
Note: Estimates were weighted by the inverse probability of selection. 1974 cohort is the base category. Each year shows the $\hat{\beta}_t$ coefficient estimated from Equation (3). 95% confidence intervals.

Though estimates for labor income are noisy for some specific cohorts, evidence seems consistent with the presence of parallel pre-trends for both outcome variables: all the coefficients for 1975-1984 cohorts are statistically non-different from zero. Moreover, the evolution of both variables shows a break around 1984 and becomes increasing, reaching significant and positive coefficients by 1989.

Next, we report the results of the sensitivity analysis proposed by Rambachan and Roth (2023). In general terms, the authors construct confidence intervals for the coefficients by taking a threshold $\bar{M}$ representing the ratio between the post-treatment parallel trend violation and the maximum pre-treatment violation. For example, if $\bar{M}$ is equal to 1 and the coefficient of the efffect statistically significant, this means that it is robust to a parallel trend violation equal to the largest pre-treatment violation. For the sake of simplicity, we only conduct the analysis for the coefficients obtained for the 1989 cohort. Results are shown in Figure 6.

Figure 6: Dynamic diff-in-diff estimates - Sensitivity analysis

(a) Years of schooling (b) Labor income



Note: Estimates correspond to 1989 cohort. 90% confidence intervals.

Considering a significance level of 10%, the effect on years of schooling is robust up to the value $\bar{M} = 0.2$. In other words, the maximum post-treatment violation of the parallel trends assumption consistent with a significant effect is 20% of the maximum violation of parallel pre-trends. Regarding labor income, the maximum threshold for a significant effect is only 10%. Although these values seem small, we believe that they are high enough to support the causal interpretation of our results given that the pre-trends estimates in our analysis do not show values significantly different from 0. Furthermore, we do not expect the groups to have been affected differently beyond the implementation of the *Bolsa Familia* program.

4.3 Placebo tests

To further test the robustness of our results, we run a series of placebo tests. First, we restrict the sample to the 1974-1984 cohorts and set the implementation of BF to different years. If our results actually capture the causal effect of BF, we should find that in this application most coefficients are not statistically significant. Results are shown in Table 5. In the case of schooling, none of the coefficients is significant at any conventional level. Regarding labor income, conditional on the covariates, we observe (weakly) significant coefficients for the 1982 and 1984 cohorts.

Table 5: Placebo test - Years of schooling and labor income

	Years of schooling		Labor income	
	Unconditional	Conditional	Unconditional	Conditional
1975	-0.215	-0.149	-26.39	-17.63
	(0.397)	(0.384)	(122.4)	(119.5)
1976	-0.0129	0.0614	-34.03	-51.54
	(0.314)	(0.304)	(84.62)	(87.89)
1977	-0.0719	-0.00916	37.76	16.40
	(0.261)	(0.253)	(73.66)	(74.59)
1978	0.0165	0.0595	197.2	162.1
	(0.226)	(0.219)	(139.1)	(126.4)
1979	0.0417	0.108	176.8*	142.8
	(0.212)	(0.202)	(104.9)	(97.58)
1980	0.100	0.151	178.6*	144.6
	(0.206)	(0.200)	(105.0)	(101.8)
1981	0.106	0.150	104.4	82.31
	(0.210)	(0.204)	(98.08)	(100.3)
1982	0.152	0.260	202.9**	182.8**
	(0.223)	(0.215)	(93.20)	(91.06)
1983	0.0538	0.0796	112.5	95.90
	(0.252)	(0.242)	(90.92)	(88.44)
1984	0.148	0.174	209.9**	163.9*
	(0.340)	(0.327)	(93.35)	(88.11)
Observations	7,425	7,425	5,739	5,739

*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates were weighted by the inverse probability of selection. Each row contains the coefficient of the interaction term corresponding to different years used as threshold to define the post-treatment period. Control variables included sex, age, age squared, household size, rural or urban location and region of residency. Cluster robust standard errors at municipality level in parentheses.

Second, we perform a leave-one-out test in order to provide additional evidence to support the parallel trends assumption. This test consists of checking whether the results are robust to the exclusion of particular cohorts. The estimates are shown in Table 6.

Table 6: Placebo test - Years of schooling and labor income

	Years of schooling		Labor income	
	Unconditional	Conditional	Unconditional	Conditional
1985	0.765***	0.781***	263.6***	267.2***
	(0.189)	(0.182)	(49.58)	(51.87)
1986	1.020***	1.027***	281.2***	287.9***
	(0.211)	(0.205)	(48.48)	(48.86)
1987	0.761***	0.815***	262.1***	268.6***
	(0.208)	(0.198)	(52.16)	(55.22)
1988	0.720***	0.716***	276.1***	263.5***
	(0.212)	(0.202)	(50.60)	(51.28)
1989	0.560***	0.574***	236.0***	247.3***
	(0.204)	(0.200)	(57.05)	(63.70)
Observations				
1985	10,007	10,007	7,657	7,657
1986	9,353	9,353	7,156	7,156
1987	9,308	9,308	7,141	7,141
1988	9,392	9,392	7,206	7,206
1989	9,393	9,393	7,207	7,207

*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates were weighted by the inverse probability of selection. Each row contains the coefficient of the interaction term corresponding to different years excluded from the sample to define the post-treatment period. Control variables included sex, age, age squared, household size, rural or urban location and region of residency. Cluster robust standard errors at municipality level in parentheses.

The estimated coefficients are very similar to those obtained for the baseline sample. The magnitude of the effect on schooling is higher when excluding the 1986 cohort and lower when excluding 1989 cohort. However, they are positive and significant. Hence, we can conclude that the estimated effect is not influenced by any particular cohort.

Finally, we estimate Equation 1 but setting non-labor income as the dependent variable. This concept aggregates different sources of income that are not expected to be affected by *Bolsa Familia* (pensions, rental income and donations). The estimates are shown in Table 7. Coefficients are small and not statistically significant (considering a confidence level of 5%). Overall, this additional evidence supports the robustness of our main results.

Table 7: Placebo test - Non-labor income

	Unconditional	Conditional
Treated.Cohort	9.816	10.93*
	(6.368)	(6.433)
Observations	10,754	10,754

*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates were weighted by the inverse probability of selection. Treated.Cohort is the coefficient of the interaction term. Control variables included sex, age, age squared, household size, rural or urban location and region of residency. Cluster robust standard errors at municipality level in parentheses.

4.4 Impact heterogeneity

Table 8 exhibits the estimates separately for men and women. Interestingly, the effect on schooling is similar among men and women. Instead, the effect on labor income is much larger among men. This difference could respond to the fact that female labor participation was notably lower than male labor participation in Brazil at the beginning of the 2010s. According to Gasparini and Marchionni (2015), male participation was above 90% in 2012, while female participation barely reached 70%.

Table 8: Diff-in-diff estimates - heterogeneous effects by gender

	Years of schooling		Labor income	
	Men	Women	Men	Women
DD	0.773***	0.821***	308.5***	191.3***
	(0.238)	(0.227)	(85.80)	(53.93)
Mean of the dependent variable	7.4	8.7	662.4	477.0
Observations	5,123	5,599	4,498	3,691

*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates were weighted by the inverse probability of selection. DD is the coefficient of the interaction term. Mean of the dependent variable corresponds to the treatment group in the pre-treatment period. Control variables included age, age squared, household size, rural or urban location and region of residency. Cluster robust standard errors at municipality level in parentheses.

Table 9 shows the diff-in-diff estimates for other subgroups. We find a positive and significant effect on years of schooling for rural and urban areas and for Afro-American and Non-Afro-American individuals. It is worth noting that the estimate is substantially larger among rural areas and Afro-Americans than among their counterparts. The effect on labor income is similar between rural and urban workers, although non-significant for rural ones, probably due to the small number of observations in that group. Unexpectedly, despite of the positive effect on schooling, the impact on labor incomes for African Americans appears to be nil. This result is consistent with the findings of Neidhöfer and Niño-Zarazúa (2019) for indigenous people in Chile, and it confirms the challenges faced by certain groups in the labor market, even when their educational achievements improve substantially.

Table 9: Diff-in-diff estimates: heterogeneous effects

(a) Years of schooling	(1) Rural	(2) Urban	(3) Afro-American	(4) Non-Afro-American
DD	1.297**	0.517***	1.229**	0.742***
	(0.639)	(0.174)	(0.536)	(0.177)
Mean of the dependent variable	5.9	8.7	8.0	8.1
Observations	1,156	9,566	940	9,782
(b) Labor income	**(1) Rural**	**(2) Urban**	**(3) Afro-American**	**(4) Non-Afroamerican**
DD	214.3	235.0***	3.169	263.5***
	(164.9)	(45.98)	(108.0)	(49.31)
Mean of the dependent variable	495.9	598.3	511.0	589.6
Observations	736	7,453	741	7,448

*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates were weighted by the inverse probability of selection. DD is the coefficient of the interaction term. Mean of the dependent variable corresponds to the treatment group in the pre-treatment period. Control variables included sex, age, age squared, rural or urban location. Cluster robust standard errors at municipality level in parentheses.

Table 10 presents the estimates for different subgroups of women. We find positive and significant effects on schooling for all of them. Interestingly, the effects on labor income are only significant for married women and those with children, i.e. the group of women with a lower labor force participation.

Table 10: Diff-in-diff estimates: heterogeneous effects for subgroups of women

	(1)	(2)	(3)	(4)
(a) Years of schooling	Married or in relationship	Single	No children	With children
DD	0.925***	0.790*	1.019**	1.003***
	(0.239)	(0.434)	(0.441)	(0.260)
Mean of the dependent variable	8.6	9.2	9.7	8.5
Observations	3,613	1,986	1,823	3,776
	(1)	(2)	(3)	(4)
(b) Labor income	Married or in relationship	Single	No children	With children
DD	254.6***	83.73	121.1	291.1***
	(63.94)	(88.90)	(112.1)	(57.54)
Mean of the dependent variable	480.7	469.7	602.2	435.1
Observations	2,185	1,506	1,385	2,306

*** p<0.01, ** p<0.05, * p<0.1

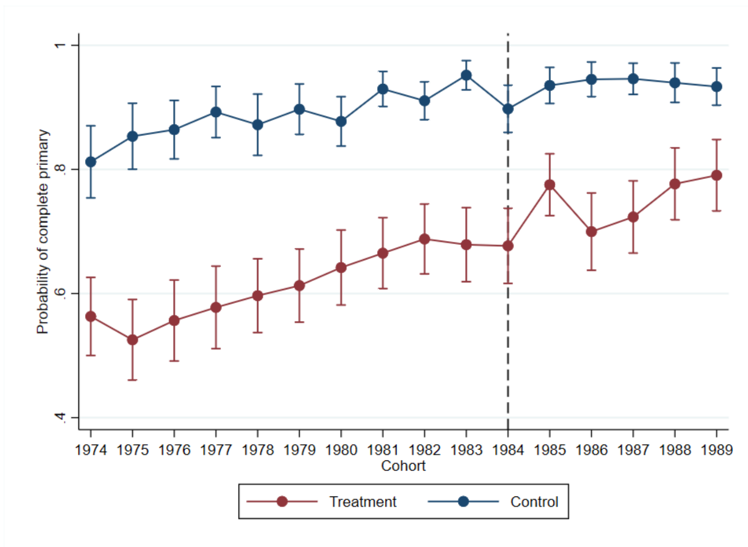
Note: Estimates were weighted by the inverse probability of selection. DD is the coefficient of the interaction term. Control variables included age, age squared, rural or urban location. Cluster robust standard errors at municipality level in parentheses.

4.5 Probability of completing formal education

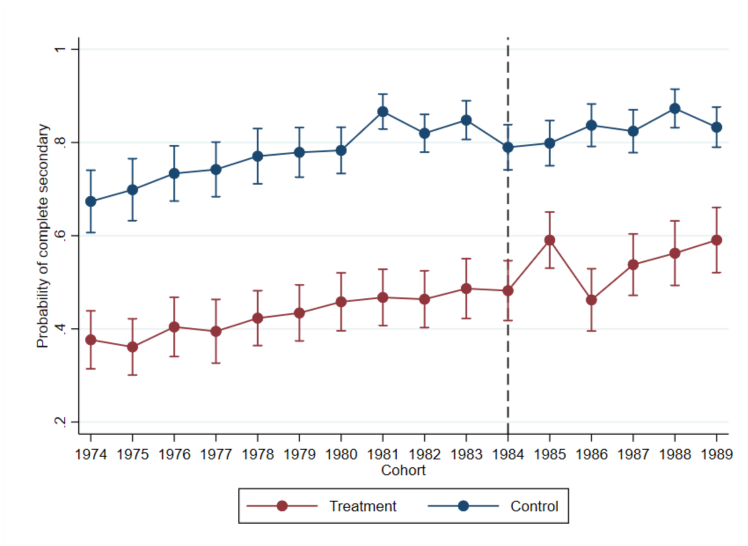
In what follows, we evaluate if the effects of the program on schooling are driven by higher completion rates of primary, secondary or tertiary education. Figure 7 shows the completion rates by cohort and educational level. The gap between treated and non-treated individual narrows after the implementation of *Bolsa Familia* in all cases. Table 11 presents the estimates for the three outcomes. We observe a positive and significant effect for the probability of completing primary education (around 9 pp.) which is similar for both men and women. Regarding secondary education, we observe a positive and significant coefficient for the whole sample (around 6 pp.). Interestingly, the effect is significant and large for women (around 10 pp.) and small and non-significant for men. Finally, we did not find any significant effect on tertiary education (though coefficients are positive). These results suggest that completion of different education levels may have played a role in securing a higher labor income, beyond the accumulation of years of schooling.

Figure 7: Completion rates by cohort and educational level

(a) Primary



(b) Secondary



(c) Tertiary

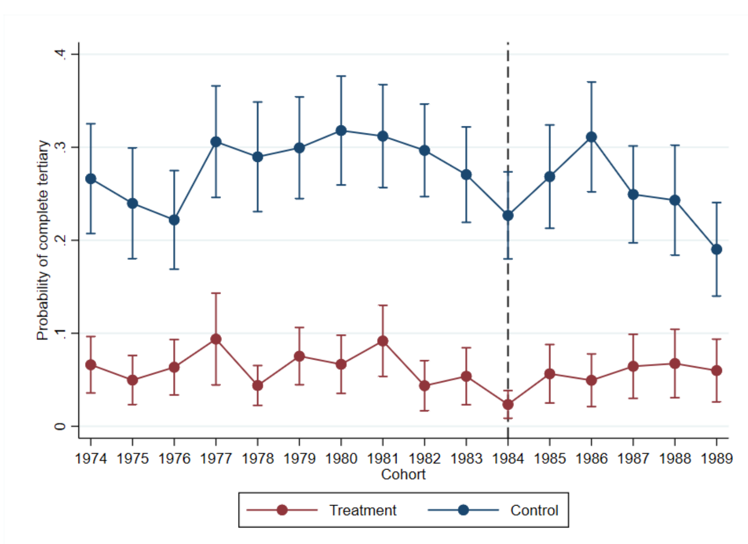


Table 11: Diff-in-diff estimates - Completion of primary, secondary and tertiary education

(a) Primary	(1) Unconditional	(2) Conditional	(3) Only men	(4) Only women
DD	0.0880***	0.0876***	0.0874***	0.0915***
	(0.0190)	(0.0186)	(0.0283)	(0.0236)
Mean of the dependent variable	0.61	0.61	0.55	0.68
(b) Secondary	**(1) Unconditional**	**(2) Conditional**	**(3) Only men**	**(4) Only women**
DD	0.0646***	0.0585***	0.0132	0.105***
	(0.0213)	(0.0205)	(0.0320)	(0.0283)
Mean of the dependent variable	0.43	0.43	0.55	0.49
(c) Tertiary	**(1) Unconditional**	**(2) Conditional**	**(3) Only men**	**(4) Only women**
DD	0.0255	0.0216	0.0200	0.0227
	(0.0167)	(0.0165)	(0.0223)	(0.0234)
Mean of the dependent variable	0.06	0.06	0.04	0.08
Observations	10,754	10,754	5,134	5,620

*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates were weighted by the inverse probability of selection. DD is the coefficient of the interaction term. Mean of the dependent variable corresponds to the treatment group in the pre-treatment period. Control variables included sex, age, age squared, household size, rural or urban location and region of residency. Cluster robust standard errors at municipality level in parentheses.

4.6 Probability of receiving social transfers

There is an ongoing debate regarding the dependency effects of social programs: specifically, whether a transfer program leads people to become more reliant on social assistance in the future. To investigate this question, we estimate Model (1) while incorporating a dummy variable that indicates whether a household is currently receiving cash transfers as depending variable. Since this information is not directly observable in the database, but the survey inquires about other sources of non-labor income, we consider that the household is a beneficiary of social transfers if the reported amount is between 35 and 336 R$ (i.e., the minimum and the maximum amount that a *Bolsa Familia* beneficiary household could obtain in 2014). Figure 8 shows the average of this variable by cohort.

While the probability of receiving social transfers remains constant across cohorts in the control group, the pattern looks more irregular among treated individuals. However, we can observe an overall decrease after the implementation of *Bolsa Familia*. Table 12 shows the point estimates for this dependent variable, which suggest that the transfer from BF received during childhood led to a significant decline of around 3.7 pp. in the probability of receiving social transfers in adulthood. This effect is larger among men (almost 6 pp.) and near to 1 pp. but non significant among women.

Figure 8: Probability of receiving social transfers by cohort

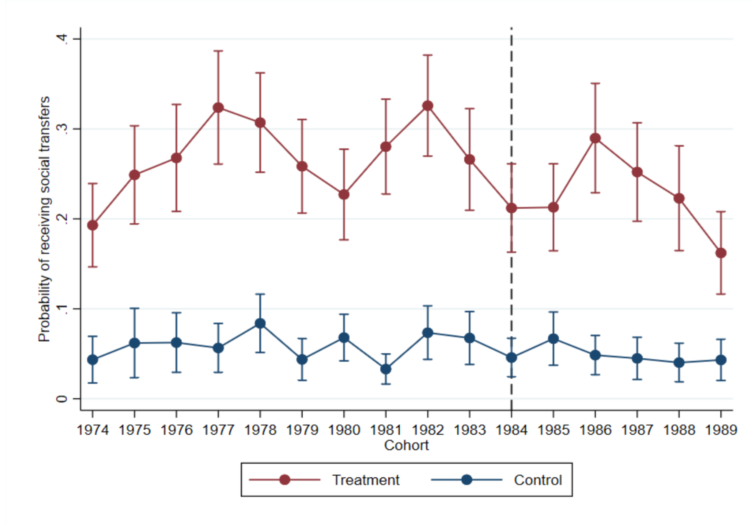


Table 12: Diff-in-diff estimates - Probability of receiving social transfers

	(1) Unconditional	(2) Conditional	(3) Only men	(4) Only women
DD	-0.0252*	-0.0370**	-0.0580***	-0.0128
	(0.0152)	(0.0153)	(0.0193)	(0.0241)
Mean of the dependent variable	0.26	0.26	0.21	0.31
Observations	10,754	10,754	5,134	5,620

Note: Estimates were weighted by the inverse probability of selection. DD is the coefficient of the interaction term. Mean of the dependent variable corresponds to the treatment group in the pre-treatment period. Control variables included sex, age, age squared, household size, rural or urban location and region of residency. Cluster robust standard errors at municipality level in parentheses.

5 Mechanisms

This section is devoted to the analysis of the potential mechanisms that could underlie the results previously found. One channel that could have led to higher labor income among treatment group members is the transition from low-productivity to higher-productivity activity sectors. We explore this possibility estimating Equation 1 using dummies for 16 productive activities as dependent variables.[4] Estimates are exhibited in Table 13.

[4]This sectors are defined according to major groups of the third revision of the International Standard Industrial Classification of All Economic Activities (ISIC).

Table 13: Diff-in-diff estimates - Activity sector

	(1) Unconditional	(2) Conditional
Agriculture, hunting and forestry	0.0180	0.00117
	(0.0162)	(0.0144)
Fishing	-0.00403**	-0.00478**
	(0.00199)	(0.00209)
Mining and quarrying	0.00205	0.00161
	(0.00372)	(0.00385)
Manufacturing	0.00275	0.00879
	(0.0199)	(0.0196)
Electricity, gas and water supply	0.0110**	0.0107**
	(0.00426)	(0.00444)
Construction	0.0322*	0.0303*
	(0.0170)	(0.0167)
Wholesale and retail trade	-0.0199	-0.0192
	(0.0223)	(0.0225)
Hotels and restaurants	-0.00162	0.00360
	(0.0109)	(0.0109)
Transport, storage and communications	0.00594	0.00431
	(0.0133)	(0.0130)
Financial intermediation	-0.00585	-0.00628
	(0.00751)	(0.00741)
Real estate, renting and business activities	-0.0198	-0.0184
	(0.0162)	(0.0160)
Public administration and defense; compulsory social security	0.0266**	0.0228**
	(0.0117)	(0.0115)
Education	-0.00656	-0.00496
	(0.0151)	(0.0150)
Health and social work	-0.000581	0.000812
	(0.0124)	(0.0122)
Other community, social and personal service activities	-0.0225**	-0.0223**
	(0.00975)	(0.00973)
Activities of private households as employers	-0.0172	-0.00771
	(0.0141)	(0.0134)
Observations	8,575	8,575

*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates were weighted by the inverse probability of selection. Each row contains the coefficient of the interaction term for each activity sector. Control variables included sex, age, age squared, household size, rural or urban location and region of residency. Cluster robust standard errors at municipality level in parentheses.

Some of the estimated coefficients support this hypothesis. We observe a significant drop in the probability of being employed in the fishing sector and in other community, social and personal service activities (both located in the lower tail of the wage distribution). On the other hand, we observe a surge in the probability of being employed in public administration and the supply of electricity, gas and water. This sectors are located in the upper tail of the wage distribution (see Table A.3). These results are in line with the findings of Behrman et al. (2011) and Parker and Vogl (2023) for Mexico, namely a drop in the probability of working in the agricultural sector, and of Barham et al. (2017) for Nicaragua, who find an increased likelihood of engaging in off-farm activities such as large plantation work, construction, security or non-agricultural self-employment.

6 Conclusions

Conditional Cash Transfers have become a key policy tool to fight poverty in developing countries in the last 25 years. While their short-term effects have been widely analyzed by the literature, less is known about their impact in the long-run and their effectiveness in achieving the goal of breaking the intergenerational transmission of poverty through the improvement of human capital formation among beneficiaries. In this paper, we make a contribution to this scarce literature studying the case of the Brazilian *Bolsa Familia* program.

We propose a novel approach that relies on the two-sample-two-stage method to identify former beneficiaries of the transfer and exploits the rich module of retrospective questions of the 2014 PNAD, the main national household survey of Brazil. Furthermore, we take advantage of the age restriction imposed by the program at implementation to adopt a difference-in-differences approach and estimate the causal effect of the program on several outcomes.

We find significant and positive effects of having received the benefit in childhood on schooling and labor income measured in adulthood. Specifically, we observe an overall increase of 0.8 years of schooling and of US$250 labor income among beneficiaries. This increase in schooling is reflected in a rise in the probability of having completed formal education (9 pp. in the the case of primary education and 6 pp. in the case of secondary). Moreover, we find a significant drop of 4 pp. in the probability of being a current beneficiary of social transfers.

Besides, we contribute to the understanding of the mechanisms underlying the sharp increase found for labor income. Beyond the accumulation of years of schooling and the surge in the probability of completion of primary and secondary education, we find that sectoral reallocation among workers plays a relevant role. Indeed, the evidence suggests that former beneficiaries of *Bolsa Familia* are less likely to work in low-productivity sectors such as fishing or community services, but more likely to be employed in higher productivity sectors such as public administration or the supply of basic services.

In summary, our findings are consistent with the arguments underlying the design of CCTs that postulate that the promotion of human capital accumulation of poor children contributes to the reduction of poverty and government dependence among future generations. Nevertheless, our findings also emphasize that, for African Americans in Brazil, the improvements in schooling caused by the program did not lead to corresponding increases in labor income. To improve the effectiveness and outreach of social programs, future research should endeavor to identify the persisting barriers facing marginalized groups and explore potential strategies for overcoming them.

References

Alam, A., Baez, J. E., and Del Carpio, X. V. (2011). Does Cash for School Influence Young Women's Behavior in the Longer Term? Evidence from Pakistan. *World Bank Policy Research Working Paper*, (5669).

Araujo, M. C., Bosch, M., and Schady, N. (2019). Can Cash Transfers Help Households Escape an Intergenerational Poverty Trap? In *The Economics of Poverty Traps*, pages 357–382. University of Chicago Press.

Attanasio, O., Sosa, L. C., Medina, C., Meghir, C., and Posso-Suárez, C. M. (2021). Long Term Effects of Cash Transfer Programs in Colombia. *National Bureau of Economic Research Working Paper Series*, (29056).

Baez, J. E. and Camacho, A. (2011). Assessing the Long-term Effects of Conditional Cash Transfers on Human Capital: Evidence from Colombia. *IZA Discussion Papers*, (5751).

Bailey, M., Hoynes, H., Rossin-Slater, M., and Walker, R. (2023). Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence From the Food Stamps Program. *The Review of Economic Studies*, page rdad063.

Baird, S., McIntosh, C., and Özler, B. (2019). When The Money Runs Out: Do Cash Transfers have Sustained Effects on Human Capital Accumulation? *Journal of Development Economics*, 140:169–185.

Barham, T., Macours, K., and Maluccio, J. A. (2017). Are Conditional Cash Transfers Fulfilling their Promise? Schooling, Learning, and Earnings after 10 Years. *CEPR Discussion Paper*, (11937).

Barham, T., Macours, K., and Maluccio, J. A. (2018). Experimental Evidence of Exposure to a Conditional Cash Transfer During Early Teenage Years: Young Women's Fertility and Labor Market Outcomes. *CEPR Discussion Paper*, (131615).

Behrman, J. R., Parker, S. W., and Todd, P. E. (2011). Do Conditional Cash Transfers for Schooling Generate Lasting Benefits? A Five-year Followup of PROGRESA/*Oportunidades*. *Journal of Human Resources*, 46(1):93–122.

Björklund, A. and Jäntti, M. (1997). Intergenerational Income Mobility in Sweden Compared to the United States. *The American Economic Review*, 87(5):1009–1018.

Brunori, P., Ferreira, F. H., and Neidhöfer, G. (2023). *Inequality of opportunity and intergenerational persistence in Latin America*. United Nations University World Institute for Development Economics Research.

Duque, V., Rosales-Rueda, M., and Sanchez, F. (2019). How Do Early-Life Shocks Interact with Subsequent Human Capital Investments? Evidence from Administrative Data. *University of Sydney, School of Economics Working Papers Series*, (17).

Filmer, D. and Schady, N. (2014). The Medium-Term Effects of Scholarships in a Low-Income Country. *Journal of Human Resources*, 49(3):663–694.

Fiszbein, A. and Schady, N. R. (2009). Conditional Cash Transfers: Reducing Present and Future Poverty. *World Bank Policy Research Report*.

Gaentzsch, A. (2020). Do Conditional Cash Transfers (CCTs) Raise Educational Attainment? An Impact Evaluation of Juntos in Peru. *Development Policy Review*, 38(6):747–765.

Garcia, S. and Saavedra, J. E. (2023). Chapter 7 - Conditional Cash Transfers for Education. volume 6 of *Handbook of the Economics of Education*, pages 499–590. Elsevier.

Gasparini, L. and Marchionni, M. (2015). *Bridging Gender Gaps? The Rise and Deceleration of Female Labor Force Participation in Latin America*. Universidad Nacional de La Plata, La Plata.

Ham, A. and Michelson, H. C. (2018). Does the Form of Delivering Incentives in Conditional Cash Transfers Matter over a Decade Later? *Journal of Development Economics*, 134:96–108.

Liu, X. (2012). Classification Accuracy and Cut Point Selection. *Statistics in Medicine*, 31(23):2676–2686.

Millán, T. M., Macours, K., Maluccio, J. A., and Tejerina, L. (2020). Experimental Long-Term Effects of Early-Childhood and School-Age Exposure to a Conditional Cash Transfer Program. *Journal of Development Economics*, 143:102385.

Millán, T. M., Barham, T., Macours, K., Maluccio, J. A., and Stampini, M. (2019). Long-Term Impacts of Conditional Cash Transfers: Review of the Evidence. *The World Bank Research Observer*, 34(1):119–159.

Neidhöfer, G. and Niño-Zarazúa, M. (2019). The Long(er)-Term Impacts of *Chile Solidario* on Human Capital and Labor Income. *Population and Development Review*, pages 209–244.

Parker, S. W. and Vogl, T. (2023). Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico. *The Economic Journal*, 133(655):2775–2806.

Rambachan, A. and Roth, J. (2023). A More Credible Approach to Parallel Trends. *The Review of Economic Studies*, 90(5):2555–2591.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature. *Journal of Econometrics*, 235(2):2218–2244.

Sanchez Chico, A., Macours, K., Maluccio, J., and Stampini, M. (2018). Six Years of Comunidades Solidarias Rurales: Impacts on School Entry of an Ongoing Conditional Cash Transfer Program in El Salvador. *IDB Working Paper Series-Inter-American Development Bank*, (IDB-WP-908).

A Appendix

As a robustness check for labor income results, we estimate model (1) using household per capita income as the dependent variable and restricting the sample to household heads. Results are shown in Table A.1. We find a positive and significant effect both for the unconditional and the conditional estimate. The latter is around US$83 and represents an increase of 10% over the average household per capita income of former BF beneficiaries.

Table A.1: Diff-in-diff estimates - Household per capita income

	(1) Unconditional	(2) Conditional
DD	124.5***	83.24**
	(45.68)	(40.63)
Mean of the dependent variable	834.35	834.35
Observations	5,100	5,100

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note: Estimates were weighted by the inverse probability of selection. DD is the coefficient of the interaction term. Control variables included sex, age, age squared, household size, rural or urban location and region of residency. Cluster robust standard errors at municipality level in parentheses.

Table A.2: Estimates of probability of participation in *Bolsa Familia* - Logistic regression

Covariates	Coefficient	Covariates	Coefficient
Father's occupational category		**State of residency**	
Informal salaried worker	0.430***	Acre	0.476***
	(0.0390)		(0.140)
Public salaried worker/Serviceman	-0.253***	Amazonas	0.534***
	(0.0817)		(0.123)
Self-employed	0.286***	Roraima	0.771***
	(0.0354)		(0.189)
Employer	-0.604***	Para	0.414***
	(0.0800)		(0.108)
Other	0.387***	Amapa	-1.075***
	(0.0933)		(0.219)
Unemployed/Inactive	0.259***	Tocantins	0.663***
	(0.0504)		(0.131)
Mother's occupational category		Maranhao	1.278***
Informal salaried worker	0.701***		(0.122)
	(0.0558)	Piaui	1.194***
Public salaried worker/Serviceman	0.219**		(0.125)
	(0.0948)	Ceara	1.438***
Self-employed	0.479***		(0.101)
	(0.0589)	Rio Grande do Norte	1.058***
Employer	-0.792***		(0.120)
	(0.213)	Paraiba	1.473***
Other	0.769***		(0.118)
	(0.0605)	Pernambuco	1.022***
Unemployed/Inactive	0.500***		(0.103)
	(0.0490)	Alagoas	1.162***
Parents' educational level			(0.121)
Incomplete primary	-0.127**	Sergipe	0.700***
	(0.0567)		(0.128)
Complete primary	-0.753***	Bahia	1.024***
	(0.0669)		(0.0982)
Incomplete secondary	-0.832***	Minas Gerais	0.483***
	(0.0724)		(0.0991)
Complete secondary	-1.392***	Espirito Santo	0.334***
	(0.0634)		(0.124)
Incomplete tertiary	-2.457***	Rio de Janeiro	-0.639***
	(0.124)		(0.114)
Complete tertiary	-3.956***	Sao Paulo	-0.465***
	(0.184)		(0.103)
Location area		Parana	-0.181*
Urban	-0.346***		(0.109)
	(0.0349)	Santa Catarina	-0.979***
Constant	-1.113***		(0.139)
	(0.119)	Rio Grande do Norte	-0.251**
			(0.107)
		Mato Grosso do Sul	-0.425***
			(0.137)
		Mato Grosso	-0.179
			(0.123)
		Goias	-0.0538
			(0.111)
		Distrito Federal	-0.704***
			(0.143)
Observations	49,728		

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note: Estimates were weighted by the inverse probability of selection. Ommited variables were formal salaried worker, no education, rural area and Rondonia state. Robust standard errors in parentheses.

Table A.3: Mean labor income by sector of activity (current US$)

Sector	Income
Fishing	**268.8**
Activities of private households as employers	314.2
Agriculture, hunting and forestry	480.8
Hotels and restaurants	563.4
Construction	647.7
Other community, social and personal service activities	**677.1**
Wholesale and retail trade	686.7
Manufacturing	761.2
Transport, storage and communications	789.1
Education	868.0
Real estate, renting and business activities	1,063.8
Electricity, gas and water supply	**1,074.6**
Health and social work	1,110.6
Public administration and defense	**1,194.6**
Mining and quarrying	1,225.9
Financial intermediation	1,474.0
Observations	70,856

Source: Own elaboration based on PNAD-2014. Note: sectors of activity correspond to major groups of the third revision of the International Standard Industrial Classification of All Economic Activities (ISIC).