Fundação Getúlio Vargas

Doctoral Thesis

Essays on Housing: Household Choices, Health, and Consumption

---

Author:
Laísa Rachter de Sousa Dias

Supervisor:
Cecilia Machado Berriel

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

Escola de Pós-Graduação em Economia

December 2020

Dados Internacionais de Catalogação na Publicação (CIP)
Ficha catalográfica elaborada peloSistema de Bibliotecas/FGV

Dias, Laísa Rachter de Sousa
Essays on housing: household choices, health, and consumption / Laísa Rachter de Sousa Dias. – 2020.

200 f.

Tese (doutorado) - Fundação Getulio Vargas, Escola Brasileira de Economia e Finanças.
Orientador: Cecilia Machado Berriel.
Inclui bibliografia.

1. Política habitacional. 2. Economia da família. 3. Saúde. 4. Consumo. 5. Avaliação de políticas públicas. I. Berriel, Cecilia Machado. II. Fundação Getulio Vargas Escola Brasileira de Economia e Finanças. III. Título.

CDD – 309.81

Elaborada por Maria do Socorro Almeida – CRB-7/4254

FGV

Laísa Rachter de Sousa Dias

"Essays on Housing: Household Choices, Health, and Consumption".

Tese apresentado(a) ao Curso de Doutorado em Economia do(a) EPGE Escola Brasileira de Economia e Finanças - FGV EPGE para obtenção do grau de Doutor(a) em Economia.

Data da defesa: 18/12/2020

ASSINATURA DOS MEMBROS DA BANCA EXAMINADORA

Presidente da Comissão Examinadora: Profº/ª Cecilia Machado Berriel

Cecilia Machado

Cecilia Machado Berriel

P/ Cecilia Machado

Francisco Junqueira Moreira da Costa

P/ Cecilia Machado

José Gustavo Féres

P/ Cecilia Machado

Breno Ramos Sampaio

P/ Cecilia Machado

Maína Celidônio de Campos

Em cumprimento Lei nº 13.979 de 06/02/20 - DOU nº 27 de 07/02/20, a Portaria MEC nº 473 de 12/05/20 - DOU nº 90 de 13/05//20 e ao Decreto nº.068 de 11/05/20 - Poder Executivo do Estado do Rio de Janeiro, DOE nº 082-A em 11/05/20 que dispõe sobre a suspensão temporária das atividades acadêmicas presenciais e a utilização de recursos tecnológicos (em conformidade à legislação vigente), face ao COVID-19, as apresentações das defesas de Tese e Dissertação, de forma excepcional, serão realizadas de forma remota e síncrona, incluindo-se nessa modalidade membros da banca e discente.

Lucas Jóver Maestri
Coordenador

Antonio de Araujo Freitas Junior
Pró-Reitor de Ensino, Pesquisa e Pós-Graduação FGV

Antonio Freitas, PhD
Pró-Reitor de Ensino, Pesquisa e Pós-Graduação
Fundação Getúlio Vargas

Instrução Normativa nº 01/19, de 09/07/19 - Pró-Reitoria FGV
Em caso de participação de Membro(s) da Banca Examinadora de forma não-presencial*, o Presidente da Comissão Examinadora assinará o documento como representante legal, delegado por esta I.N.
*Skype, Videoconferência, Apps de vídeo etc

Acknowledgments

Agradeço à Deus, sem o qual nenhum sonho seria possível.

Agradeço ao meu marido, Arthur, pelo amor e apoio incondicional – obrigada por caminhar ao meu lado nesses últimos 12 anos, por ser meu parceiro na vida e nos meus interesses de pesquisa. Seu entusiasmo pela pesquisa acadêmica me inspira. Agradeço também ao meu filho Joaquim – a sua chegada no meio do doutorado me motivou a encontrar a minha melhor versão.

Agradeço aos meus pais, Ida e Jairo, pelo amor, pela dedicação e por terem investido tudo o que puderam na minha educação. O esforço e investimento de vocês foram fundamentais para que eu chegasse até aqui. À toda minha família, irmãs, cunhados e sobrinhos, obrigada pelo apoio incondicional.

Agradeço à minha orientadora, Cecilia Machado, pelas oportunidades, ensinamentos, conselhos e principalmente pela confiança. Agradeço também aos professores da FGV que tanto contribuíram para a minha formação. Em especial, agradeço ao Luis Braido, Francisco Costa, Cezar Santos, Carlos Eugênio, André Trindade, Marcelo Santana e João Victor pelas discussões e sugestões enriquecedoras ao longo do desenvolvimento da minha pesquisa. Agradeço também a todos aos funcionários da Escola de Pós Graduação em Economia da FGV, em especial a Ana Lúcia, Andrea e Vera. Agradeço também aos professores Carlos Eugênio e Lucas Maestri, coordenadores da pós-graduação no início e no final da minha trajetória como aluna de doutorado, pelo compromisso com a excelência do programa de pós-graduação da EPGE.

Meu muito obrigado tambeém aos meus amigos da FGV que tornaram essa caminhada mais leve. Em especial, à Kátia, Renata, Luana, Rafael, Heron, Valdemar e Rafaela.

Por último, agradeço ao apoio financeiro da CAPES e da FGV que viabilizaram esse trabalho.

Summary

This thesis contains three independent but interrelated papers on the economic effects of housing.

The first paper evaluates the effects of a housing program that built houses for low-income families from the city of Rio de Janeiro (Brazil). We explore the lotteries used to select the program's beneficiaries to provide evidence of its effects on location, housing quality, housing costs, and household choices. The program induced households to move to less populated, more impoverished, and more distant neighborhoods. However, it increased the houses' quality in which these households lived and decreased their housing costs. Increases in other expenditures did not compensate for the decline in housing costs. Furthermore, we find the program did not influence labor force participation and income and weakly increased teenagers' enrollment. Overall, our evidence contributes to understanding the mechanisms through which housing programs affect well-being.

The second paper examines the effects of better houses on infant health also in the context of Brazil's *Minha Casa Minha Vida* program, which built roughly 900,000 houses to poor households in Brazil during the period 2010-2017. We use a regression discontinuity design and administrative data to estimate the program's effects on health at birth and infant health. We find the program reduced the share of households living in inadequate houses by 18 percentage points. We find this improvement in housing conditions led to increases in birth weight and decreases in infant (before 1 year) mortality caused by conditions originating in children's perinatal period. We find no effect of the program in children with more than one year. Our results point out the importance of better houses in improving health at birth.

The third paper investigates the effect of real-estate prices on non-durable consumption in Brazil. For that, we explore a state-level panel of the determinants of non-durable consumption growth during the period 2008-2017. We estimate the effect of house prices on consumption using the reduced-form equation proposed by Campbell (2007), which is derived from simulating a theoretical model of housing and consumption choice under debt constraints. Due to data limitations, use data aggregated at the state level to estimate our panel-data regressions. Our results suggest that changes in house prices significantly affect non-durable consumption in Brazil. The magnitudes are quantitatively close to the effects found for the U.K. by Campbell (2007). Furthermore, we document that the effect of house prices on non-durable consumption is asymmetric, stronger in the "bust" than in the "boom" phase of the business cycle. This difference in the effects during different phases of the business cycle suggests that borrowing constraints might explain the effects of house prices on non-durable consumption.

Keywords: *Housing Policies, Household Choices, Health, Consumption, Impact Evaluation*

Better Neighborhoods or Better Houses?
The Effects of Housing Policies on Poor Households in Brazil*

Abstract

This paper evaluates the effects of a housing program that built houses for low-income families from the city of Rio de Janeiro (Brazil). We explore the lotteries used to select the program's beneficiaries to provide evidence of its effects on location, housing quality, housing costs, and household choices. The program induced households to move to less populated, more impoverished, and more distant neighborhoods. However, it increased the houses' quality in which these households lived and decreased their housing costs. Increases in other expenditures did not compensate for the decline in housing costs. Furthermore, we find the program did not influence labor force participation and income and weakly increased teenagers' enrollment. Overall, our evidence contributes to understanding the mechanisms through which housing programs affect well-being.

Keywords: *Housing Policies, Houses, Neighborhoods, Schooling, Labor Supply*

*We are grateful to Luís Braido, Francisco Costa, Marcelo Sant'anna, and André Trindade for comments and suggestions. We gratefully acknowledge financial support from CAPES and FGV. We thank Brenda Prallon and Nícolas Pinto for their help in digitizing and organizing the data. All errors are our own.

1 Introduction

The United Nations estimates that over 800 million people live in slums located in the cities throughout the world (UN-HABITAT, 2015). The lack of sanitation, the overpopulation, and the poor public services that define these neighborhoods are thought to increase poverty, deteriorate health, and stimulate crime (Wilson, 1987; Jencks & Mayer, 1990; Glaeser, 2011; Marx et al., 2013). Throughout the world, governments have responded to this issue by offering poor households houses on peripheries (Barnhardt et al., 2017). However, to the extent that these policies relocate households further from job opportunities, there is a concern that they might have unintended consequences (Glaeser, 2011; Picarelli, 2019).

This paper contributes to the literature on the consequences of housing policies in developing countries by documenting the direct and indirect effects of a large-scale housing program called *Minha Casa Minha Vida* (hereafter, MCMV) implemented in Brazil in the 2010s. The MCMV consists of different initiatives that subsidized home purchases. Our research focuses on poor households (total income up to R$ 1,600 per month or US$ 300 at the current exchange rates) living in the municipality of Rio de Janeiro. These households were selected to receive highly subsidized apartments built specifically for the program using lotteries.[1] We explore two lotteries that took place in the year 2013 to examine the direct effects of this program on neighborhood quality, housing quality, and housing costs and its indirect effects on school enrollment, labor force participation, and income.

Our investigation uses a novel dataset linking the lotteries' official records, data of the program's contracts, geo-coded information on neighborhoods, jobs, and schools, as well as socioeconomic information from the *Cadastro Único* (Brazil's unified registry of beneficiaries of social policies) from the period 2012-2018. This dataset enables us to compare

[1] We choose Rio de Janeiro mainly for two reasons. First, it was one of the few cities in Brazil to make available the list of participants and lottery winners with name and CPF (taxpayer registration number). Second, the municipality organized general lotteries with clear rules based on the Brazilian Federal Lottery.

winners ("treatment group"), and losers ("control group") of these lotteries both before and after the lotteries occurred and the units built under the program were delivered. To ensure comparability between these groups, we focus on beneficiaries of the *Bolsa Família* program (Brazil's flagship social policy) who entered into the *Cadastro Único* before the MCMV started. Because beneficiaries of the *Bolsa Família* program must update their information in 24-month intervals, the first restriction ensures we observe most of these households multiple times. Moreover, because households were supposed to register in the *Cadastro Único* to participate in MCMV lotteries, the second restriction ensures we focus on households who entered our data because of the program. After these restrictions, we provide evidence that the treatment and the control groups' demographic and economic outcomes were comparable for the lotteries that occurred in 2013.[2]

We divide our investigation into four parts. First, we examine the program's take up. Using contract records, we estimate households in the treatment group are 54 p.p. more likely to sign a contract to purchase a house built under the MCMV program than households in the control group. Using administrative data, we estimate that these households are roughly 30 p.p. more likely to live or move to the neighborhood where the MCMV projects were built. There is no evidence these numbers decrease in the first three years after the households move, contrasting with the evidence of increasing program exit over time documented by Barnhardt et al. (2017) in India.

Second, we document the MCMV effects on neighborhood quality. Combining the MCMV's records, administrative data, and geo-coded information of neighborhoods and the location of existing schools and job opportunities, we provide evidence the program moved households to neighborhoods that are less populated, poorer, and more distant from jobs and schools.

[2] Our results are robust to using other lotteries. However, we opt to exclude them from our empirical investigation because we cannot ensure the treatment and control groups are balanced. We discuss this issue in detail in Section 4.

Third, we document the MCMV effects on housing quality and housing costs. Combining MCMV's records and administrative data, we find that households in the treatment group have houses 26% larger, 6 p.p. more likely to have wood/tile floor, and 3 p.p. more likely to be connected to the sewage system than the households in the control group. We further find these households reduce their expenditures with rents by R$ 37.8-43.8 (70-80% of the control mean). While increases in spending with utilities mitigate the reduction in rents, expenditures in general fall by R$ 28.9-33.6 (10% of the control mean). These findings highlight the heterogeneous effects of the program on houses and neighborhoods.

Fourth, to understand how the heterogeneous effects of the MCMV on houses and neighborhoods map into economic outcomes, we examine the program's short-run effects on labor supply, income, and school enrollment. Using administrative data, we find null effects of the program on school labor supply and income both immediately and some years after the treatment. The decrease in total expenditures combined with the null effect on income indicates that households might be increasing savings or investments due to the program. Furthermore, we find evidence the program increases the enrollment of teenagers. However, we find no effects on enrollment in high school and high school completion rates, indicating age-grade distortion might be growing.

The evidence provided in this paper contributes to the growing literature on the effects of housing programs on socio-economic outcomes for low-income households. The literature on housing programs in developing countries highlight these programs often move families to more isolated neighborhoods (e.g., Barnhardt et al. (2017), Picarelli (2019), and Franklin (2019)). This might deteriorate their job prospects (e.g., Picarelli (2019)) and induce program exit (e.g., Barnhardt et al. (2017)). Consistent with this literature, our evidence suggests the MCMV moves households to more isolated neighborhoods. However, we find that the program improves housing quality and reduces housing costs, benefits not previously documented in the literature. Furthermore, we find no evidence of program exit or reductions in employment. Indeed, our results on economic outcomes are closer to the ones from studies in developed countries (e.g., Jacob & Ludwig (2012), Oreopoulos (2003), Jacob (2004) and Chetty et al. (2016)). However, opposite to these studies, where families were induced to move to better neighborhoods, we find null effects on economic outcomes in a setting in which households move to more isolated neighborhoods. Taken together, these findings suggest that improvements in housing quality and declines in housing costs might be compensating for increased isolation.

This paper further contributes to the literature on slums (e.g., Marx et al. (2013)). This literature emphasizes that closeness to city centers is important to ensure that poor households benefit from urban living. This is consistent with urban economic models, which predict that employment and income would decline in response to increases in the distance to job opportunities (e.g., Alonso et al. (1964)). However, our findings show that living close to city centers provided negligible benefits for households in employment and income. A possible explanation is that our effects are estimated for a set of households who did not have good labor market prospects.

Finally, this paper adds to the literature evaluating the consequences of the MCMV program. Other studies use the lotteries from Rio de Janeiro and some other selected municipalities to examine the program's effects on formal employment using data from RAIS (a matched employer-employee registry of employment). Mata & Mation (2018) and Chagas et al. (2019) find small negative effects of the MCMV on formal employment. Pacheco (2019) also finds negative effects of the MCMV on formal employment right after the houses are delivered but finds these effects revert in three years. She further documents the program moves households further from job opportunities. Leape (2020) extends the former work adding more lotteries the took place in Rio de Janeiro (but balancing the treatment and control groups using a propensity-score method) and finds that moving to a MCMV unit increased the likelihood of employment by 2% after four years.[3]

[3] Less related to our work, Teixeira (2019) examine the effects of the MCMV on credit access, and Bueno et al. (2018) on political preferences.

This literature focuses on outcomes measured using RAIS because it is hard to follow the lotteries' participants through time using other datasets. We complement this literature on two dimensions. Methodologically, we provide evidence of how it is possible to use restrictions to eliminate selection of the MCMV beneficiaries into and out of the *Cadastro Único* and follow them through time in this dataset. This enables us to eliminate some of the differences between lotteries' winners and losers documented in the existing literature and conduct more thorough randomization checks.[4] Empirically, we examine the effects of the program on more outcomes. This enables us to document effects on housing quality, housing costs, formal and informal employment, and enrollment not previously studied in this program's literature.[5]

The rest of this paper is divided into six sections. Section 2 describes the institutional background of the program. Section 3 describes the expected effects of the program. Section 4 presents the data, describes the sample selection, and tests the balance of the sample. Section 6 present the results. Section 7 concludes the paper.

[4]Mata & Mation (2018) check randomization only for demographic characteristics and did not find balance on two of the six lotteries they analyze, while Chagas et al. (2019) documents *Cadastro Único* outcomes are not balanced in the lotteries he studies.

[5]This comes at the cost of restricting attention to lotteries' winners and losers of the MCMV who are also beneficiaries of the *Bolsa Família* program. This sub-population is probably poorer than the population targeted by the MCMV. We discuss this in detail in Section 6.

2 Institutional Context

This section describes the institutional background of the *Minha Casa Minha Vida* (MCMV) program focusing on the features relevant to our empirical investigation. Appendix A provides a detailed description of the program.

2.1 The *Minha Casa Minha Vida Program*

The *Minha Casa Minha Vida* (MCMV) program was created in the late 2000s to provide housing for low and middle-income households in Brazil. In the period 2009-2018, the program financed the construction of 3.95 million houses at a total cost of about R$430 billion (US$80.3 billion at the current exchange rates) (PAC, 2018).[6]

The MCMV offered different types of subsidies for households in different segments determined according to their income levels.[7] Its resources came from the federal government budget and are managed and channeled mainly by *Caixa*, a stated-owned bank specialized in mortgage financing. This bank is responsible for certifying construction companies, contracting the development of housing projects, and providing subsidies for eligible households.[8]

Roughly 30% of these houses were built and sold to households in the program's segment 1 (income up to R$ 1,600 per month or US$ 320 at the current exchange rates). Subsidies for households in this segment could go up to 90% of the unit cost. There were no down payment requirements, and monthly installments were capped at 5% of the household income (or R$ 25). Because the demand for houses built for households in this segment typically exceeded the supply of units, municipalities organized lotteries to select its

[6] Appendix Table A1 reports information on the number of units financed for urban poor households by the MCMV until 2017.

[7] Appendix Table A2 describes these subsidies in detail as well as the average value of the units built for these different groups.

[8] Table A4 reports segment 1 corresponds to 32% of the of units (constructed and under construction) and 20% of the investment in the MCMV Program. It further reports that the typical value of a unit of segment 1 is roughly 50% of the typical value of a unit in segment 2 or 3.

beneficiaries.[9]

Housing units built through MCMV must have at least two bedrooms, a living room, a kitchen, and a bathroom. Its minimum surface area is 37 m^2 (roughly 400 sq^2). These houses are built in housing projects which have from a few dozen to more than a thousand units. The location of these projects must comply with minimum requirements regarding environmental planning, sewage treatment, connection to the electricity grid and the water network, access roads, and public transportation.[10]

Households were required to register for the lotteries either online or at municipal offices. In theory, households in the *Cadastro Único* eligible for the program were registered automatically. However, in practice, it is unclear whether local governments (responsible for selecting the beneficiaries) registered these households. Indeed, we observe the that not all households eligible for the MCMV observed in the *Cadastro Único* were in the lists of lotteries. When housing projects were close to completion, local officers organized the lottery among registered households to select beneficiaries. People forcibly displaced from their homes or individuals with disabilities are prioritized in separate lotteries. Apart from this, the allocation mechanism is straightforward. If the last two digits of the participant's registration number matched the Federal Lottery draw's last two digits, the household is selected. When the units' construction finishes, winners of the lotteries are invited to sign their contracts with *Caixa*. At this moment, officials check the household's eligibility. Importantly, units cannot be legally sold or rented.

2.2 The Lotteries in the Municipality of Rio de Janeiro

Our investigation focuses on the lotteries in Rio de Janeiro, Brazil's second-largest municipality, with 6.7 million inhabitants. This municipality received 2% of the units and

[9]Housing lotteries were implemented by Ordinance n. 140 from the Ministry of Cities enacted on 2010.
[10]Appendix Figure A2 shows pictures of a typical house plan, building and surroundings of a MCMV project.

2.45% of the MCMV program's total funding in the period 2009-2018. In total, 27,843 low-income households received units from the program. This corresponds to more than 1% of the number of households of the municipality. The typical unit built in Rio de Janeiro was priced at about R$ 62,566 and had square footage of about 45 m^2.[11] We focus on this municipality for two reasons. First, Rio de Janeiro publicly released the records of the winners and losers of the lotteries that took place in the city, giving us a comparable control group to estimate causal effects. Second, the municipality is known for the long-lasting prevalence of slums in the surroundings of its most important neighborhoods (Perlman, 2010; Monteiro & Rocha, 2017).

A total of 12 lotteries occurred in the period 2011-2015.[12] Three lotteries occurred in 2011, one in 2012, two in 2013, none in 2014, and six in 2015. These 12 lotteries selected households to live in 31 different projects delivered between the years of 2012-2018.[13] Demand exceeded the supply of units in all lotteries with 0.1% to 4.2% of the subscribers being selected. Winners were contacted by phone or letter by public officials who offered them a house of the program. Typically, there was a period of one to two years between the lotteries and the signing of the contracts.

Figure 1, Panel A depicts the location of these housing projects. It further shows the spatial distribution of the households who register for the lotteries. The projects are concentrated in the municipality's western neighborhoods while the subscribers come mostly from the municipality's western and northern neighborhoods.[14]

While we observe data from all lotteries from 2011, 2012, 2013, and 2015, our invest-

[11] We obtain the average value of the unit for Rio de Janeiro from the dataset of contracts provided by *Caixa*. The average value for Brazil is R$ 59,509

[12] Here we focus on general lotteries. There were specific lotteries for disabled and elderly, and for people living in risk areas. Information on the subscribers is public available in the website `http://www.rio.rj.gov.br/web/smhc/menu-minha-casa-minha-vida#`.

[13] Appendix Table A.3 presents details on the lotteries and the projects

[14] Appendix B reports summary statistics of projects location combining spatial data with the MCMV units location and population census data by census tract to describe the neighborhoods where the MCMV units were built.

igation focuses on the two lotteries that occurred in the year of 2013. We exclude the lotteries that occurred in 2011 due to absence of pre-program information of their winners and losers. This occurs because we start observing information on demographic and economic outcomes of most of the households in 2012. However, this is the year in which the winners of the 2011 lotteries received their units. Because we do not know exactly when households moved to the units (we approximate these dates using date on contract signature), it is not possible to test whether the winners and losers of the lotteries were comparable before the MCMV was implemented.

We exclude the lotteries that occurred in 2012 and 2015 due to a combination incomplete information and implementation problems. The 2012 housing project was invaded before the units were delivered (see Appendix A for details) and the municipality included the housing project in a later lottery that took place in January 14, 2015. The lists of subscribers of the 2015 lotteries have a number of problems (e.g., repeated identifiers, incomplete names – see Appendix A for details). This puts into question the integrity of the lotteries.[15]

Figure 1, Panel B depicts the location of the projects allocated in the 2013 lotteries and the origins of its subscribers. As most MCMV projects implemented in the municipality, the projects are located in Rio de Janeiro's western zone. The origins of the subscribers are also almost identical to the origins of the subscribers in general. This reflects the fact that the list of participants does not change much. Indeed, a large share of the households participate in multiple lotteries. Te similarities between the projects' location and the subscribers of the 2013 lotteries and the rest of the lotteries indicates the effects from these lotteries is informative of the effects of the MCMV in the municipality in general.

[15] Appendix D provide evidence it is possible to balance the lists of winners and losers of the 2011 and 2015 lotteries by resorting excluding beneficiaries not observed before the treatment and lotteries with more implementation problems. It further shows that including these lotteries does not influence our results.

3 The Expected Effects of the *Minha Casa Minha Vida*

Before we present the results, we briefly discuss the expected effects of the MCMV projects. We begin by laying out the expected effects of this program on neighborhood quality, housing quality, and housing costs. We then discuss how economic theory and existing evidence suggest effects on these three dimensions influence economic choices.

The geographic concentration of the MCMV projects in the peripheries documented in Section 2 suggests the program might induce households to move to worse neighborhoods. An influential body of work discusses the influence neighborhoods – through their influence on individuals' preferences, social connections, and access to public services – exert on outcomes like schooling, labor force participation, and income (Jacobs, 1970; Wilson, 1987; Jencks & Mayer, 1990; Massey & Denton, 1993; Kowarick, 2002). Recent empirical studies exploring exogenous improvements in neighborhood quality induced by housing programs in the U.S. document that better neighborhoods improve the educational outcomes of children (Chetty et al. (2016) and Chyn (2018)) but do not improve (and might even deteriorate) the labor market outcomes of adults (e.g., Katz et al. (2001), Kling et al. (2007), Jacob & Ludwig (2012), Ludwig et al. (2013)). Thus, to the extent that the MCMV induces households to move to poorer and more distant neighborhoods, we expect this mechanism to (weakly) deteriorate both the educational outcomes of children and adults' labor market outcomes.

However, the houses built under the MCMV program might be of better quality than the houses poor households typically reside in the slums of Rio de Janeiro. There is a long line of empirical studies that discuss how houses with proper sanitation, lighting, and ventilation, might positively affect individual health and well-being (e.g., Galiani & Schargrodsky (2004) and Kling et al. (2007)). In particular, access to proper sanitation has an essential role in reducing the incidence of communicable diseases due to oral contamination (Cutler & Miller, 2005; Alsan & Goldin, 2019). These effects on health might positively affect the school achievement of children (e.g., Miguel & Kremer (2004), Bleakley (2007)) and the labor supply and earnings of adults (e.g., Currie & Madrian (1999)). Thus, to the extent that the MCMV induced households to move to better houses, we expect this mechanism (weakly) to improve both the educational outcomes of children and the labor market outcomes of adults.

Moreover, the MCMV relieves households of the burden of rent. This reduction in housing expenditures represents an increase in non-labor income. Households will respond to this increase by investing more in human and physical capital and decreasing their labor supply. Empirical evidence from other settings suggest poor households typically respond to increases in non-labor income by investing more both on human and physical capital (e.g., Gertler et al. (2012) and Blattman et al. (2014)). In line with these studies, Kumar (2019) finds positive effects on human capital investments of subsidized housing lotteries in Mumbai (India) despite these lotteries induced households to move to worse neighborhoods. She interprets these effects as a consequence of income effects esteeming from the program's subsidies. Empirical evidence also suggests that poor households respond to increase in non-labor income by reducing their labor supply (e.g., Alzúa et al. (2016)).[16] Thus, to the extent that the MCMV reduces housing expenditures, we expect this mechanism to (weakly) improve the educational outcomes of children and to (weakly) deteriorate the labor market outcomes of adults.

We expect the effects of the MCMV on economic outcomes to reflect the combination of the neighborhood, house, and income effects described in the previous paragraphs.[17] We use our extremely detailed data on neighborhoods and households to investigate how

[16] It is important to note that in the presence of credit constraints, households might increase their labor supply in response to increases in non-labor income. See Banerjee et al. (2020) for evidence on this.

[17] A fourth mechanism highlighted in the literature is that housing programs influence economic decisions by changing tenure security. Moving from slums with poorly defined property rights to projects with well-defined property rights might influence households' economic decisions as suggested by evidence from previous land titling programs (Field, 2007; Galiani & Schargrodsky, 2010). However, we do not expect this mechanism to be relevant in our setting since households cannot sell or formally rent the houses built under the MCMV program.

the MCMV indeed influenced neighborhood quality, housing quality, housing costs, and their combined effect on school enrollment of the children, labor force participation of the adults, and the household's overall income.

4 Data Construction

4.1 Data Sources

We use data from multiple data sources. To obtain information on housing quality, expenditures, enrollment, labor force participation, and income of the lotteries' subscribers, we combine publicly-available records from the MCMV lotteries with microdata of the beneficiaries of the *Bolsa Família* program which registered in the program and, therefore, entered in the *Cadastro Único* (Brazil's unified registry of beneficiaries of social policies) before the MCMV started. To understand if and when the households selected in the lotteries program signed a contract to purchase houses constructed for the MCMV program, we match this data with official data of the program's contracts obtained from *Caixa*. Furthermore, to generate information on neighborhood quality, we merge information of the location of the households with neighborhood-level characteristics computed using the 2010 Population Census and information on distance to jobs and education institutions from Pereira et al. (2020). We describe each of these data sources in detail below.

Lotteries. Rio de Janeiro's municipal government provides the list of participants (winners and losers) in the form of PDF files.[18] It is the main source of information on "treated" and "non-treated" individuals we use in the research. We digitized this data to obtain the CPF (taxpayer registration number), the full name, and the treatment status of each individual who subscribed in the lottery. We further use the lottery records to obtain the name of the housing projects with units allocated through each lottery.

Cadastro Único. The Ministry of Citizenship's unified register of social beneficiaries provides demographic and economic information of the low-income population of Brazil. It is the main source of information on demographic, households and houses' characteristics, expenditures, and economic outcomes we use in this research.

[18]See `http://www.rio.rj.gov.br/web/smhc/menu-minha-casa-minha-vida#`.

The *Cadastro Único* was created in 2001 and, since 2003, it is the tool used for identifying and monitoring beneficiaries of the *Bolsa Família* program (Mostafa & Sátyro, 2014). While its focus is on the *Bolsa Família* beneficiaries, this registry is increasingly used to identify and monitor beneficiaries of other programs. Currently, more than 20 programs run by the Federal Government and numerous programs run by local governments use the *Cadastro Único* to track their beneficiaries.

In 2018, this dataset contained information of 23 million households, 13.8 million of which were beneficiaries of the *Bolsa Família Program*. Its data is grouped into six categories: personal identification, household identification, household characteristics, schooling, work, and income. Supplementary information on expenditures, participation in social programs, and vulnerability (homeless, engaged in child labor, etc.) has been collected either for particular groups of households or at specific dates.

We use *Cadastro Único* extractions for the period 2012-2018. Each extraction is a cross-section including the most recent information of each household from the registry. Importantly, it contains the date in which the information was updated. This will be essential to enable us to follow households included in this dataset over time. At the individual-level, we obtain the following information from the *Cadastro Único*: household identifier, CPF (taxpayer registration number)[19], NIS (social registration number), full name, demographic information (age, sex, marital status etc.), enrollment, and employment. At the household-level, we obtain the following information: household identifier, date of the update house characteristics, participation in the *Bolsa Família*, expenditures, and income per capita.

To ensure comparability between our treatment and control groups, we focus on beneficiaries of the *Bolsa Família* program (Brazil's flagship social policy). Because households must update their information every 24-months to continue eligible for the federal pro-

[19]The CPF identifying is missing for almost 50% of the individuals registered

grams, this enables us to follow a group of households over time. This will be typically the case of the households receiving the *Bolsa Família* program. Figure 2 provides evidence of these households update their information much more frequently than the other households.

Contracts. *Caixa* provides information on the mortgages signed by the beneficiaries of the MCMV program. We obtain the NIS (social registration number), the date the contract was signed, the value of the mortgage, the subsidy, and the name and address of the housing project of each beneficiary of the MCMV program. The information on the address is incomplete, but we retrieve the complete address using geocoding tools from Google Maps Geocoding API.

Neighborhoods. We use tract-level information on demographic, economic and tract characteristics from the 2010 Population Census to examine the average characteristics of the neighborhoods in which "treatment" and "control" households are located both before and after the MCMV. We extract the following indicators from the census: the share of poor households, the average household income, the share of black individuals, the average schooling, and the share of households located in streets with paved roads, the share of households located in streets with garbage collection, and the share of households located in streets with open sewage.

We also explore data on access to job opportunities and educational institutions constructed by Pereira et al. (2020). These authors combine data on firm location coming from the Ministry of Labor's administrative data (*RAIS*) and the Ministry of Education's administrative data (*Censo Escolar*) with geo-coded timetables of public transportation to built a dataset containing information on the share of employment opportunities and educational facilities accessible in 30, 60, 90, and 120 minutes at a 200m $\times$ 200m resolution. We aggregated this data at the neighborhood level.

4.2 Data Linkages

Our procedure to match the different sources of data described in the previous section has four different steps.

In the first step, we use the CPF and the name to match the lotteries' records with data from the *Cadastro Único*. We define a household as being treated if an individual of this household participated and won one of the MCMV lotteries. We define a household as being non-treated if an individual of this household participated but lost one of the MCMV lotteries.

There are two concerns with this match. The first is that the CPF is optional in the *Cadastro Único*, available for about 50-55% of the individuals. We mitigate this problem using both the CPF and the names of the individuals to match these datasets. The second concern with the match is that the *Cadastro Único* does not contain information of the full set of individuals who subscribed to the lotteries. In theory, participants must enroll in the *Cadastro Único* to participate in the lotteries. However, in practice, participants were only required to register in the *Cadastro Único* to receive the houses from the program.[20] This endogenous selection into the *Cadastro Único* might unbalance the characteristics of households in the treatment and control groups, threatening our research design. We deal with this issue by focusing on the beneficiaries of the *Bolsa Família* program who entered into the *Cadastro Único* before the first MCMV lottery in Rio de Janeiro (June 11, 2011). These restrictions ensure treatment and control households were randomly allocated in the lotteries' records and the records matched with the *Cadastro Único*. Using these restrictions, we obtain a match rate of 10.5% for the lotteries that occurred in 2013. For the lottery that occurred in October 2013, we find 60 treated households and 49,654 control households. For the one that occurred in December 2013, we find 297 treated units and

[20]For instance, our data indicates that the probability of finding households in the treatment group in any extraction of the *Cadastro Único* is much higher than the likelihood of finding households in the control group.

51,394 control units.

In the second step, we use the NIS and the name of the individuals to match the our data with the contract-level information of the MCMV.[21] We define a household as having received a house from the program if a member of the household signed a contract with *Caixa* to receive a house built by the MCMV. We found that about 57% of the treated households signed contracts to purchase the program's units.

In the third step, we use neighborhood codes from IBGE to match our data with data on neighborhood quality from the 2010 Population Census. In the baseline, there are no neighborhood codes in the *Cadastro Único* for about 15% of our data, implying we do not have neighborhood information for them. However, we do have neighborhood information for most of our sample in the post-treatment period. This is due to improvements in the *Cadastro Único* information over time.

In the fourth step, we use the more recent information to build a panel dataset containing information of the treatment and control households pre and post-treatment. For each household, we define the pre-treatment (post-treatment) periods as the observations with information updated prior to (after) April 27, 2015. This is the last date of delivery of the projects for which the households in our sample subscribed. Our empirical analysis will typically focus on the first and the last observations of each household.

Our empirical investigation focuses on *Bolsa Família*'s beneficiaries which, as reported in Figure 2, help us to track the households in our sample through time. However, there is still some attrition in our data. Indeed, we do not find roughly 30% of the households from our matched dataset. Figure 3 explains the sample construction. The black rectangle denotes the *Cadastro Único* and blue rectangle denotes the lotteries. The gray area represents the final match.

After linkages, we are left with 50 treatment (34,883 control) units for the lottery from

[21] All individuals registered in *Cadastro Único* have a NIS.

October 2013 and 224 treatment (36,080 control) units for the lottery of December 2013. We merge the lists and remove observations which appear in both. Because the lists of subscribers are remarkably similar, this reduces a lot the number of observations in our sample. Our final sample has 36,470 households which observe in the pre and post-treatment periods. The treatment group has 274 observations and the control group 36,196 observations in each period. Figure 4 reports a histogram of the year of the updates in the first pre-treatment period (which we use to test the balance of our sample) and last post-treatment period (which we use to test the MCMV's effects). Information from the pre-treatment period typically comes from the year 2011, while information from the post-treatment period is divided in the years 2015-2018 (less than 12 months, 12-24 months, and more than 24 months after the treatment).

5 Empirical Framework

Random assignment implies it is possible to estimate the effects of the MCMV comparing the outcomes of winners and losers of the program's lotteries. Our baseline specification estimates these intent-to-treat (ITT) effects using the following equation:

$$y_i = \alpha + \beta T_i + \gamma \mathbf{X}_i + \epsilon_i, \tag{1}$$

in which y_i is an outcome of interest for household i, T_i is a dummy indicating whether a member of the household i was offered a housing built under MCMV program, $\mathbf{X}_i$ is a vector of pre-determined controls included to improve precision, and ϵ_i is a error term.

The coefficient of interest in equation (1) is β. This coefficient captures the effect of being offered a house built under the MCMV program on outcome y_i. This coefficient is identified under the hypothesis that the lotteries were well-implemented, that is, T_i is not correlated with ϵ_i.

Equation (1) is estimated using household's i most recent information. However, households update their information in different periods, implying their exposure to the program is different. Thus, it is possible to estimate the intent-to-treat (ITT) for different exposures using the following equation:

$$y_i = \alpha + \sum_{s \in S} \beta_s \left(T_i \times E_{is} \right) + \gamma \mathbf{X}_i + \epsilon_i, \tag{2}$$

in which E_{is} is a dummy denoting whether the household was exposed to the program in the intervals $S = \{$0-24 months, 24+ months$\}$. This coefficient is identified under the hypotheses that the lotteries were well-implemented and the timing in which households are observed is exogenously determined. These hypotheses imply $T_i \times E_{is}$ is not correlated with ϵ_i.

5.1 Multiple Testing

Due to the large number of outcomes of interest used in our empirical investigation, inference using conventional methods will probably generate over-rejection of the null hypotheses. This happens because the FWER (Family-Wise Error Rate) – the probability of rejecting at least one null hypothesis when all the null hypothesis are true – increases when the number of hypotheses increases.

To deal with this issue, we use Romano-Wolf step-down procedure to obtain p-values corrected for multiple testing (Romano & Wolf, 2005a,b, 2016). This procedure is a more powerful method to control the FWER than other procedures (e.g., Bonferroni and Holm) that assume the test statistics are independent because it uses re-sampling (bootstrap) to incorporate information about the joint dependence structure of the test statistics of the different hypothesis being tested. It is also more general than other procedures based on re-sampling (e.g., Westfall et al. (1993)) because its algorithm is more general.

The Romano-Wolf correction has been increasingly used in empirical work (e.g., Mazzocco & Saini (2012); Gertler et al. (2014); Olken et al. (2014); Attanasio et al. (2017)). We employ this correction for the groups of outcomes for which we have more than one indicator (neighborhood quality, housing quality, housing costs). The corrected p-values should be interpreted as the significance level that would have to be applied to the entire family of hypotheses if we were to accept the null that the effect is zero.

5.2 Balance

To test the integrity of our research design, we estimate equation (1) using pre-treatment indicators. The results are reported in Table 1. We test for pre-treatment differences in five different groups of outcomes: demographics, neighborhood and neighbors characteristics, house characteristics, housing costs, and enrollment and employment. For each outcome, we compute the mean of the outcome for the control group (column 1), the mean of the outcome for the treatment group (column 2), and the mean differences between the treatment and control groups (column 3). For each group of outcomes, we further report a test of joint significance of the differentes between the treatment and control groups.[22]

Table 1 reports that pre-treatment differences were negligible for all groups of outcomes. Panel A reports pre-treatment differences in demographic characteristics. There are almost no prime-aged males in our sample. More than 90% of the households are headed by females and spouses are present in only 20% of them. Heads have 39 years on average, children under the age of 6 years are present in 45% of the households, and the average household size is about 3.80. Mean differences between the treatment and the control group are economically and statistically irrelevant for all but one outcome. The exception is the number of dwellers which is 0.16 higher in the treatment than in control group. This difference is statistically significant at the 10% level. However, a joint test of significance rejects the hypothesis that these differences are jointly different from zero.

Panels B and C report pre-treatment differences in neighborhood and housing quality. Households lived in neighborhoods with close to 100,000 inhabitants, where almost 80% of the households were connected to the sewage network, 44% of the inhabitants were white, and their income was slightly over R$ 1,400 per month. Their houses had 3.8 rooms and 1.3 dorms. 98% of them had a bathroom, and 56% had either wood or tile floor. Mean differences between treatment and controls groups are minor and not significant at the usual statistical levels.

Panel D reports pre-treatment differences in housing costs. Households spent approximately R$54 on rent, R$20 with electricity, R$37 on gas, and R$ 5.5 with water and sanitation. Panel E reports pre-treatment differences in school enrollment, labor force participation, and income. 90% of the children with 6-17 years were enrolled at school. Comparable

[22]To test the joint significance of the mean differences between the treatment and control groups for each group of outcomes, we estimate the mean differences jointly using Seemingly Unrelated Regression (SUR) and test whether the differences are jointly different from zero using a chi-squared statistic.

figures are found for males and females. 51% of the (female) heads of the households with 25-64 years were employed. Mean differences between treatment and controls groups are minor and not significant at the usual statistical levels.

Taken together, the evidence from Table 1 indicates that the characteristics of the treatment and control groups of the 2013 lotteries are balanced. As discussed before, it is harder to confirm balance for the other MCMV lotteries due to a combination of missing information and implementation issues. However, in Appendix D (Tables D1 and D8), we provide evidence that the 2011 and the 2015 are reasonably balanced once we exclude households without pre-program information and lotteries with problems in the implementation.

6 Results

We present our results in three parts. We begin by discussing MCMV's take-up. We then document the program's effects on neighborhood quality, housing quality, and housing costs. Finally, we examine whether these changes in neighborhood quality, housing quality, and housing costs influence labor force participation of adults, income, and schooling decisions.

6.1 Take-up

The first row of Table 2 reports the share of the control households who signed contracts with *Caixa* (column 1), the share of treated households who signed contracts with *Caixa* (column 2), and the difference in these probabilities (column 3). The control mean is 0% and the treatment mean is 54%. This 54 p.p. difference is highly significant, thereby indicating the lotteries did increase the probability the households benefits from the MCMV program. Furthermore, it is worth noting the control mean indicates there are no "always takers" of the 2013 lotteries, while the treatment mean indicates that "never takers" amount to 46% of the subscribers of these lotteries.

While useful, the contract-level information does not enable us to examine whether the households who signed contracts indeed live in the units built by the MCMV. Indeed, it is possible that households abandon the MCMV houses after a period. Moreover, the contract-level information does not enable us to understand whether the program moved the households to other neighborhoods or simply re-located them in their neighborhood or origin.

Thus, we investigate other measures of take-up built using information of the households' location. The remaining rows of Table 2 reports the results. 6-7% of the treatment and control households lived in the neighborhoods of *Cosmos* and *Santíssimo* (where the units allocated in the 2013 lotteries were built) before the program. This probability increases to 8% for the controls households and to 40% of the treatment households after the program. We further document that 5% of the control households and 35% of the treatment households moved to these neighborhoods in the period of analysis. These measures indicate a take up of 31-32 percentage points.

One concern with policies that move households to other neighborhoods is whether households quit the program as their exposure to it increases. We estimate equation (2) to obtain the effects of the MCMV on the probability of moving to the neighborhoods in which the program's houses were built in different time horizons. As shown in Figure 5, the probability of moving to these neighborhoods increases with the program's exposure. Hence, there is no evidence of program exit. This contrasts with the evidence of increasing program exit over time documented by Barnhardt et al. (2017) for a program in Mumbai (India).

Together, our measures indicate the program's take-up is between 31-54%. These figures are in the range of the rates estimated in the literature. They are higher than the 19%-48% take-up of housing vouchers found in studies analyzing experiments in the U.S. (Rubinowitz & Rosenbaum, 2002; Kling et al., 2007; Jacob & Ludwig, 2012) but are lower than the 66% take-up of houses found in Mumbai (India) (Barnhardt et al., 2017).

6.2 Neighborhood Quality, Housing Quality, and Housing Costs

Table 3 reports the effects of the MCMV on neighborhood quality. Panel A examines whether the program moved households to different neighborhoods, Panel B examines whether the program moved households to locations with different neighbors, and Panel C reports whether the program moved households to locations closer or further from job opportunities and schools. Each row reports the results for a different outcome of interest. Column 1 presents the mean of the control group, column 2 reports the results of a bivariate regression of outcome of interest on a treatment indicator, and column 3 adds a control for the outcome measured in the baseline. Inference is based on Romano-Wolf *p*-values corrected for multiple testing.

Panel A shows that the MCMV moved households to neighborhoods that were less populated. The population of the neighborhoods in which the households from the treatment group reside is roughly 12-17,000 smaller than the population of the neighborhoods in which the households from the control group reside. However, the neighborhoods in which treatment and controls households live are no different in terms of access to the sewer network and slightly better in terms of access to the water network. Indeed, using a neighborhood index combining these three variables[23], we find only weak evidence the program moved families to worse neighborhoods in terms of population and infrastructure. The differences between the treatment and control group are small, being significant at the 10% level in the specification without controls and not significant in the specification with controls.

Panel B shows that the MCMV moved households to neighborhoods with a lower share of white residents, labor force participation, and income. The typical head of the household of the neighborhoods in which treatment units live is 2.1 p.p. less likely to be white, 1.1 p.p. less likely to work, and receives R$ 104-106 less than the typical head of the household of the neighborhoods in which control units live. Indeed, using a neighbors index combining these four variables, we find that the program moves households to neighborhoods with resident with worse outcomes.

The findings from Panels A and B indicate that the the MCMV moved households to neighborhoods less populated and with worse socioeconomic according to the 2010 Population Census. However, one concern with these measures is that they do not capture the changes in these neighborhoods that might have occurred after 2010. While this has

[23]We perform a principal component analysis of the neighborhood variables, and and define the neighborhood index as the first principal component. We use the same strategy to build the other indexes used in this section.

the benefit of insuring re-location of treatment households contaminates neighborhood quality, it has the cost of not picking up other changes in neighborhood quality that might have happened after 2010.

Panel C deals with this issue by examining whether the program moved households to neighborhoods closer or further to job opportunities and schools in the post-treatment period. We find the treatment group lives in neighborhoods in which the share of job opportunities (schools) accessible (using public transportation) in less than 90 minutes is 4-6 p.p. (1-3 p.p.) smaller than the in the neighborhoods in which the control group lives. For job opportunities, these results are statistically significant at the 1% level in all specifications while, for educational facilities, these results are statistically significant at the 1% level using controls and at the 10% level using controls.

Together, these results provide robust evidence that the MCMV induced households to re-locate to worse and more distant neighborhoods. This movement is predicted to decrease the consumption of amenities and decrease income net of transportation costs (Alonso et al., 1964; Straszhem, 1987; Glaeser, 2000). This indicates that consumption-maximizing households will choose to move to worse and more distant neighborhoods only if increases in housing quality or decreases in housing costs compensates them for the losses in the consumption of amenities and income.

Tables 4 and 5 test these conjectures. Table 4 reports the effects of the MCMV on housing quality. It tests the effects of the program on the following outcomes: a dummy indicating whether the house has wood or tile floor, a dummy indicating whether the house's sidewalk is paved, a dummy indicating whether the house is connected to the sewage network, the number of rooms, the number of bedrooms, a dummy indicating whether the house has metered electricity, and a housing index. The structure of the table is identical to the structure of Table 3.

We find the program moves households to better houses: the probability of having a wood or tile floor increases by 7 p.p. ($\approx$ 12% of the control mean), the probability of having a paved sidewalk increases by 7 p.p. ($\approx$ 8% of the control mean), and the probability of being connected to the sewage network increases by 3 p.p. ($\approx$ 3% of the control mean). There is also evidence the MCMV moves households to bigger houses. The number of bedrooms increases by 0.19-0.20 ($\approx$ 15% of the control mean) and the number of rooms increases by 0.34-0.35 ($\approx$ 10% of the control mean). This provides compelling evidence that the program positively affected the quality of the house in terms of size and construction materials and moved families to places with better infrastructure. Furthermore, we find the program also induced the formalization of the households by increasing the probability of the household having metered electricity (as opposed to a irregular connection) by 11-12 p.p. ($\approx$ 20% of the control mean). This reinforces the interpretation the program moves households from slums (or other informal settlements) closer to the city center to apartments located in the outskirts of the city. Together, estimates obtained using a house index built combining all outcomes discussed above indicate that the MCMV increases housing quality by about 0.34-0.39 standard deviations.

Table 5 reports the effects of the MCMV on housing costs. It tests the effects of the program on the following outcomes: expenditures with rent, expenditures with water, expenditures with electricity, expenditures with gas, and total expenditures. The structure of the table is identical to the structure of Table 3. The table shows that increases in housing quality are not the only benefit households obtain from the MCMV program. The program also reduces housing costs significantly. The treatment group spends R$37.8-43.8 per month less in rent than the control group. It is possible this decreases in rents is at least partially compensated by increases in other housing costs. For instance, the increase in the number of rooms and the decrease in illegal connections might mechanically increase the expenditures with water and electricity. Indeed, we find water expenditures are about R$ 5.7-5.8 and electricity expenditures R$12 higher in the treatment group compared to the control group. The former effect is statistically significant at the 5% level while the latter is significant at the 10% level. Interestingly, however, gas expenditures decline by about R$6.0-6.1. This is consistent with the MCMV moving households out of slums in which drug dealers and/or paramilitaries monopolize the gas distribution, thereby increasing its price.

The magnitude of the effects on water, electricity, and gas expenditures indicates the MCMV generated net decreases in housing costs. Moreover, there is no evidence that other expenditures (e.g., food) increase significantly. As a consequence, we find that the treatment reduces total expenditures by R$ 28.9-33.6. This effect is statistically significant at the 5% level. This indicates the program had a sizable income effect. To get a sense of the magnitude of this effect, it is useful to compute Wald (IV) estimates of the effects of the MCMV subsidies on expenditures. The midpoint of the take up estimates reported before is 0.42, implying the that receiving a house from the MCMV program reduces costs by almost R$75. This effect corresponds to one quarter of the household's reported expenditures.

Appendix C report the effects of the MCMV on neighborhood quality (Figure C1) and housing quality and costs (Figure C2). The effects on neighborhood quality are stable. Interestingly, the effects on the index of housing quality and on housing costs move in opposite directions. The effects on the index of housing quality increase from 0.30 to 0.45 standard deviations. The effects on housing and total costs, on their turn, decline significantly from R$ 40 (R$ 50) to less than R$ 20 (R$ 10) per month.

6.3 Labor Force Participation, Income, and Schooling

Our findings so far indicate that the MCMV moved households to worse neighborhoods but to better and less expensive houses. To understand whether theses chances influence economic choices, we examine the program's effects on female labor force participation, household income, and schooling decisions.

We begin by analyzing the effects of the MCMV on the employment of adults. We expect the deterioration of neighborhood quality and the declines in housing costs to (weakly) decrease employment and the improvements in housing quality to (weakly) increase employment. Table 6 presents the results. As discussed in section 5, we do not observe adult males in our sample of *Bolsa Família* beneficiaries. Thus, we focus on female employment. Panel A uses the labor force participation of the head of the household as the outcome of interest while Panel B uses the share of adult females that work as the outcome of interest. Both panels report results for prime-aged females in general (25-64 years) and for younger (25-44 years) and older (45-64 years) women.

We find no effects of the MCMV on female employment in general. The coefficients are close to zero in most of the specifications. Point estimates are typically positive for younger women but negative for older women. Our results are consistent to the findings of Pacheco (2019) that suggests no persistent impact of the MCMV on employment in the formal market. She finds a negative effect of the program on employment. However, this effect diminishes over time and disappears after three years.[24]

Then, we examine the effects of the MCMV on income. We expect the effects of the program on income to be qualitatively similar than the effects on employment. Table D13 reports the results. There is no effect on wages of the head of the household, total wages of the household, and income per capita. This is consistent with the absence of effects on employment in general. However, it is worth noting the effects on income represent the combined effects of changes in labor supply in the extensive margin (documented in Table 6), changes in the labor supply in the intensive margin (for which there is no information), and changes in hourly wages (for which there is no information). Thus, the null effect on income suggests the responses in these other margins are not relevant as well.

[24] Pacheco (2019) suggests that there was an adaptation by the people drawn by the program through accessibility via individual motorized transport. This evidence is corroborated by work of Mata & Mation (2018) that shows an increase in the purchase of motorcycles by MCMV participants.

The combination of the decrease in expenditures documented before with the stability in income indicates that households disposable income increased due to the MCMV. This is suggestive evidence that the program might be increasing investments (in durable goods, small businesses etc.) or savings, thereby increasing income and consumption in the long run. This increase in investments is consistent with the evidence from Gertler et al. (2012) who document that poor households respond to permanent income increases by increasing investments.

Finally, we analyze the effects of the MCMV on schooling decisions. We expect the deterioration of neighborhood quality to (weakly) decrease enrollment and the improvements in housing quality and the decline in housing costs to (weakly) increase enrollment. Table 8 reports the results. It presents the effects of the MCMV on enrollment both for children in general as well as for boys and girls separately. The gender split is motivated by previous studies that finds that girls' school outcomes respond much more than boys' to housing programs (e.g., Kling et al. (2005)) and other interventions (e.g., Anderson (2008)). Notice that the number of observations is different in each row because the number of households with school-aged children in general, boys, and girls is different. We find some evidence that the MCMV increases the enrollment of teenagers. However, this does not translate into increases in the enrollment of teenagers on high school nor on increases in high school graduation rates among young adults. These pieces of evidence indicate the program might be increasing age-grade distortion.

7 Conclusion

This paper contributes to the literature on the consequences of housing policies in low and middle-income countries by documenting the direct and indirect effects of a large-scale housing program called *Minha Casa Minha Vida* (MCMV) on poor households from the city of Rio de Janeiro (Brazil). We explore two lotteries used to select beneficiaries which took place in 2013 to investigate this program's direct effects of this program on neighborhood quality, housing quality, and housing costs and its indirect effects on school enrollment, labor force participation, and income.

Combining multiple sources of information, we are able to track winners (treatment group) and losers (control group) of these lotteries over time. We find that 34-54% of the households in the treatment group purchase and move to the highly subsidized houses built by the program. We then document the MCMV moved these households to neighborhoods that are less populated, poorer, and more distant from schools and job opportunities but to houses that are larger and of better quality. Moreover, we find that the program decreases significantly housing costs.

Turning to economic outcomes, our findings suggest that, despite moving households to more distant and isolated neighborhoods, there is no evidence the MCMV decreases children's school enrollment, adults' labor supply, and the household's overall income in the short run. This indicates that either improvements in housing quality and declines in housing costs might be compensating the deterioration in neighborhood quality or that neighborhood quality does not affect economic outcomes in the short run. Furthermore, our evidence suggests the declines in housing costs are not generating increases in current expenditures with transportation and food.

The evidence provided in this paper contributes to the growing literature on the effects of housing programs on poor households. Our findings on economic outcomes are consistent with other studies which find null effects of housing programs on labor force participation and income (e.g., Jacob & Ludwig (2012), Oreopoulos (2003), Jacob (2004) and Chetty et al. (2016) for the U.S., Barnhardt et al. (2017) for India, and Franklin (2019) for Ethiopia). However, in our setting, households moved to worse neighborhoods, while households moved to better neighborhoods in the settings of the aforementioned studies. Moreover, we find no evidence of increases in program exit in the short run and document significant gains in terms of housing quality and reduced housing expenditures not documented previously. Documenting these responses and investigating how they influence households in the long run is an important agenda for future research.

References

Alonso, W. et al. (1964). Location and land use. toward a general theory of land rent. *Location and land use. Toward a general theory of land rent.*.

Alsan, M. & Goldin, C. (2019). Watersheds in child mortality: the role of effective water and sewerage infrastructure, 1880–1920. *Journal of Political Economy*, *127*(2), 586–638.

Alzúa, M. L., Amendolaggine, J., Cruces, G. & Greppi, C. (2016). Housing subsidies, labor supply and household welfare. experimental evidence from argentina.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, *103*(484), 1481–1495.

Attanasio, O., Guarín, A., Medina, C. & Meghir, C. (2017). Vocational training for disadvantaged youth in colombia: A long-term follow-up. *American Economic Journal: Applied Economics*, *9*(2), 131–43.

Banerjee, A., Karlan, D., Trachtman, H. & Udry, C. R. (2020). *Does poverty change labor supply? evidence from multiple income effects and 115,579 bags* (Tech. Rep.). National Bureau of Economic Research.

Barnhardt, S., Field, E. & Pande, R. (2017). Moving to opportunity or isolation? network effects of a randomized housing lottery in urban india. *American Economic Journal: Applied Economics*, *9*(1), 1–32.

Blattman, C., Fiala, N. & Martinez, S. (2014). Generating skilled self-employment in developing countries: Experimental evidence from uganda. *The Quarterly Journal of Economics*, *129*(2), 697–752.

Bleakley, H. (2007). Disease and development: evidence from hookworm eradication in the american south. *The quarterly journal of economics*, *122*(1), 73–117.

Bueno, N. S., Nunes, F. & Zucco Júnior, C. (2018). Do governments make dreams come true? an analysis of the minha casa minha vida housing program.

Chagas, A., Squarize, L. & Rocha, G. M. (2019). *Housing program and social conditions impact: Evidences from minha casa minha vida program lotteries in brazil* (Tech. Rep.). University of São Paulo (FEA-USP).

Chetty, R., Hendren, N. & Katz, L. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, *106*(4), 855–902.

Chyn, E. (2018, October). Moved to opportunity: The long-run effects of public housing demolition on children. *American Economic Review*, *108*(10), 3028-56. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.20161352 doi: 10.1257/aer.20161352

Currie, J. & Madrian, B. C. (1999). Health, health insurance and the labor market. *Handbook of labor economics*, *3*, 3309–3416.

Cutler, D. & Miller, G. (2005). The role of public health improvements in health advances: the twentieth-century united states. *Demography*, *42*(1), 1–22.

Field, E. (2007). Entitled to work: Urban property rights and labor supply in peru. *The Quarterly Journal of Economics*, *122*(4), 1561–1602.

Franklin, S. (2019). The demand for government housing: Evidence from lotteries for 200,000 homes in ethiopia. In *Technical report.* Working Paper.

Galiani, S. & Schargrodsky, E. (2004). Effects of land titling on child health. *Economics & Human Biology*, *2*(3), 353 - 372. Retrieved from http://www.sciencedirect.com/science/article/pii/S1570677X04000541 (Special Issue on Child Health in Latin America) doi: https://doi.org/10.1016/j.ehb.2004.10.003

Galiani, S. & Schargrodsky, E. (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics*, *94*(9-10), 700–729.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., … Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in jamaica. *Science*, *344*(6187), 998–1001.

Gertler, P., Martinez, S. W. & Rubio-Codina, M. (2012). Investing cash transfers to raise long-term living standards. *American Economic Journal: Applied Economics*, *4*(1), 164–92.

Glaeser, E. (2000). Urban and regional growth. *The Handbook of Economic Geography*, 83–98.

Glaeser, E. (2011). *Triumph of the city: How urban spaces make us human*. Pan Macmillan.

Habitat, U. (2013). Scaling up affordable housing supply in brazil: the'my house my life'programme. *Nairobi: UNON. Retrieved March*, *29*, 2015.

Jacob, B. A. (2004). Public housing, housing vouchers, and student achievement: Evidence from public housing demolitions in chicago. *American Economic Review*, *94*(1), 233–258.

Jacob, B. A. & Ludwig, J. (2012). The effects of housing assistance on labor supply: Evidence from a voucher lottery. *American Economic Review*, *102*(1), 272–304.

Jacobs, J. (1970). The economy of cities.

Jencks, C. & Mayer, S. E. (1990). The social consequences of growing up in a poor neighborhood. In L. Lynn & M. G. H. McGeary (Eds.), *Inner city poverty in the united states* (p. 111-186).

Katz, L. F., Kling, J. R. & Liebman, J. B. (2001). Moving to opportunity in boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics*, *116*(2), 607–654.

Kling, J. R., Liebman, J. B. & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, *75*(1), 83–119.

Kling, J. R., Ludwig, J. & Katz, L. F. (2005). Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment. *The Quarterly Journal of Economics*, *120*(1), 87–130.

Kowarick, L. (2002). Viver em risco: sobre a vulnerabilidade no brasil urbano. *Novos Estudos*, *63*(1), 9–29.

Kumar, T. (2019). *The intergenerational effects of subsidizing homeownership: evidence from a program in mumbai* (Tech. Rep.). Working Paper.

Leape, J. H. (2020). *Winning the housing lottery in rio de janeiro: curse or cure?* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R. & Sanbonmatsu, L. (2013). Long-term neighborhood effects on low-income families: Evidence from moving to opportunity. *The American Economic Review*, *103*(3), 226–231.

Marx, B., Stoker, T. & Suri, T. (2013). The economics of slums in the developing world. *Journal of Economic Perspectives*, *27*(4), 187–210.

Massey, D. S. & Denton, N. A. (1993). *American apartheid: Segregation and the making of the underclass*. Harvard University Press.

Mata, D. & Mation, L. (2018). Labor market effects of public housing: Evidence from large-scale lotteries. In *Technical report.* Working Paper.

Mazzocco, M. & Saini, S. (2012). Testing efficient risk sharing with heterogeneous risk preferences. *American Economic Review*, *102*(1), 428–68.

Miguel, E. & Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, *72*(1), 159–217.

Monteiro, J. & Rocha, R. (2017). Drug battles and school achievement: evidence from rio de janeiro's favelas. *Review of Economics and Statistics*, *99*(2), 213–228.

Mostafa, J. & Sátyro, N. (2014). *Cadastro único: a registry supported by a national public bank* (Tech. Rep.).

Olken, B. A., Onishi, J. & Wong, S. (2014). Should aid reward performance? evidence from a field experiment on health and education in indonesia. *American Economic Journal: Applied Economics*, *6*(4), 1–34.

Oreopoulos, P. (2003). The long-run consequences of living in a poor neighborhood. *The quarterly journal of economics*, *118*(4), 1533–1575.

PAC. (2018). Pac 7o balanço: 2015-2018. *Brazilian federal government*, 64.

Pacheco, T. S. (2019). *Moradia, localização e o programa habitacional" minha casa minha vida" no município do rio de janeiro* (Unpublished doctoral dissertation).

Pereira, R. H., Braga, C. K. V., Serra, B. & Nadalin, V. G. (2020). Desigualdades socioespaciais de acesso a oportunidades nas cidades brasileiras–2019.

Perlman, J. (2010). *Favela: Four decades of living on the edge in rio de janeiro*. Oxford University Press.

Picarelli, N. (2019). There is no free house. *Journal of Urban Economics*, *111*, 35 - 52. Retrieved from http://www.sciencedirect.com/science/article/pii/S0094119019300245 doi: https://doi.org/10.1016/j.jue.2019.04.002

Romano, J. P. & Wolf, M. (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, *100*(469), 94–108.

Romano, J. P. & Wolf, M. (2005b). Stepwise multiple testing as formalized data snooping. *Econometrica*, *73*(4), 1237–1282.

Romano, J. P. & Wolf, M. (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters*, *113*, 38–40.

Rubinowitz, L. S. & Rosenbaum, J. E. (2002). *Crossing the class and color lines: From public housing to white suburbia*. University of Chicago Press.

Straszhem, M. (1987). The theory of urban residential location. In *Handbook of regional and urban economics* (Vol. 2, pp. 717–757). Elsevier.

Teixeira, L. I. (2019). *Essays on credit policies* (Unpublished doctoral dissertation).

UN-HABITAT. (2015). Un-habitat global activities report 2015. increasing synergy for greater national ownership. *United Nations Human Settlements Programme (UN-Habitat)*, 31.

Westfall, P. H., Young, S. S. & Wright, S. P. (1993). On adjusting p-values for multiplicity. *Biometrics*, *49*(3), 941–945.

Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. University of Chicago Press.

Table 1: Descriptive Statistics and Randomization Check

	(1)	(2)	(3)	(4)
	Control	Treatment	T-C	N
Panel A: Demographics				
Female head	0.96	0.97	0.01	36470
	[0.18]	[0.17]	(0.01)	
Age	38.72	38.27	-0.45	36470
	[9.37]	[8.80]	(0.53)	
Spouse (0/1)	0.20	0.20	-0.00	36470
	[0.40]	[0.40]	(0.02)	
Children 0-6 (0/1)	0.45	0.46	0.01	36470
	[0.50]	[0.50]	(0.03)	
Dwellers	3.80	3.96	0.16*	36470
	[1.54]	[1.52]	(0.09)	
Joint significance test (p-value)			0.633	
Panel B: Neighborhoods				
Population (in 1000s)	102.18	97.63	-4.55	31615
	[92.37]	[90.85]	(6.09)	
Sewage	0.78	0.77	-0.01	31615
	[0.21]	[0.22]	(0.01)	
Water	0.99	0.99	0.00	31615
	[0.03]	[0.02]	(0.00)	
Sh. Work (head)	0.86	0.86	0.00	31615
	[0.03]	[0.03]	(0.00)	
Avg. Income (head)	1417.8	1430.1	12.35	31615
	[725.6]	[705.8]	(47.34)	
Sh. white	0.44	0.44	0.00	31615
	[0.10]	[0.11]	(0.01)	
Joint significance test (p-value)			0.874	

Descriptive Statistics and Randomization Check (continuation)

	(1)	(2)	(3)	(4)
	Control	Treatment	T-C	N
Panel C: Housing Quality				
Wood/Tile (0/1)	0.55	0.56	0.01	36470
	[0.50]	[0.50]	(0.03)	
Sewage (0/1)	0.90	0.91	0.00	36470
	[0.30]	[0.29]	(0.02)	
Paving (0/1)	0.68	0.69	0.01	36470
	[0.46]	[0.46]	(0.03)	
Electricity (meter 0/1)	0.56	0.57	0.01	35983
	[0.50]	[0.50]	(0.03)	
Dorms	1.32	1.35	0.02	31316
	[0.87]	[0.54]	(0.04)	
Rooms	3.83	3.85	0.02	35983
	[1.53]	[1.03]	(0.06)	
Dwellers per room	1.09	1.12	0.03	35983
	[0.61]	[0.62]	(0.04)	
Joint significance test (p-value)			0.395	
Panel D: Housing Costs				
Rent	54.76	63.28	8.52	32329
	[111.38]	[127.11]	(8.30)	
Electricity	20.25	20.65	0.40	32329
	[57.58]	[33.42]	(2.17)	
Gas	37.50	36.06	-1.44	32329
	[46.44]	[11.86]	(0.77)	
Water	5.51	6.26	0.75	32329
	[16.05]	[15.74]	(1.03)	
Joint significance test (p-value)			0.366	
Panel E: Enrollment and LFP				
School enrollment (%)	0.89	0.91	0.01	32758
	[0.25]	[0.22]	(0.01)	
Female LFP (Head, 25-64)	0.51	0.52	0.01	28919
	[0.50]	[0.50]	(0.03)	
Joint significance test (p-value)			0.532	

Notes: Column 1 reports the mean of each indicator in the control group. Column 2 reports the mean of each indicator in the treatment group. Column 3 reports mean differences between the treatment groups and the control group. Column 4 reports the number of observations of each indicator. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Standard deviations are reported in brackets and robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

Table 2: Take-up

	(1)	(2)	(3)
	Control	Treatment	T-C
Signed	0.00	0.57	0.57***
	[0.00]	[0.50]	(0.03)
Lives in MCMV Neighborhood (pre)	0.06	0.07	0.00
	[0.24]	0.24]	(0.02)
Lives in MCMV Neighborhood (post)	0.08	0.40	0.32***
	[0.27]	[0.49]	(0.03)
Moved to MCMV Neighborhood	0.05	0.35	0.31***
	[0.21]	[0.48]	(0.03)
N	36196	274	36470

Notes: Column 1 reports the mean of each take-up indicator in the control group. Column 2 reports the mean of each take-up indicator in the treatment group. Column 3 reports mean differences between the treatment groups and the control group. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Standard deviations are reported in brackets and robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

Table 3: Neighborhood Quality

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Panel A: Neighborhood				
Population	102.18	-17.32***	-12.55***	36470
	(92.37)	(4.38)	(4.68)	
		[0.005]	[0.010]	
Sewage	0.775	0.007	0.010	36470
	(0.206)	(0.013)	(0.009)	
		[0.589]	[0.292]	
Water	0.986	0.005***	0.003**	36470
	(0.027)	(0.001)	(0.001)	
		[0.007]	[0.027]	
Neighborhood Index	-0.000	-0.088*	-0.048	36470
	(1.000)	(0.053)	(0.053)	
Panel B: Neighbors				
LFP (Head)	0.859	-0.011***	-0.011***	36470
	(0.034)	(0.002)	(0.002)	
		[0.002]	[0.002]	
Income (Head)	1417.77	-104.02***	-106.67***	36470
	(725.59)	(37.63)	(34.28)	
		[0.005]	[0.007]	
White (%)	0.442	-0.022***	-0.021***	36470
	(0.100)	(0.006)	(0.006)	
		[0.002]	[0.002]	
Neighbors Index	0.000	-0.253***	-0.262***	36470
	(1.000)	(0.060)	(0.060)	
Panel C: Access to Opportunities				
Jobs: 90 minutes	0.277	-0.037***	-0.061***	36470
	(0.198)	(0.012)	(0.011)	
		[0.005]	[0.002]	
Schools: 90 minutes	0.274	-0.012*	-0.031***	36470
	(0.131)	(0.006)	(0.006)	
		[0.067]	[0.002]	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses and Romano-Wolf *p*-values correcting for multiple testing in each group of outcomes are reported in brackets. ***p<0.01; **p<0.05; *p<0.10

Table 4: Housing Quality

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Wood/Tile Floor (0/1)	0.55	0.07***	0.07***	36470
	(0.50)	(0.02)	(0.02)	
		[0.01]	[0.00]	
Paving (0/1)	0.68	0.07***	0.07***	36470
	(0.46)	(0.02)	(0.02)	
		[0.00]	[0.00]	
Sewage (0/1)	0.90	0.03***	0.03**	36470
	(0.30)	(0.30)	(0.30)	
		[0.01]	[0.02]	
Dorms	1.32	0.20***	0.19***	35903
	(0.87)	(0.03)	(0.03)	
		[0.00]	[0.00]	
Rooms	3.83	0.35***	0.34***	35903
	(1.53)	(0.06)	(0.05)	
		[0.00]	[0.00]	
Electricity - meter (0/1)	0.56	0.12***	0.11***	35903
	(0.50)	(0.03)	(0.03)	
		[0.00]	(0.00)	
House Index	-0.11	0.39***	0.34***	35903
	(1.09)	(0.05)	(0.05)	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses and Romano-Wolf *p*-values correcting for multiple testing are reported in brackets. ***p<0.01; **p<0.05; *p<0.10

Table 5: Housing Costs

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Rent	54.76	-37.84**	-43.80***	36470
	(111.38)	(9.09)	(10.93)	
		[0.02]	[0.01]	
Water	5.51	5.81**	5.74**	36470
	(16.05)	(2.17)	(2.29)	
		[0.02]	[0.01]	
Electricity	20.25	12.40*	12.85*	36470
	(57.58)	(4.30)	(4.55)	
		[0.08]	[0.10]	
Gas	37.50	-6.11**	-6.00**	36470
	(46.44)	(1.43)	(1.46)	
		[0.02]	[0.03]	
Total	296.79	-28.92*	-33.59**	36470
	[199.32]	[16.85]	[16.24]	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses and Romano-Wolf *p*-values correcting for multiple testing are reported in brackets. ***p<0.01; **p<0.05; *p<0.10

Table 6: Labor Force Participation

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Panel A: Head (0/1)				
25-64	0.44	0.02	0.00	33022
	(0.50)	(0.03)	(0.03)	
25-44	0.45	0.04	0.02	19483
	(0.50)	(0.04)	(0.04)	
45-64	0.40	-0.02	-0.03	13539
	(0.49)	(0.05)	(0.05)	
Panel B: All (%)				
25-64	0.424	0.02	0.02	33581
	(0.49)	(0.03)	(0.03)	
25-44	0.434	0.04	0.04	21356
	(0.49)	(0.04)	(0.04)	
45-64	0.370	-0.02	-0.08	13975
	(0.48)	(0.05)	(0.06)	

Notes: Panel A reports the effects of the MCMV on the labor force participation of the heads of household. Panel A reports the effects of the MCMV on the labor force participation of adults in general. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

Table 7: Income

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Wage (Head)	273.30	-0.64	-2.97	35278
	(394.01)	(23.93)	(24.86)	
Wage (Household)	327.08	-3.78	0.20	35278
	(449.50)	(27.45)	(26.85)	
Income per capita	159.37	-12.54	-13.62	36470
	(199.60)	(11.97)	(11.36)	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

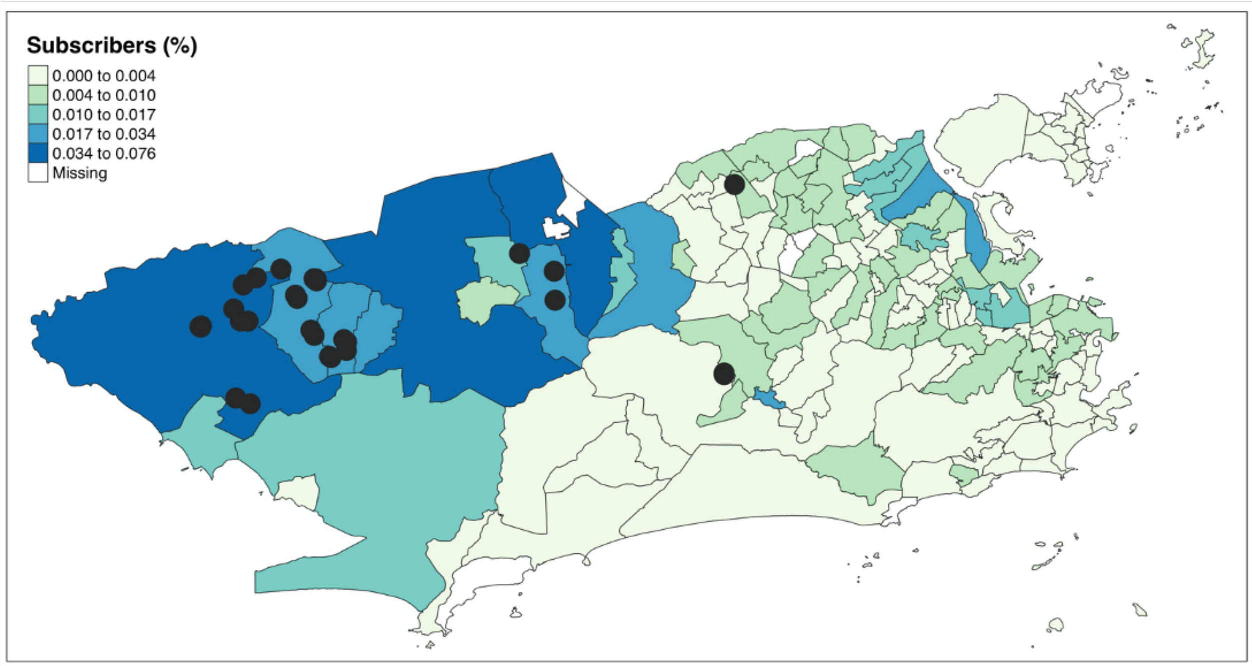
Table 8: Education

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Enrollment, 4-18	0.89	0.01	0.01	28982
	(0.25)	(0.01)	(0.01)	
Boys, 4-18	0.89	0.01	0.02*	20300
	(0.28)	(0.01)	(0.01)	
Girls, 4-18	0.90	-0.01	-0.00	19727
	(0.27)	(0.02)	(0.02)	
Pre-School, 4-6	0.62	-0.02	0.03	5639
	(0.48)	(0.09)	(0.09)	
Enrollment, 7-15	0.96	-0.00	-0.01	22788
	(0.17)	(0.01)	(0.01)	
Boys, 7-15	0.96	-0.00	-0.01	14391
	(0.19)	(0.02)	(0.02)	
Girls, 7-15	0.96	-0.01	-0.01	13908
	(0.19)	(0.02)	(0.02)	
Elementary, 7-15	0.86	0.01	0.01	22788
	(0.30)	(0.01)	(0.01)	
Boys, 7-15	0.86	0.01	0.00	14391
	(0.33)	(0.02)	(0.02)	
Girls, 7-15	0.86	0.01	0.00	13908
	(0.33)	(0.02)	(0.02)	
Enrollment, 16-18	0.94	0.03**	0.07***	13845
	(0.23)	(0.01)	(0.01)	
Boys, 16-18	0.94	0.04**	0.06***	7610
	(0.23)	(0.01)	(0.01)	
Girls, 16-18	0.94	0.02	0.09***	7111
	(0.23)	(0.02)	(0.02)	
High School, 16-18	0.33	0.03	0.10	13845
	(0.46)	(0.05)	(0.10)	
Boys, 16-18	0.29	-0.05	-0.14	7610
	(0.45)	(0.06)	(0.14)	
Girls, 16-18	0.36	0.02	-0.04	7157
	(0.47)	(0.06)	(0.13)	
High School Graduate, 19-24	0.32	0.00	-0.07	12828
	(0.45)	(0.05)	(0.09)	
Boys, 19-24	0.26	-0.02	-0.20	7015
	(0.43)	(0.06)	(0.15)	
Girls, 19-24	0.36	0.02	-0.04	7157
	(0.47)	(0.06)	(0.13)	

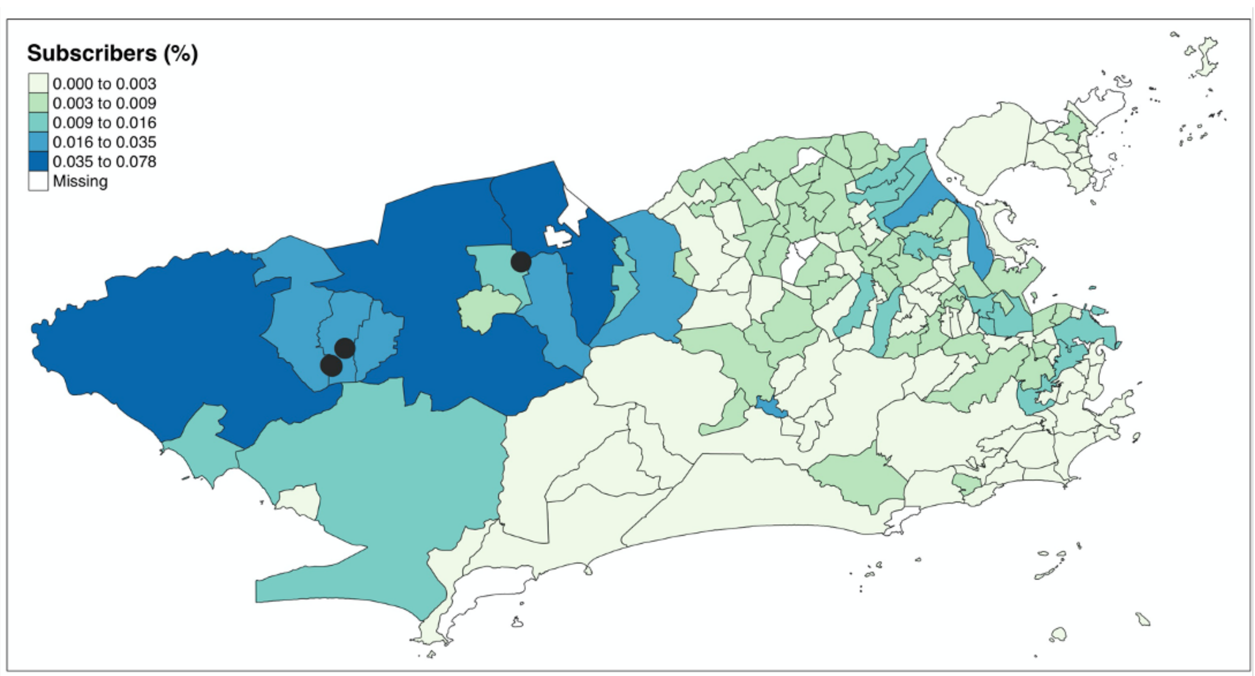
Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

Figure 1: Projects' Location and Origins of Subscribers

(a) **All lotteries**

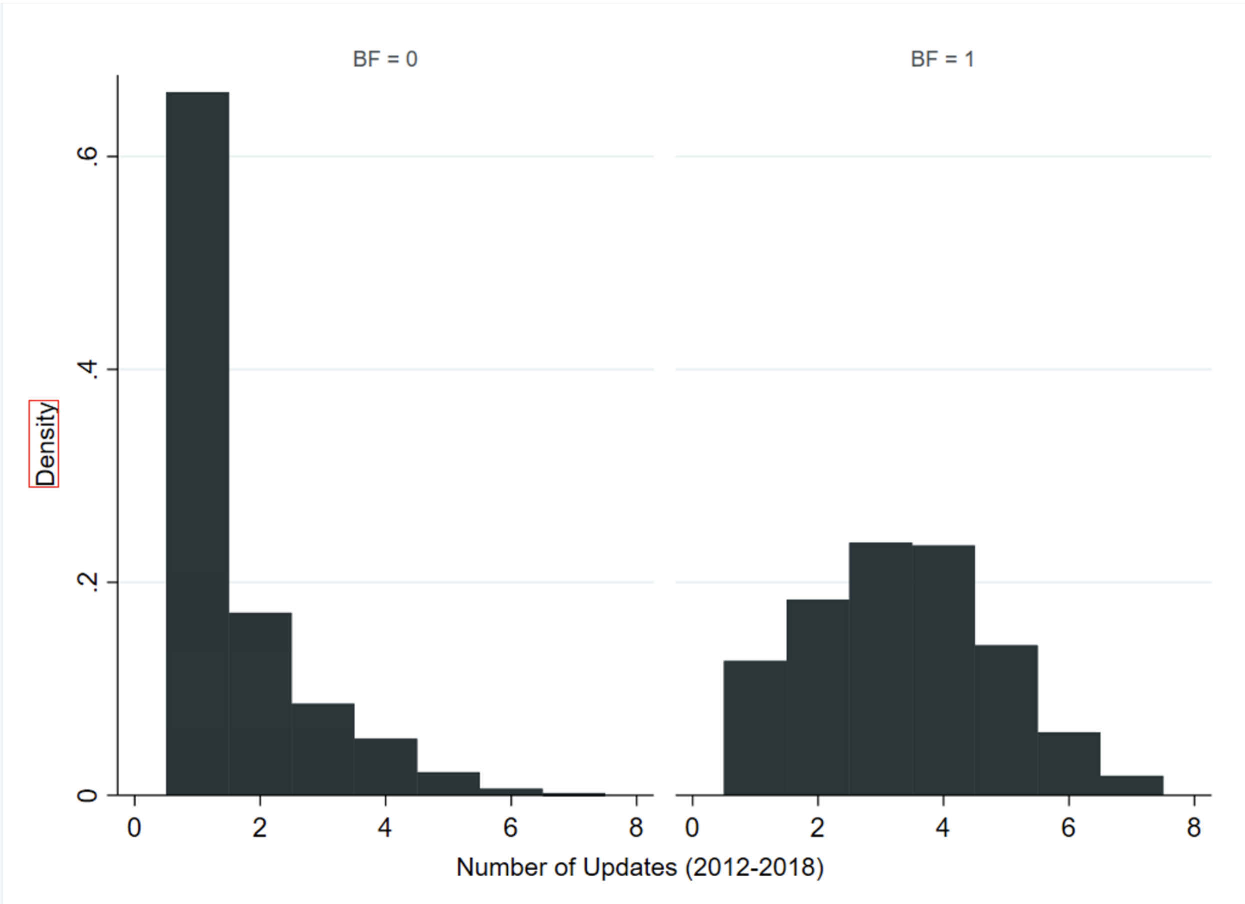


(b) **Sample lotteries**



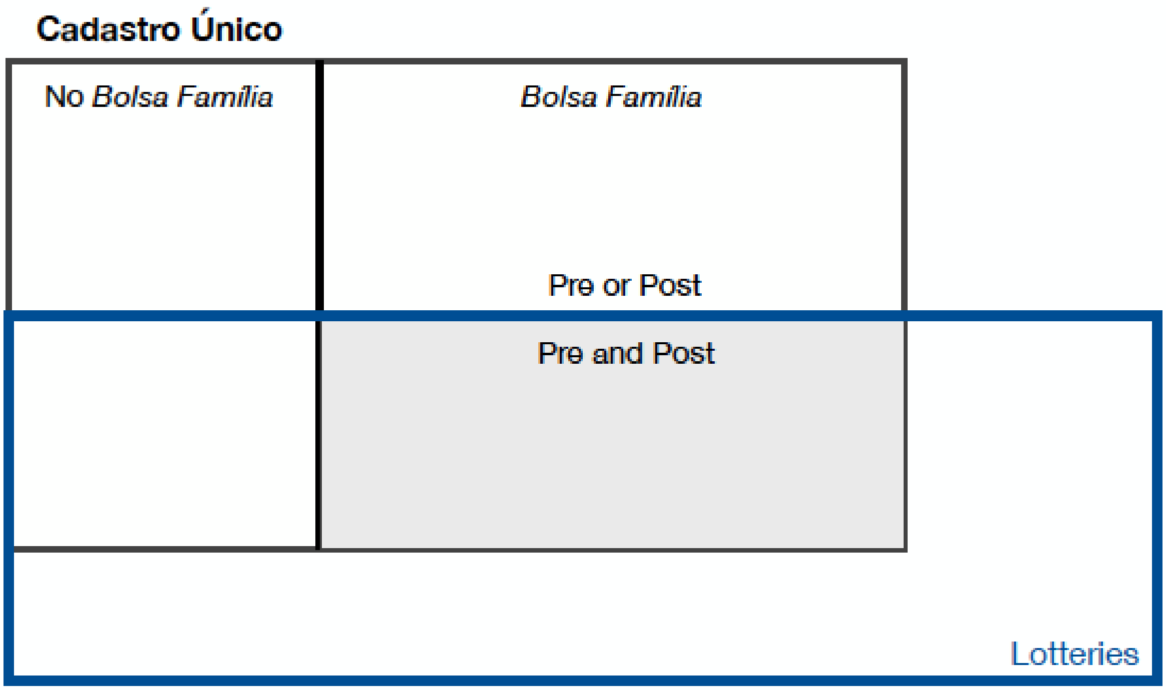
Note: Panel A plots the location of the MCMV units allocated through the regular lotteries that occurred during the period 2011-2015 and the neighborhood of residence of their subscribers. Panel B plots the location of the MCMV units allocated through the two lotteries used in our empirical investigation and the neighborhood of residence of their subscribers.

Figure 2: Updates



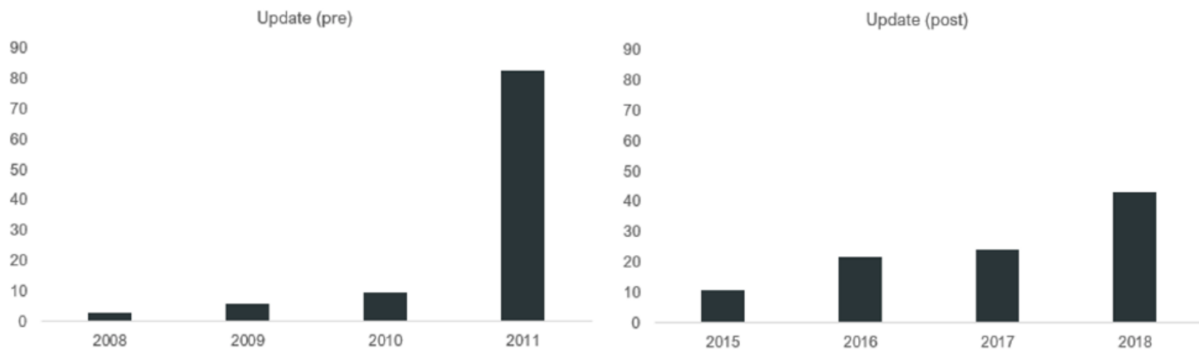
Notes: The figure reports the number of updates in the *Cadastro Único* by *Bolsa Família* status.

Figure 3: Sample Construction



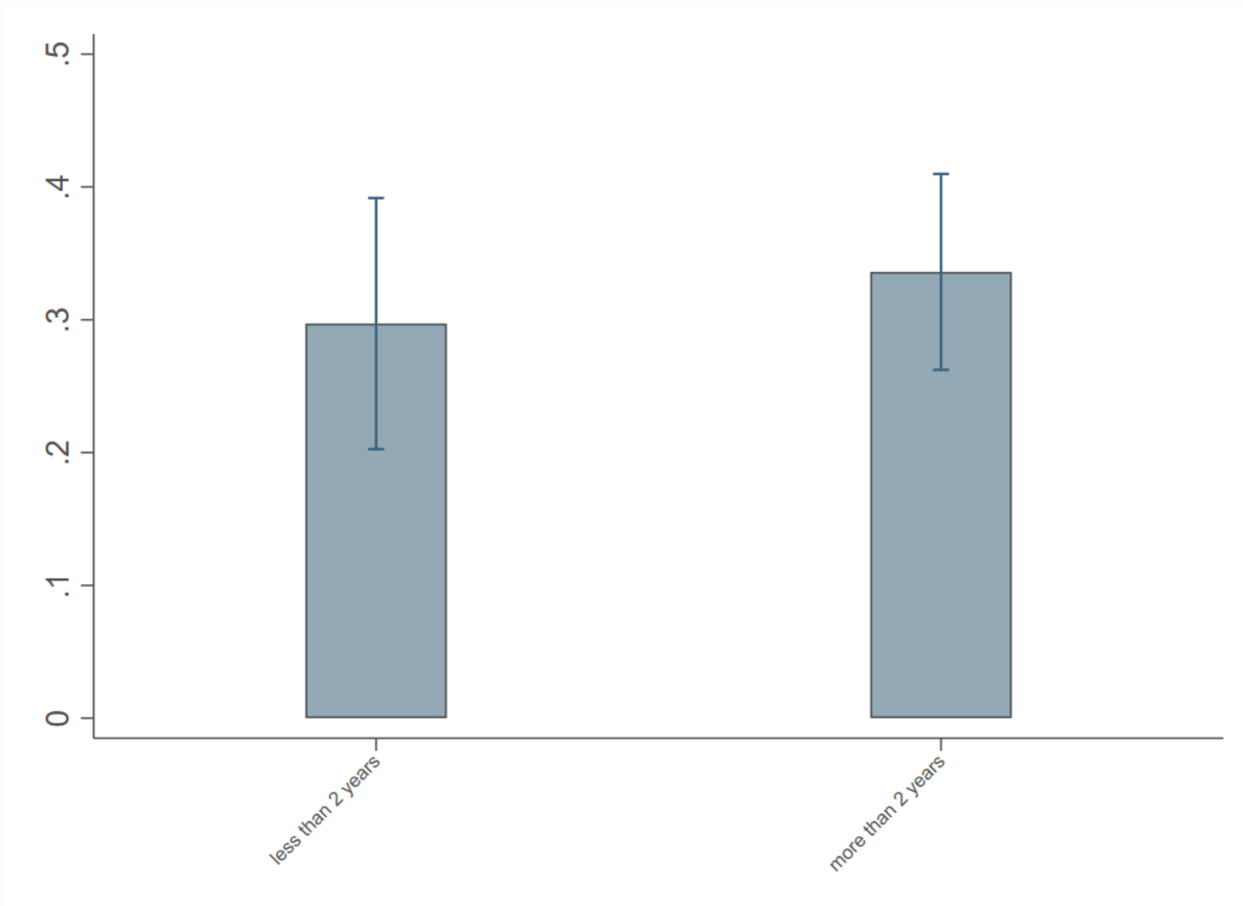
Notes: The depicts the sample construction. The black rectangle denotes the *Cadastro Único* and the blue rectangle the list of subscribers of the lotteries. The gray area represents the final sample.

Figure 4: Updates (pre and post treatment)



Notes: The figure reports the the year of the updates in the first pre-treatment period (used to test the balance of our sample) and last post-treatment period (used to test MCMV's effects). Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment.

Figure 5: The Evolution of Take-Up



Notes: The figure plots the coefficients obtained from estimating equation (2) using a dummy indicating of whether household moved to the neighborhoods in which the program's houses were built as dependent variable. The bars denote the coefficient and the capped lines their 95% confidence intervals.

Appendix to "Better Neighborhoods or Better Houses?"

A MCMV Description

In the main text, we present the aspects of the *Minha Casa Minha Vida* (MCMV) program more relevant to our empirical investigation. In this appendix, we provide a more complete description of the program. We begin by describing its rules, financing, and scale. We then describe the program in Rio de Janeiro (Brazil) and the lotteries used to select its participants.

A.1 The *Minha Casa Minha Vida Program*

Creation. The *Minha Casa Minha Vida* (MCMV) program was created by the Federal Law n. 11,977 in 2009. Its aim is to provide housing for low and middle-income households in Brazil. In the period 2009-2018, the program financed the construction of about 5.5 million houses at a total cost of R$464 billion (US$ 92.8 billion at the current exchange rates).

Segment 1. As mentioned in section 2, the MCMV offered different types of subsidies for households depending on the family's income level. These different brackets are shown in Table A2. In the program's segments 2 and 3, private developers sell units directly for households with income below R$ 5,000 (about US$ 1,250) with *Caixa* offering mortgages with subsidized rates. In segment 1, municipal governments allocate units to households with income below R$ 1,600 (about US$ 400) with *Caixa* financing the construction. Subsidies for segment 1 could go up to 90% of the construction cost. There were no down payment requirements, and monthly installments were capped at 5% of the household income (or R$ 25). Our investigation focuses on segment 1 of the MCMV program due to its focus on the poor population and participants' randomized assignment. There were three different initiatives focused on building houses for households poor households eligible for subsidies of the program's segment 1: MCMV-FAR, MCMV-

Sub-50, and MCMV-Entities. This paper focuses on MCMV-FAR which targeted poor households living in municipalities with more than 50,000 inhabitants. This modality concentrated more than 85% of the units built for segment 1.[25] The program's investments were more intensive in the country's largest municipalities due to the concentration of households living in inappropriate houses. Figure A1 shows the geographic distribution of the program in Brazil. Table A1 shows the number of MCMV contracts signed by year from 2010 to 2017.[26]

Funding. The program's resources come from the federal government budget.[27] These resources are managed and channeled mainly by *Caixa*, a stated-owned bank specialized in mortgage financing.[28] *Caixa* is also the financial institution responsible for most social programs of the Federal Government like the *Bolsa Família* (conditional cash transfer).

Execution. Private developers present the project to *Caixa*, and the bank is responsible for certifying construction companies, contracting the development of housing projects, and providing funding subsidized for eligible households. Municipalities are responsible for selecting the beneficiaries, guaranteeing compliance with urban regulation, guaranteeing the provision of infrastructure and public goods nearby housing projects, and improving the feasibility of developments, for instance, by donating land or providing tax cuts.

Application. Households were required to register for the lotteries either online or at municipal offices. In theory, households in the *Cadastro Único* (an administrative registry for managing the payment of federal government programs) eligible for the program were

[25]The MCMV-Entities also targeted households living in municipalities with more than 50,000 inhabitants. However, unlike the MCMV-FAR, the projects' execution and the selection of beneficiaries occurs through social movements. The MCMV-*Sub-50* initiative targeted poor households living in municipalities with less than 50,000 inhabitants.

[26]According to data on signed contracts provided by *Caixa*, the 309 contracts signed in 2009 were from the MCMV-Entities modality.

[27]The federal government transferred resources to three funds (FAR, FDS and FGTS - the FFF funds) to subside housing mortgages, to fund *FGHab* to provide guarantees for those mortgages, and to BNDES to finances urban infrastructure.

[28]*Caixa* is Brazil's largest mortgage lender, responsible for about 70% of Brazil's home mortgages.

registered automatically. However, it is unclear whether local governments (responsible for selecting the beneficiaries) registered these households in practice. We observe in the data is the beneficiaries eventually register in the *Cadastro Único*. Local officers organized lotteries among registered households to select beneficiaries. Households removed to construct infrastructure projects or individuals with disabilities are prioritized in the lotteries.

Lotteries. Housing lotteries were implemented by Ordinance n. 140 from the Ministry of Cities (from Abril 2010). For each housing project, at least 6% of housing units should be allocated to people with special needs and older people. If there is excess demand for these two groups, lotteries must also be used. Only families affected by natural disasters and reallocated due to federal-level infrastructure projects do not need to apply for lotteries to receive housing units. Every lottery must indicate a waiting list corresponding to 30% of the number of winners. The Federal Law n. 11,977 created three nationally defined priority criteria: families living in risk-prone areas, female-headed families, and families with people with disabilities. The law also allowed local governments to stipulate (up to) three additional priority criteria. Some municipalities have chosen local priority criteria; others have not. Lotteries can use numbers drawn from the results of other lotteries, or municipalities can carry out lotteries by themselves (that usually takes place in sports arenas and is supervised by a *Caixa* worker). For instance, Rio de Janeiro uses results from a famous national lottery run by *Caixa*, and separate lists of the general lotteries (with no priority criteria) and the lists of the special lotteries (with priority criteria).

Winners. When the units' construction finishes, all households selected to receive a housing unit of a particular housing project are invited to sign their contracts and receive their units on the same date. Units are identical and, given the large pool of applicants, selected households are unlikely to know each other. Thus, households typically do not reallocate units among themselves. Lottery results must be published in official registers (*"Diários Oficiais"*). After enrolling winners and the waiting list in *Cadastro Único*, municipalities send the list to *Caixa*. The bank then verifies compliance with the income threshold and other data by using different registries. When compliance is verified, *Caixa* authorizes the credit.

Subsidy. The design of housing subsidies is such that monthly installments should be several times lower than rent values. The lower the household income, the greater the subsidy. Subsidies are considerable: if total household income is within the segment 1 range, up to 90% of the housing price is subsidized. Subsidies were also designed to reach a wide range of the population. For instance, the segment 1 income range corresponded to the percentile 63 of the income distribution according to 2010 Brazil's 2010 Population Census. Every beneficiary must pay monthly installments (reduced by subsidies), lasting up to 120 months. Mortgage installments are set to be 5% of households' gross income. They are adjusted annually by a below-market interest rate, usually below inflation, to provide negative real interest rates (Resolution 477, October 2013). If the borrower does not pay the installments or uses units for other purposes, *Caixa* forecloses the unit.[29]

Houses. The MCMV program's housing units must have at least two bedrooms, a living room, a kitchen, and a bathroom. Its minimum surface area is 37 m^2 (roughly 400 sq^2). These houses are built in housing projects from a few dozen to more than one thousand units. MCMV's law established minimum requirements for the project's location. Projects should be located either inside the current urban network or in expansion areas indicated in the municipality's current urban planning. Moreover, these projects must comply with minimum requirements regarding environmental planning, sewage treatment, the electricity grid, the water network, access roads, and public transportation.[30] There is a price cap for housing units set by the federal government, which differs by state, municipal size, and housing type (house or apartment). The price cap influences the feasibility

[29] The comprehensive legislation regarding the PMCMV can be found at https://www.caixa.gov.br/Downloads/habitacao-minha-casa-minha-vida/_Legislacao_FAR.pdf.

[30] These requirements were instituted by the Provisional Measure 459 enacted in March 2009. This provisional measure was later converted into the Law #12,424 enacted in June 2011.

of housing projects and thus interferes with their location and scale. In large-size municipalities, where available land is scarce and more expensive, most housing projects are higher-density developments in the suburbs or surrounding municipalities with inadequate provision of urban infrastructure and public services ((Habitat, 2013)). In this scenario, a housing development usually comprises hundreds of housing units, so hundreds of families move virtually simultaneously to suburb areas. Figure A2 shows a MCMV house and its surroundings in Rio de Janeiro.

A.2 The Lotteries in the Municipality of Rio de Janeiro

Our investigation focuses on the lotteries in Rio de Janeiro, Brazil's second-largest municipality, with 6.7 million inhabitants. This municipality received 2% of the units and 2.45% of the MCMV program's total funding in 2009-2018. In total, 27,843 low-income households received units from the program. This corresponds to more than 1% of the number of families of the municipality. The typical unit built in Rio de Janeiro was priced at about R$ 51,644.00 and had square footage of about 45 m^2. We focus on this municipality for two reasons. First, the municipality publicly released the winners' and losers' records of the lotteries. Second, the municipality is known for the long-lasting prevalence of slums in the surroundings of its most important neighborhoods (Perlman, 2010; Monteiro & Rocha, 2017).

We analyze the general lotteries' list (with no priority criteria) from 2011 to 2015, corresponding to contracts signed between 2012-2017. A total of 11 general lotteries occurred in the period 2011-2015.[31] Three lotteries occurred in 2011, one in 2012, two in 2013, none in 2014, and five in 2015 (three of them considered age as a priority criterion in the selection process). Demand exceeded the supply of units in all lotteries, with 0.1% to 4.2% of the subscribers being selected. These 12 chosen lotteries households to live in 32 different

[31] We focus on general lotteries. There were specific lotteries for the disabled and elderly and for people living in risk areas. Information on the subscribers is publicly available on the website http://www.rio.rj.gov.br/web/smhc/menu-minha-casa-minha-vida#.

projects delivered between the years of 2012-2017. Figure A3 shows the location of these housing projects are spatially concentrated in the municipality's western zone.

As mentioned before, to select the beneficiaries, Rio de Janeiro uses results from a nationwide famous lottery run by *Caixa*. For these general lotteries, the allocation mechanism is straightforward. If the last two digits of the participant's registration number matched the Federal Lottery draw's last two digits, the household is selected. Applications are free of charge, and participants must not be the owner of a housing unit. Non-winners automatically participate in future lotteries. Winners who chose to withdraw from the rest of the process also automatically participate in future lotteries. Figure A5 shows a notice of the December 21, 2013 lottery that occurred in Rio de Janeiro indicating the Federal Lottery result that would determine the winners among the participants.

Table A4 presents information on the 10 lotteries from 2011 to 2015 for the low-income population (general lotteries) that we analyze in this paper. The houses delivered throughout 2012-2015 were clustered in the western zone of Rio de Janeiro. Only the region of Santa Cruz received 18 out of 26 units of Segment 1 of MCMV considered in these notices.

As explained in the main text, we use only data on the 2013 lottery. The 2011 lotteries selected beneficiaries of houses which were delivered during the year of 2012, the first year for which we observe the outcomes of most of the treatment and control households in our data. This precludes us from testing whether the treatment and control units we find in our match procedure were comparable before the MCMV happened. The 2012 lottery selected beneficiaries for for 240 houses to be constructed in Guadalupe in north of the municipality of Rio de Janeiro. However, the project was invaded after completion, first by 200 low-income families that lived in a slum nearby and, later, by armed gangs. These invasions had widespread media coverage[32], delayed the delivery of the units June 2015[33], and led to the organization of another lottery to select beneficiaries. It is not clear

[32] https://cutt.ly/ahpATJs; https://cutt.ly/KhpAOlo

[33] We obtain this information from the contract-level data of the program

how this lottery the took into account the 2012's lottery. This precludes us from using it.

The 2015 lotteries also had a series of problems. The lottery from notice 004/2015, which occurred on January 21, 2015, has 1848 duplicates in CPF. Besides, we observe several winners duplicated in the list of subscribers, with different registration numbers (the registration number is used to define the winner according to the Federal lottery). There are also several people whose digits were drawn in the Federal lottery and, still, they are not in the list of winners.[34] Therefore, we are not confident we can use these registers given this evidence that the lists of winners are not compatible with the exogenous rule for the lotteries.

Despite these issues, in the appendix C, we provide evidence that adjusting the lists of subscribers from the 2011 and 2015 lotteries and including them in our analysis does not change much the results. For 2011, we obtain reasonably balanced lists of winners and losers by excluding the households which we only observe after the MCMV houses were delivered. For 2015, we obtain reasonably balanced lists of winners and losers by excluding the more problematic lotteries. We then show the impacts of the MCMV estimated using these lotteries are qualitatively and quantitatively similar to the impacts estimated using all lotteries.

Table A3 presents the number of subscribers and winners on each lottery. Subscribers that were not selected remain on the list in subsequent raffles. The number of subscribers is much higher than the number of winners, representing from 0.1% to 4.2% of the total number of subscribers. The share of winners falls over the years (except in 2015(5)), and the main reason for that is that the waiting lists are cumulative - people who do not win a lottery remain on the list to participate in the subsequent raffle.

[34]In the 2015 notices, there are 942 people enrolled whose registration number ended in numbers selected in the Federal lotteries but did not appear in the winner's list

Table A1: Number of Signed Contracts by year

Year	# Houses	% Signed Contracts
2010	9986	1.1%
2011	102956	11.3%
2012	145713	15.9%
2013	133955	14.6%
2014	167268	18.3%
2015	170535	18.6%
2016	152013	16.6%
2017	32314	3.5%

Notes: The table presents the number of MCMV signed contracts by year for the urban population in Brazil, 2009-2017.

Table A2: Income threshold for eligibility

Monthly Family Income	**Benefits**
Up to R$ 1,600.00 (Faixa 1)	Up to 90% of the value of the property. The portion paid by the beneficiary is 5% of the monthly income with a minimum benefit of R$ 25 (divided into 120 months), without interest.
Up to R$ 3,100.00 (Faixa 2)	Subsidy with 5% interest per year.
Up to R$ 5,000.00 (Faixa 3)	Subsidy with 6%-7% interest per year.

Notes: Information compiled using data from the Ministry of Cities (2017)

Table A3: Number of subscribers by lottery

Year	Winners	Losers	Total
2011 (1)	2,983	295,149	298,132
2011 (2)	6,505	318,788	325,293
2011 (3)	14,055	337,343	351,398
2012	414	413,774	414,188
2013 (1)	566	471,902	472,468
2013 (2)	2,457	489,173	491,630
2015 (1)	3,337	664,958	668,295
2015 (2)	275	552,588	552,863
2015 (3)	2,226	554,184	556,410
2015 (4)	1,111	554,034	555,145
2015 (5)	7,305	554,719	562,024
2015 (6)	569	568,599	569,168

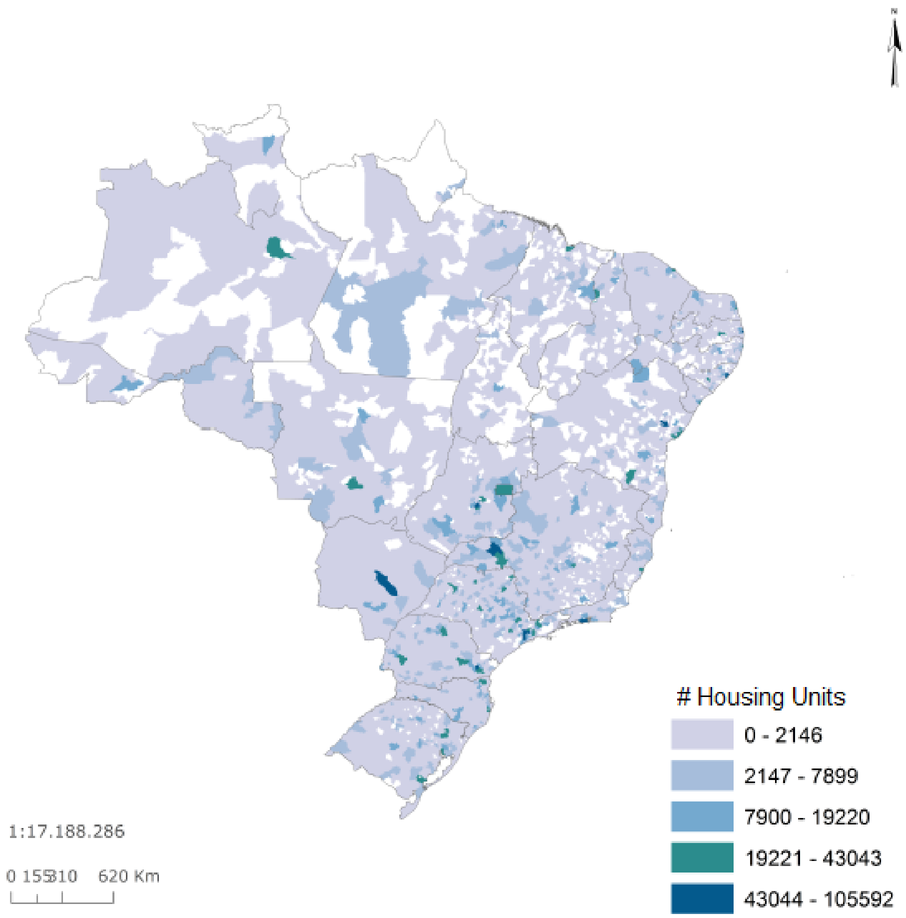
Notes: The table reports the number of winners (column 1), losers (column 2), and subscribers (column 3) of each of the general MCMV lotteries that occurred from 2011 to 2015 in the municipality of Rio de Janeiro (RJ).

Table A4: Lotteries Details

Notice	Project Name	Lottery date	Construction Year	Signature
2011/003	Park Imperial	11/06/2011	2009	23/10/2012
	Park Royal		2009	24/10/2012
	Destri		2009	09/01/2012
	Toledo		2009	28/06/2012
	Rio Bonito		2010	29/10/2012
	Estoril		2009	29/05/2012
2011/006	Sevilha	13/08/2011	2009	25/06/2012
	Taroni		2009	11/01/2012
	Cascais		2009	30/07/2012
	Toledo		2009	28/06/2012
2011/009	Vidal	02/11/2011	2009	11/04/2012
	Évora		2009	11/07/2012
	Zaragoza		2009	09/07/2012
	Park Imperial		2009	23/10/2012
	Park Royal		2009	24/10/2012
	Toledo		2009	28/06/2012
	Estoril		2009	29/05/2012
	Sevilha		2009	25/06/2012
	Cascais		2009	30/05/2012
2013/003	Vivendas das Garças	02/10/2013	2011	23/10/2014
2013/006	Recanto do Paçuaré I	21/12/2013	2011	27/04/2015
	Recanto do Paçuaré I I		2011	27/04/2015
	Vivenda dos Pintassilgos		2012	24/10/2014
	Vivenda das Gaivotas		2012	22/04/2015
2015/004	Mikonos	21/01/2015	2011	24/12/2015
	Dellos		2011	20/04/2016
	Santorine		2011	30/03/2016
	Vivenda das Cotovias		2012	04/05/2016
	Vivenda das Coleirinhas		2012	11/05/2016
	Vivenda dos Colibris		2013	13/05/2016
2015/007	Recanto do Paçuaré I	07/03/2015	2011	27/04/2015
	Recanto do Paçuaré II		2011	27/04/2015
2015/018	Mikonos	11/04/2015	2011	24/12/2015
	Dellos		2011	20/04/2016
	Santorini		2011	30/03/2016
2015/019	Vivendas das Cotovias	15/04/2015	2012	04/05/2016
	Vivendas das Coleirinhas		2012	11/05/2016
	Vivendas dos Colibris		2013	13/05/2016

Notes: The table reports the date of each lottery we use in our empirical investigation. It further reports the name of projects allocated in each of these lotteries, the year their construction started, and the dates in which contracts started to be signed with beneficiaries.

Figure A1: Geographic Distribution of Housing Units



Notes: The figure reports the regional distribution of the number of MCMV signed contracts using administrative data provided by Caixa (2017).

Figure A2: MCMV units in Rio de Janeiro

(a) Buildings

(b) Living Room

(c) Kitchen

(d) Buildings and Surroundings

Notes: House plan and housing units built through MCMV in Rio de Janeiro.

Figure A3: Projects’Location by Census Tract

Notes: Location of the units drafted through lotteries that occurred from 2011 to 2015 in the municipality of Rio de Janeiro.

Figure A4: 2013/006 Notice

PREFEITURA DA CIDADE DO RIO DE JANEIRO
SECRETARIA MUNICIPAL DE HABITAÇÃO

EDITAL DE DIVULGAÇÃO DE RESULTADO DE SORTEIO REFERENTE AO EDITAL 006/2013 PUBLICADO NO DIÁRIO OFICIAL DO MUNICIPIO EM 20/12/2013

A Prefeitura da Cidade do Rio de Janeiro, através da Secretaria Municipal de Habitação, torna público o resultado do sorteio da extração da Loteria Federal nº. 04.825 da Caixa Econômica Federal – CAIXA, realizado às 18:00 horas do dia 21/12/2013.

Resultado do 1º. Prêmio	10.352
Resultado do 2º. Prêmio	45.790
Resultado do 3º. Prêmio	87.091
Resultado do 4º. Prêmio	88.678
Resultado do 5º. Prêmio	88.701

Conforme prevê o Edital 006/2013 publicado no Diário Oficial do Município em 20/12/2013, foram sorteados os candidatos cujos três últimos algarismos do número que lhes foi atribuído corresponder às centenas 352, 790, 091, 678 e 701.
Posteriormente, os candidatos sorteados receberão no endereço cadastrado, carta-convite para comparecimento à reunião, onde serão informados acerca de todas as regras do PMCMV, em especial critérios de enquadramento, documentação necessária e prazo-limite para entrega da documentação, o qual, se descumprido, caracterizará desistência.

Notes: Image from notice of the December 21, 2013 general lottery from Rio de Janerio.

Figure A5: Eviction in MCMV housing project - Guadalupe

Notes: Image from the *Guadalupe* housing project during its eviction in November 2014. Image from TV Globo, obtained in https://noticias.uol.com.br/album/2014/11/12/predios-do-minha-casa-minha-vida-sao-invadidos-no-rio.htm?foto=1.

B MCMV Locations

In this section, we combine geo-coded information of the MCMV projects' with tract-level information of the Population Census 2010 to describe the neighborhoods where the MCMV units were built. Figure B1 panel a shows the distribution of income across the city, panel b the share of whites, panel c shows in yellow the census tracts that do not have full access to sewage, and panel d shows in yellow streets that has paved sidewalk. The location of these residential complexes is concentrated in areas with low-income families and with little infrastructure.

Table B1 presents several features for locations with no MCMV projects (column (1)), for areas with MCMV projects (column(2)), the difference between census tracts with and without projects, and the difference of theses variables within neighborhoods. Panel A of Table B1 has information on the average income and the share of female and literates. The average income of census tracts with no projects is R$ 813,94 higher than the average income of census tracts with projects. The percentage of whites is ten percentage points larger in locations with no projects. Females represent 47% of residents in census tracts without projects, and 44% of residents in places with MCMV projects and literates represent 97% and 96%, respectively. The share of females and literates are not significantly different. As column (3) shows, while the locations with no projects have higher income, the average characteristics within an area are similar between residents.

Panel B of Table B1 has information on street identification, the share of streets in the census tract with a paved sidewalk, maintenance hole, and share of tree-lined streets. The Table shows that locations with no projects have more roads with name identification than places with MCMV projects (difference of 13 percentage points). The percentage of streets with maintenance holes in locations with no projects is 15 percentage points higher than in areas with no projects. Finally, sites with no projects have more tree-lined streets than locations with MCMV projects (13 percentage points of difference). In all variables tested, areas with no projects have fewer public services than places with projects. Again, we see in column (3) that while the sites with no projects have less public services offer, the average characteristics within a location is similar between residents.

Panel C of Table B1 has information on the share of renters, the share of households with water sanitation, and the share of households with sewage sanitation. Locations with no projects have more renters (8 percentage points) and a higher share of households with sewage sanitation (14 percentage points). The difference within a location is substantially different in terms of sewage, as shown in column (4).

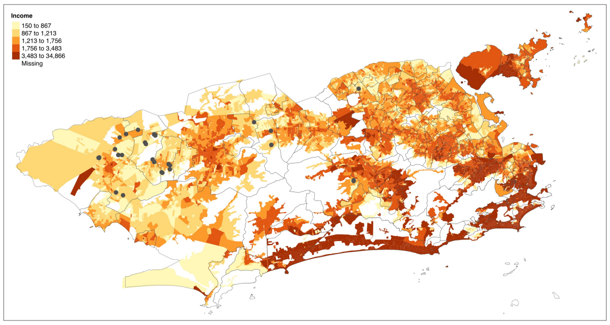
Table B1: Census Tracts with and without MCMV projects

	No Projects	Projects	Diff.	Diff. (within)
	(1)	(2)	(3)	(4)
Panel A: Neighbors				
Income	2374.32	1100.01	-1274.31***	25.26
	[2375.57]	[285.48]	(55.67)	(63.45)
White	0.54	0.38	-0.16***	0.02
	[0.21]	[0.09]	(0.02)	(0.02)
Female	0.47	0.44	-0.03	-0.01
	[0.11]	[0.12]	(0.02)	(0.02)
Literate	0.97	0.96	-0.01	0.01
	[0.05]	[0.04]	(0.01)	(0.01)
Panel B: Neighborhood				
Street Id.	0.14	0.43	0.30***	0.10*
	[0.27]	[0.34]	(0.06)	(0.06)
Sidewalk	0.11	0.41	0.30***	0.16***
	[0.27]	[0.37]	(0.07)	(0.06)
Manhole	0.14	0.45	0.32***	0.16**
	[0.28]	[0.38]	(0.07)	(0.07)
Trees	0.24	0.55	0.31***	-0.04
	[0.35]	[0.37]	(0.06)	(0.07)
Panel C: Houses				
Rental	0.22	0.11	-0.11***	-0.02
	[0.13]	[0.07]	(0.01)	(0.01)
Water	0.98	0.98	0.00	0.00
	[0.10]	[0.03]	(0.01)	(0.01)
Sewage	0.90	0.63	-0.27***	-0.17***
	[0.22]	[0.34]	(0.06)	(0.06)
Observations	10,247	34	10,247	10,247

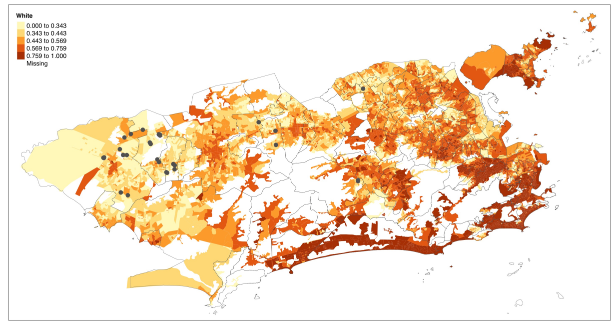
Notes: Column 1 reports the reports the mean of the outcome for census tracts without projects, column 2 for census tracts with projects, column 3 reports the differences in means between these census tracts, column 4 reports the within-neighborhood differences in means between these census tracts. Standard deviations are reported in brackets and standard errors in parenthesis. ***p<0.01; **p<0.05; *p<0.10

Figure B1: MCMV housing projects and Neighborhood characteristics

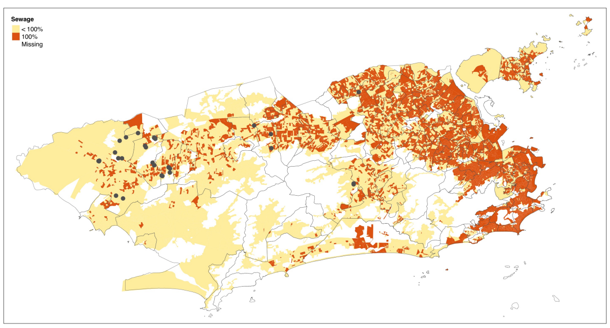
(a) Income



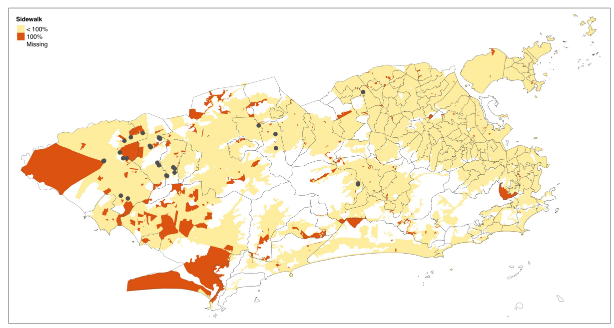
(b) Share of whites



(c) Sewage



(d) Paved Sidewalk

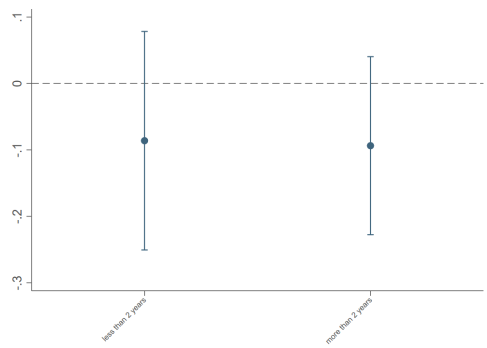


Note: Each panel reports the spatial distribution of a different socioeconomic at the census-tract level in the municipality of Rio de Janeiro. The black dots represent the location of the MCMV projects.

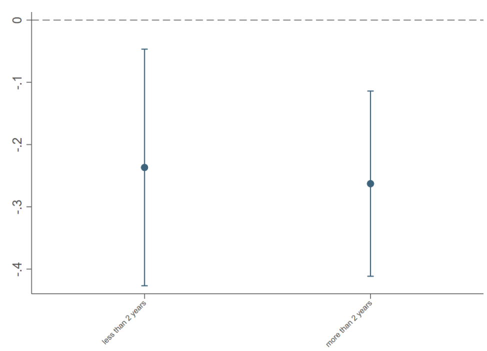
C Exposure Effects

Figure C1: Neighborhoods

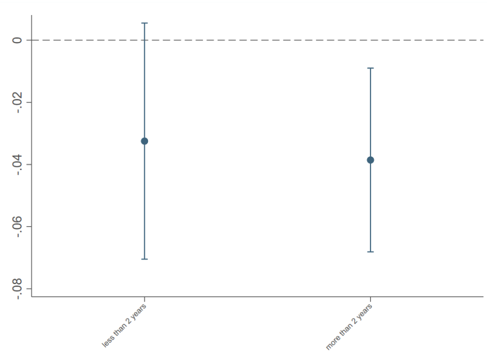
(a) Neighborhood index



(b) Neighbors index



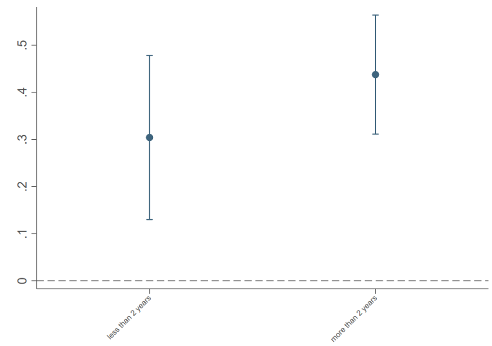
(c) Job accessibility



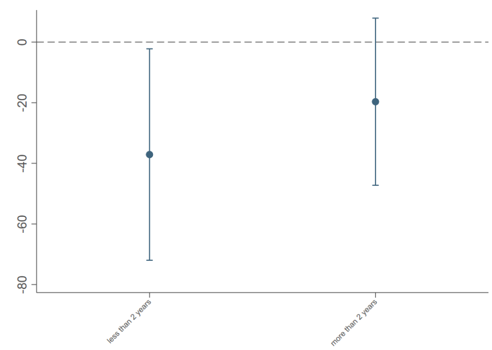
Notes: The figure plots the coefficients obtained from estimating equation (2) using a different indicators of neighborhood quality as dependent variables. The point denote the coefficient and the capped lines their 95% confidence intervals.

Figure C2: Housing Quality and Costs

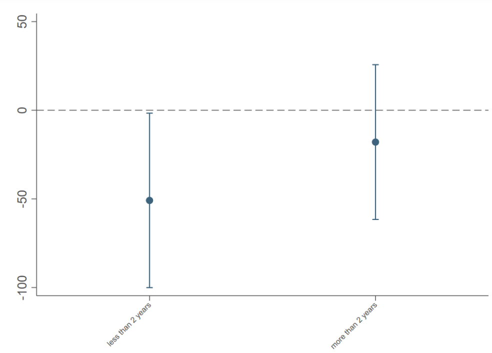
(a) House index



(b) Housing Costs



(c) Total Costs



Notes: The figure plots the coefficients obtained from estimating equation (2) using a different indicators of housing quality and costs as dependent variables. The point denote the coefficient and the capped lines their 95% confidence intervals.

D Results using Other Lotteries

Tables D1-D7 replicate the effects of the MCMV for the 2011 lotteries. Table D1 restricts the sample to households which updates their records before the first date of delivery of the units allocated using these lotteries. The other tables use the full sample.

Table D1: Descriptive Statistics and Randomization Check

	(1)	(2)	(3)	(4)
	Control	Treatment	T-C	N
Panel A: Demographics				
Female head	0.96	0.95	-0.00	15069
	[0.20]	[0.21]	(0.01)	
Age	38.80	39.02	0.22	15069
	[9.56]	[9.38]	(0.33)	
Spouse (0/1)	0.23	0.22	-0.01	15069
	[0.42]	[0.42]	(0.01)	
Children 0-6 (0/1)	0.42	0.40	-0.02	15070
	[0.49]	[0.49]	(0.02)	
Dwellers	3.69	3.75	0.06	15070
	[1.52]	[1.45]	(0.05)	
Joint significance test (p-value)			0.457	
Panel B: Neighborhoods				
Population (in 1000s)	108.45	99.57	-8.89**	10402
	[94.79]	[90.67]	(3.79)	
Sewage	0.78	0.78	0.00	10402
	[0.20]	[0.19]	(0.01)	
Water	0.99	0.98	-0.00*	10402
	[0.03]	[0.04]	(0.00)	
Sh. Work (head)	0.86	0.86	0.00	10402
	[0.03]	[0.03]	(0.00)	
Avg. Income (head)	1399.76	1451.65	51.89*	10402
	[678.93]	[649.53]	(27.14)	
Sh. white	0.44	0.45	0.01**	10402
	[0.09]	[0.10]	(0.00)	
Joint significance test (p-value)			0.131	

Descriptive Statistics and Randomization Check (continuation)

	(1)	(2)	(3)	(4)
	Control	Treatment	T-C	N
Panel C: Housing Quality				
Wood/Tile (0/1)	0.43	0.43	-0.00	15070
	[0.50]	[0.50]	(0.02)	
Sewage (0/1)	0.89	0.90	0.01	15070
	[0.32]	[0.30]	(0.01)	
Paving (0/1)	0.54	0.55	0.01	15070
	[0.50]	[0.50]	(0.02)	
Electricity (meter 0/1)	0.58	0.57	-0.01	14860
	[0.49]	[0.49]	(0.02)	
Dorms	1.34	1.32	-0.03	10280
	[0.70]	[0.53]	(0.02)	
Rooms	3.78	3.82	0.04	14858
	[1.49]	[1.68]	(0.06)	
Dwellers per room	1.09	1.10	0.01	14858
	[0.66]	[0.65]	(0.02)	
Joint significance test (p-value)			0.520	
Panel D: Housing Costs				
Rent	55.97	60.39	4.42	11066
	[106.45]	[111.05]	(4.48)	
Electricity	21.34	22.40	1.07	11764
	[38.44]	[35.61]	(1.41)	
Gas	36.25	34.32	-1.93	13280
	[53.43]	[12.40]	(0.65)	
Water	6.05	6.52	0.47	10914
	[15.32]	[16.17]	(0.66)	
Joint significance test (p-value)			0.473	
Panel E: Enrollment and LFP				
School Enrollment (%)	0.86	0.85	-0.01	13206
	[0.29]	[0.30]	(0.01)	
LFP (Head, 25-64)	0.54	0.56	0.03	9252
	[0.50]	[0.50]	(0.02)	
Joint significance test (p-value)			0.479	

Notes: Column 1 reports the mean of each indicator in the control group. Column 2 reports the mean of each indicator in the treatment group. Column 3 reports mean differences between the treatment groups and the control group. Column 4 reports the number of observations of each indicator. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Standard deviations are reported in brackets and robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

Table D2: Neighborhood Quality

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Panel A: Neighborhood				
Population	104.548	-0.607	1.591	28542
	(93.534)	(2.062)	(1.454)	
		[0.756]	[0.501]	
Sewage	0.776	0.008	0.004	28542
	(0.203)	(0.004)	(0.003)	
		[0.224]	[0.501]	
Water	0.987	-0.001	0.001	28542
	(0.026)	(0.001)	(0.001)	
		[0.494]	[0.501]	
Neighborhood Index	-0.008	0.020	0.022	28542
	[1.008]	[0.022]	[0.022]	
Panel B: Neighbors				
LFP (head)	0.860	-0.005***	-0.005***	28542
	(0.033)	(0.001)	(0.001)	
		[0.002]	[0.002]	
Income (head)	1421.942	-38.834**	-44.917***	28542
	(720.219)	(15.989)	(10.390)	
		[0.032]	[0.002]	
White (%)	0.444	-0.010 ***	-0.010***	28542
	(0.100)	(0.002)	(0.002)	
		[0.002]	[0.002]	
Neighbors Index	0.026	-0.114 **	-0.105**	28542
	[1.012]	[0.024]	[0.024]	
Panel C: Access to Opportunities				
Jobs: 90 minutes	0.283	-0.020 ***	-0.019 ***	28542
	(0.197)	(0.005)	(0.003)	
		[0.002]	[0.002]	
Schools: 90 minutes	0.277	-0.008 ***	-0.011 ***	28542
	(0.131)	(0.003)	(0.002)	
		[0.007]	[0.002]	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses and Romano-Wolf *p*-values correcting for multiple testing in each group of outcomes are reported in brackets. ***p<0.01; **p<0.05; *p<0.10

Table D3: Housing Quality

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Wood/Tile Floor (0/1)	0.53	0.04***	0.04***	28647
	(0.50)	(0.01)	(0.01)	
		[0.00]	[0.00]	
Paving (0/1)	0.66	0.02**	0.01	28647
	(0.47)	(0.01)	(0.01)	
		[0.02]	[0.25]	
Sewage (0/1)	0.90	0.01	0.01	28647
	(0.30)	(0.30)	(0.30)	
		[0.11]	[0.25]	
Dorms	1.33	0.11***	0.11***	28180
	(0.75)	(0.01)	(0.01)	
		[0.00]	[0.00]	
Rooms	3.83	0.18***	0.15***	28180
	(1.50)	(0.02)	(0.02)	
		[0.00]	[0.00]	
Electricity - meter (0/1)	0.58	0.07***	0.06***	28180
	(0.49)	(0.01)	(0.01)	
		[0.00]	[0.00]	
House Index	-0.13	0.20***	0.20***	28180
	(1.06)	(0.02)	(0.02)	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses and Romano-Wolf p-values correcting for multiple testing are reported in brackets. ***p<0.01; **p<0.05; *p<0.10

Table D4: Housing Costs

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Rent	59.27	-14.52*	-17.70***	28657
	(114.80)	(3.94)	(3.94)	
		[0.01]	[0.00]	
Water	5.90	3.56***	3.72***	28647
	(16.19)	(0.69)	[0.72)	
		[0.00]	[0.00]	
Gas	37.61	-1.89*	-1.89*	28647
	(53.70)	(0.69)	(0.72)	
		[0.08]	[0.10]	
Electricity	21.39	6.31	6.34	28647
	(64.15)	(2.59)	(2.86)	
		[0.12]	[0.13]	
Total	298.99	-0.44	-3.54	28647
	(296.78)	(7.23)	(7.19)	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses and Romano-Wolf *p*-values correcting for multiple testing are reported in brackets. ***p<0.01; **p<0.05; *p<0.10

Table D5: Female LFP

	(1) Control Mean	(2) Bivariate	(3) Baseline	(4) N
Panel A: Head (0/1)				
25-64	0.43	0.01	0.01	26305
	[0.49]	(0.01)	(0.01)	
25-44	0.44	0.01	0.00	15356
	[0.50]	(0.02)	(0.02)	
45-64	0.39	0.02	0.01	10949
	[0.49]	(0.02)	(0.02)	
Panel B: All (%)				
25-64	0.415	0.01	0.00	26869
	[0.485]	(0.01)	(0.01)	
25-44	0.424	0.01	0.01	16857
	[0.492]	(0.02)	(0.02)	
45-64	0.366	0.02	0.00	11385
	[0.481]	(0.02)	(0.02)	

Notes: Panel A reports the effects of the MCMV on the labor force participation of the heads of household. Panel A reports the effects of the MCMV on the labor force participation of adults in general. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

Table D6: Income

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Wage (head)	283.52	10.10	-0.84	28368
	[396.74]	(11.83)	(12.99)	
Wage (Household)	340.89	8.74	-7.11	28368
	[451.32]	(13.23)	(13.01)	
Income per capita	168.36	1.05	-3.44	28647
	[200.97]	(4.65)	(4.64)	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

Table D7: Education

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Enrollment, 4-18	0.89	-0.01	-0.01 *	22610
	[0.25]	(0.00)	(0.00)	
Boys, 4-18	0.89	-0.01	-0.01	15675
	[0.28]	(0.01)	(0.01)	
Girls, 4-18	0.90	-0.01	-0.01	15235
	[0.28]	(0.01)	(0.01)	
Pre-School, 4-6	0.61	-0.03	-0.06	4251
	[0.48]	(0.05)	(0.05)	
Enrollment, 7-15	0.96	-0.00	-0.00	17461
	[0.18]	(0.00)	(0.00)	
Boys, 7-15	0.95	-0.00	0.00	11003
	[0.20]	(0.00)	(0.00)	
Girls, 7-15	0.96	-0.01	-0.01	10563
	[0.19]	(0.01)	(0.01)	
Elementary, 7-15	0.85	-0.01	-0.01	17461
	[0.32]	(0.01)	(0.01)	
Boys, 7-15	0.85	0.00	0.01	11003
	[0.34]	(0.01)	(0.01)	
Girls, 7-15	0.85	-0.02	-0.02	10563
	[0.34]	(0.01)	(0.01)	
Enrollment, 16-18	0.94	-0.01	0.00	10770
	[0.23]	(0.02)	(0.02)	
Boys, 16-18	0.94	-0.02	0.01	5942
	[0.24]	(0.03)	(0.03)	
Girls, 16-18	0.94	-0.01	-0.03	5526
	[0.23]	(0.04)	(0.04)	
High School, 16-18	0.32	0.01	0.02	10770
	[0.45]	(0.02)	(0.03)	
Boys, 16-18	0.28	0.00	0.02	5942
	[0.44]	(0.03)	(0.06)	
Girls, 16-18	0.34	-0.02	0.04	5405
	[0.47]	(0.03)	(0.06)	
High School graduate, 19-24	0.31	0.00	0.05	9863
	[0.45]	(0.02)	(0.03)	
Boys, 19-24	0.26	0.02	0.03	5470
	[0.43]	(0.03)	(0.06)	
Girls, 29-24	0.34	-0.02	0.04	5405
	[0.47]	(0.03)	(0.06)	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

Tables D8 -D14 replicate the effects of the MCMV for the 2015 lotteries. The sample is restricted to the four lotteries without known implementation issues.

Table D8: Descriptive Statistics and Randomization Check

	(1)	(2)	(3)	(4)
	Control	Treatment	T-C	N
Panel A: Demographics				
Female head	0.96	0.97	0.00	42728
	[0.18]	[0.17]	(0.01)	
Age	38.57	39.14	0.57	42728
	[9.36]	[9.23]	(0.41)	
Spouse (0/1)	0.21	0.18	-0.02	42728
	[0.40]	[0.39]	(0.02)	
Children 0-6 (0/1)	0.46	0.45	-0.00	42729
	[0.50]	[0.50]	(0.02)	
Dwellers	3.80	3.82	0.02	42729
	[1.54]	[1.45]	(0.06)	
Joint Test (p-value)			0.177	
Panel B: Neighborhoods				
Population (in 1000s)	100.69	102.62	1.93	37060
	[91.53]	[95.29]	(4.55)	
Sewage	0.77	0.77	-0.00	37060
	[0.21]	[0.21]	(0.01)	
Water	0.99	0.98	-0.00	37060
	[0.03]	[0.03]	(0.00)	
Sh. Work (head)	0.86	0.86	0.00	37060
	[0.03]	[0.03]	(0.00)	
Avg. Income (head)	1415.31	1468.85	53.54	37060
	[732.79]	[750.71]	(35.83)	
Sh. white	0.44	0.45	0.01	37060
	[0.10]	[0.10]	(0.00)	
Joint Test (p-value)			0.116	

Descriptive Statistics and Randomization Check (continuation)

	(1)	(2)	(3)	(4)
	Control	Treatment	T-C	N
Panel C: Housing Quality				
Wood/Tile (0/1)	0.55	0.50	-0.04**	42728
	[0.50]	[0.50]	(0.02)	
Sewage (0/1)	0.90	0.90	-0.00	42728
	[0.30]	[0.31]	(0.01)	
Paving (0/1)	0.68	0.65	-0.03	42728
	[0.47]	[0.48]	(0.02)	
Electricity (meter 0/1)	0.56	0.58	0.02	42146
	[0.50]	[0.49]	(0.02)	
Dorms	1.32	1.30	-0.03	36706
	[0.84]	[0.50]	(0.02)	
Rooms	3.83	3.73	-0.10*	42147
	[1.51]	[1.09]	(0.05)	
Dwellers per room	1.09	1.12	0.03	42146
	[0.61]	[0.63]	(0.03)	
Joint Test (p-value)			0.101	
Panel C: Housing Costs				
Rent	53.96	56.23	2.27	37894
	[110.95]	[116.15]	(5.48)	
Electricity	20.02	21.11	1.09	38698
	[54.99]	[39.36]	(1.86)	
Gas	37.47	37.38	-0.10	40619
	[46.79]	[12.07]	(0.59)	
Water	5.36	5.11	-0.26	37727
	[15.88]	[14.31]	(0.68)	
Joint Test (p-value)			0.633	
Panel D: Enrollment and LFP				
Enrollment (sh)	0.89	0.88	-0.01	39034
	[0.25]	[0.27]	(0.01)	
LFP (female, 25a64)	0.51	0.52	0.01	33843
	[0.50]	[0.50]	(0.02)	
Joint Test (p-value)			0.867	

Notes: Column 1 reports the mean of each indicator in the control group. Column 2 reports the mean of each indicator in the treatment group. Column 3 reports mean differences between the treatment groups and the control group. Column 4 reports the number of observations of each indicator. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Standard deviations are reported in brackets and robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

Table D9: Neighborhood Quality

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Panel A: Neighborhood				
Population	100.692	13.642***	13.762***	42714
	(91.526)	(4.117)	(3.594)	
		[0.005]	[0.002]	
Sewage	0.771	0.001	0.002	42714
	(0.208)	(0.009)	(0.007)	
		[0.918]	[0.833]	
Water	0.986	-0.000	0.001	42714
	(0.028)	(0.001)	(0.001)	
		[0.918]	[0.716]	
Neighborhood Index	-0.005	0.093**	0.098**	42714
	(1.000)	(0.044)	(0.044)	
Panel B: Neighbors				
LFP (head)	0.858	-0.007***	-0.007***	42714
	(0.034)	(0.002)	(0.002)	
		[0.002]	[0.002]	
White (%)	0.442	-0.015***	-0.020***	42714
	(0.101)	(0.004)	(0.004)	
		[0.005]	[0.002]	
Income (head)	1415.310	-35.513	-68.993***	42714
	(732.787)	(32.517)	(18.791)	
		[0.262]	[0.002]	
Neighbors Index	0.020	-0.148**	-0.208**	42714
	(1.013)	(0.045)	(0.045)	
Panel C: Access to Opportunities				
Jobs: 90 minutes	0.274	-0.028***	-0.041***	42714
	(0.198)	(0.009)	(0.006)	
		[0.005]	[0.002]	
Schools: 90 minutes	0.271	-0.018***	-0.024***	42714
	(0.132)	(0.006)	(0.004)	
		[0.005]	[0.002]	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses and Romano-Wolf p-values correcting for multiple testing in each group of outcomes are reported in brackets. ***p<0.01; **p<0.05; *p<0.10

Table D10: Housing Quality

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Wood/Tile Floor (0/1)	0.55	0.04**	0.06***	42728
	(0.50)	(0.02)	(0.02)	
		[0.06]	[0.00]	
Paving(0/1)	0.68	0.04***	0.05***	42728
	(0.47)	(0.02)	(0.01)	
		[0.04]	[0.00]	
Sewage (0/1)	0.90	0.01	0.01	42728
	(0.30)	(0.30)	(0.30)	
		[0.43]	[0.30]	
Dorms	1.32	0.07***	0.09***	42728
	(0.84)	(0.02)	(0.02)	
		[0.02]	[0.00]	
Rooms	3.83	0.17***	0.20***	42728
	(1.51)	(0.04)	(0.04)	
		[0.00]	[0.00]	
Electricity - meter (0/1)	0.56	0.08***	0.07***	42728
	(0.50)	(0.02)	(0.02)	
		[0.00]	[0.00]	
House Index	-0.11	0.17***	0.23***	42728
	(1.09)	(0.04)	(0.04)	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses and Romano-Wolf *p*-values correcting for multiple testing are reported in brackets. ***p<0.01; **p<0.05; *p<0.10

Table D11: Housing Costs

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Rent	53.96	-27.33*	-31.97*	42728
	(110.95)	(6.69)	(6.69)	
		[0.09]	[0.08]	
Water	5.36	4.16*	4.66*	42728
	(15.88)	(1.24)	(1.31)	
		[0.08]	[0.08]	
Gas	37.47	-2.90	-2.80	42728
	(46.79)	(1.10)	(1.14)	
		[0.22]	[0.23]	
Electricity	20.02	22.65	25.06	42728
	(54.99)	(15.51)	(17.57)	
		[0.22]	[0.23]	
Total	296.01	2.50	2.98	42728
	(200.36)	(19.87)	(19.10)	
Cost Index	-0.22	0.04	0.06	42728
	(0.64)	(0.05)	(0.05)	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses and Romano-Wolf *p*-values correcting for multiple testing are reported in brackets. ***p<0.01; **p<0.05; *p<0.10

Table D12: Female LFP

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Panel A: Head (0/1)				
25-64	0.44	-0.02	-0.02	39692
	[0.50]	(0.02)	(0.02)	
25-44	0.45	-0.04	-0.04	23582
	[0.50]	(0.03)	(0.03)	
45-64	0.40	0.01	0.01	16110
	[0.49]	(0.03)	(0.04)	
Panel B: All (%)				
25-64	0.424	-0.01	-0.01	40359
	[0.486]	(0.02)	(0.02)	
25-44	0.435	-0.04	-0.04	25846
	[0.494]	(0.03)	(0.03)	
45-64	0.371	0.02	-0.05	16641
	[0.482]	(0.03)	(0.04)	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

Table D13: Income

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Wage (head)	273.44	-17.33	-18.49	42368
	[395.47]	(17.15)	(18.26)	
Wage (Household)	328.04	-20.74	-20.71	42368
	[449.86]	(19.95)	(19.45)	
Income per capita	158.36	11.48	10.94	42729
	[197.70]	(9.22)	(9.20)	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

Table D14: Education

	(1)	(2)	(3)	(4)
	Control Mean	Bivariate	Baseline	N
Enrollment, 4-18	0.89	-0.00	-0.00	34057
	[0.25]	(0.01)	(0.01)	
Boys, 4-18	0.89	-0.01	-0.00	23908
	[0.28]	(0.01)	(0.01)	
Girls, 4-18	0.89	0.00	-0.00	23171
	[0.28]	(0.01)	(0.01)	
Pre-School, 4-6	0.62	-0.00	-0.05	6694
	[0.48]	(0.08)	(0.08)	
Enrollment, 7-15	0.96	0.00	0.00	26843
	[0.17]	(0.00)	(0.00)	
Boys, 7-15	0.96	-0.00	0.01 ***	16969
	[0.19]	(0.00)	(0.00)	
Girls, 7-15	0.96	0.01 ***	0.01 ***	16398
	[0.19]	(0.00)	(0.00)	
Elementary, 7-15	0.86	-0.00	0.01	26843
	[0.31]	(0.01)	(0.01)	
Boys, 7-15	0.86	-0.00	0.01	16969
	[0.33]	(0.02)	(0.01)	
Girls, 7-15	0.86	0.01	-0.01	16398
	[0.33]	(0.02)	(0.02)	
Enrollment, 16-18	0.94	-0.02	-0.08 *	16208
	[0.23]	(0.04)	(0.04)	
Boys, 16-18	0.94	-0.00	0.07 ***	8951
	[0.23]	(0.01)	(0.01)	
Girls, 16-18	0.94	-0.05	-0.19 **	8284
	[0.23]	(0.09)	(0.09)	
High School, 16-18	0.34	-0.00	-0.02	16208
	[0.46]	(0.03)	(0.06)	
Boys, 16-18	0.29	0.02	0.17	8951
	[0.45]	(0.05)	(0.14)	
Girls, 16-18	0.35	0.04	0.04	8332
	[0.47]	(0.05)	(0.11)	
High School graduate, 19-24	0.32	0.01	0.07	14963
	[0.45]	(0.04)	(0.07)	
Boys, 19-24	0.26	0.00	0.16	8208
	[0.43]	(0.05)	(0.12)	
Girls, 29-24	0.35	0.04	0.04	8332
	[0.47]	(0.05)	(0.11)	

Notes: Each row reports the results for a different outcome. Column 1 reports the control mean. Column 2 reports the coefficient of a bivariate regression of each outcome on a treatment dummy. Column 3 add the outcome measure in the baseline as a control. Sample is restricted to households observed in the *Cadastro Único* pre and post-treatment. Robust standard errors are reported in parentheses. ***p<0.01; **p<0.05; *p<0.10

The Effects of Better Houses on Infant Health*

Abstract

This paper examines the effects of better houses on infant health in the context of Brazil's *Minha Casa Minha Vida* program, which built roughly 900.000 houses to poor households in Brazil during the period 2010-2017. We use a regression discontinuity design and administrative data to estimate the program's effects on health at birth and infant health. We find the program reduced the share of households living in inadequate houses by 18 percentage points. We find this improvement in housing conditions led to increases in birth weight and decreases in infant mortality (before 1 year) caused by conditions originating in children's perinatal period. We find no effect of the program in children with more than one year. Our results point out the importance of better houses in improving health at birth.

Keywords: *Housing Policies, Discontinuity, Health Outcomes*

*We gratefully acknowledge financial support from CAPES and FGV

1 Introduction

Health at birth is an important determinant of physical and mental health, human capital accumulation, and income (Gluckman et al., 2005; Cunha & Heckman, 2007; Currie, 2009). Nevertheless, while there is a growing body of empirical work documenting the role of shocks during fetal development on health at birth (e.g., (Almond & Currie, 2011; Almond et al., 2018)), there is much less evidence on the role of the environment in which the mothers live on fetal development and health at birth.

In this paper, we examine the effects of housing conditions on health at birth, exploring exogenous changes in housing conditions coming from investments of the *Minha Casa Minha Vida* Program (hereafter, MCMV). The MCMV is a series of initiatives introduced in the late 2000s focused on helping households become homeowners. It is divided into different segments according to the income of the beneficiaries. We focus on segment I of the program. In this segment, the federal government provides funds for the construction of heavily subsidized houses for poor households (monthly income below R$ 1,600 or US$ 320 at the current exchange rates). We obtain causal estimates of the construction of these houses exploring differences in the MCMV rules that facilitated municipalities with a population above 50,000 inhabitants to obtain funds from the program. This enables us to use a Regression Discontinuity (RD) design to estimate the effects of the MCMV program on health outcomes for municipalities close to the 50,000 inhabitants cutoff.

We begin by exploring the MCMV contracts' data to document the program's investments increase at the 50,000 population threshold. We find that the number of houses delivered by the program increases by 300-350 units during the period 2011-2017 at the 50,000 inhabitants threshold. This corresponds to 14-18% of the housing deficit of the typical municipality to the left of the discontinuity.

We then explore data on birth outcomes to document the program's effects on health at birth. We find that the birth weight increases by 12.9-15.6 grams at the 50,000 inhabitants threshold during 2011-2017. This effect is robust to different bandwidths and weighting procedures and statistically significant at the 5% levels regardless of the specification. Its magnitude corresponds to 0.4-0.5% of the mean birth weight in the sample. This is comparable to the effect of fasting during Ramadan on birth weight (see Almond & Mazumder (2011)) and the effect of job losses through announced notices during pregnancy (see Carlson (2015)). The increase in birth weight is driven by a combination of gestational length changes and birth weight conditional on gestational age (small for gestational age, thereafter SGA).[1] The share of pregnancies of less than 32 weeks decreases 0.2 p.p. (13% of the mean) while the share of SGA births decreases 0.8 p.p. (7.8% of the mean) at the discontinuity.

The effects of the MCMV on birth weight reflect a combination of the effects of houses and their construction. Delivering houses might improve infant health by enhancing access to sanitation, increasing in housing quality, reducing housing costs, and improving in property rights security. On its turn, constructing houses might influence infant health through changes in labor market conditions that affect employment and earnings. We use the timing of the effects of the MCMV to provide suggestive evidence of the importance of these mechanisms. We explore the fact that the program is expected to temporarily influence labor market conditions during the construction of the houses, but to influence living conditions permanently once the houses are delivered. Thus, we expect the effects of the construction to be stronger in the program's early periods (when the program's investments are at their peak, but the number of units delivered is modest) and the effects of houses to be stronger in the program's final periods (when the program's investments fall but the number of units delivered is considerable). We find the MCMV effects increase weakly through time, going from a statistically insignificant effect of fewer than 10 grams in 2011 to a statistically significant effect of more than 20 grams in 2017. This is suggestive evidence that the effects are driven primarily by houses.

[1] SGA newborns are those who are smaller in size than normal for the gestational age, commonly defined as a weight below the 10th percentile for the gestational age (Villar et al., 2014).

We further examine the effects of the MCMV on children's morbidity and mortality from 0-1 and 1-5 years. We test the program's effects on overall hospitalization and mortality rates and hospitalization and mortality rates by causes. We focus on causes most strongly connected to housing conditions and sanitation, and consider infectious, nutritional, and respiratory diseases, and perinatal origin's affections (also the main drivers of infant/child morbidity and deaths (Organization et al., 2019)). For children between 0-1 year, we find no effects of the MCMV on morbidity. However, there is a negative and statistically significant effect on mortality due to perinatal conditions of 1.1 deaths per 1,000 births.[2] This is consistent with the findings that the program improves health at birth. For children between 1-5 years, we find no effects either on morbidity or mortality.

There are several studies documenting positive effects of housing programs on different measures of adult health (e.g., Katz et al. (2001), Ludwig et al. (2013), Gale (2018) for the U.S., Barnhardt et al. (2017) for India, and Franklin (2019) for Ethiopia). There are also several studies documenting the positive effects of slum upgrading initiatives on the prevalence of diarrhea and respiratory problems on children (e.g., Cattaneo et al. (2009) and Galiani et al. (2017)). However, there is considerably less evidence linking housing policies and improvements in houses with health at birth (Vogl, 2007; Galiani & Schargrodsky, 2004). We contribute to this literature by documenting the meaningful effects of a housing program in Brazil on infant health.

The health externalities our work uncovers have important implications for the debate on the design of housing policies. The UN estimates that close to 900 million people live in these poor housing conditions in cities throughout the developing world. To deal with this issue, governments typically invest heavily in constructing houses for poor households in the cities' peripheries. However, there is concern these programs hurt households as moving to peripheries might increase the distance to job opportunities, thereby reducing

[2]Hospitalizations or deaths due to perinatal conditions are hospitalizations or deaths connected to the health at birth. While most of these events occur in the neonatal period (up to 28 days of life), they can happen at all ages.

employment and earnings and inducing households to return to their original neighborhoods (e.g., Barnhardt et al. (2017) and Picarelli (2019)). We contribute to this literature by documenting that, despite their negligible or negative effect on adults' economic outcomes, the construction of houses for poor households improves the health outcomes of children. Because improvements in infant health generate long-run benefits in terms of human capital and income (e.g., Gould et al. (2011) and Lavy et al. (2016)), this suggests that the long-run return of these programs might differ substantially from their short-run return. This distinction between effects on adults and children has proved important in other settings (e.g., Chetty et al. (2016) and Kumar (2019)), and is suggestive housing programs might have substantially different intra and inter-generational effects.

This discussion of the effectiveness of housing programs in general mirrors the effectiveness of the MCMV program in particular. The literature on the program finds no effects of the MCMV on employment and earnings (Pacheco, 2019; Squarize Chagas et al., 2019; Belchior, 2019). Our work shows that, despite its negligible effect on adults' economic outcomes, the MCMV improves the health outcomes of children.

The rest of the paper proceeds as follows. Section 2 provides a description of MCMV program. Section 3 provides a theoretical discussion of how house infrastructure may affect health outcomes. Sections 4 describes the data construction. Section 5 presents the empirical strategy. Section 6 present the results and discussion. Section 7 concludes.

2 Context and Background

This section describes the institutional background of the *Minha Casa Minha Vida* (MCMV) program focusing on the features relevant to our empirical investigation.

2.1 The *Minha Casa Minha Vida Program*

The *Minha Casa Minha Vida* (MCMV) program was created in the late 2000s to provide housing for low and middle-income households in Brazil. In the period 2010-2017, the program financed the construction of about 5.5 million houses at a total cost of R$464 billion (US$ 92.8 billion at the current exchange rates).

The MCMV is divided into four segments according to the income of the beneficiaries. Segment 1 covers households with income up to R$ 1,600 per month (US$ 320 at current exchange rates); Segment 1.5 covers households with income up to R$ 2,600 per month (US$ 520); Segment 2 covers households with income up to R$ 4,000 per month (US$ 800); and Segment 3 covers households with income up to R$ 9,000 per month (US$ 1,800). Each segment has access to different types of benefits. For segment 1, the government subsidizes 90% of the cost of the houses and provides financing to the other 10% at zero interest rates. For the other segments, the government provides financing to the households at subsidized rates starting at 5% per year.[3]

Our work focuses on households in segment 1. We observe roughly 900,000 units built with MCMV financing were destined for households in this segment. Different from the other segments, its houses are not sold in the market. Instead, they are allocated by local governments and, to a lesser extent, non-government organizations. Its resources come from the federal government budget and are managed and channeled by *Caixa*, a stated-owned bank specialized in mortgage financing. This bank is responsible for certifying

[3]The interest rate in Brazil on December 2010 was 10.66% per year (see https://www.bcb.gov.br/controleinflacao/historicotaxasjuros.

construction companies, contracting housing projects, and providing funding subsidized for an eligible household.[4]

Three different initiatives focused on building houses for households in segment 1: MCMV-FAR, MCMV-*Sub-50*, and MCMV-Entities. The MCMV-FAR targets poor households living in municipalities with more than 50,000 inhabitants. This is the main MCMV initiative focused on segment 1. Data on MCMV contracts from *Caixa* indicates that a total of 785,286 units were built under this initiative until 2017. This represents more than 86% of the units built for segment 1. Local governments run the MCMV-FAR. They are responsible for contracting construction companies to implement the projects and for selecting beneficiaries. Municipalities must follow guidelines issued by *Caixa*. However, there is no direct interference of neither the federal government nor *Caixa* in this process. Households were required to register for the program either online or at municipal offices. Local officers organized the selection of beneficiaries (ideally using the lotteries) among registered households. The project's construction typically ends one to two years after the selection of beneficiaries occurred. When the units' construction finishes, the beneficiaries are invited to sign their contracts with *Caixa*.

The MCMV-Entities also targeted households living in municipalities with more than 50,000 inhabitants. However, unlike the MCMV-FAR, the projects' execution and the selection of beneficiaries focuses on the active participation of homeless people's movements, such as Homeless Workers' Movement. These social movements engage in the development of the housing project, managed the project's execution and budget, and select beneficiaries. It is a small initiative with 22,035 units being built under it until 2017. This represents 2.4% of the segment's 1 contracts and less than 1% of the total resources invested in the program (Tatagiba & Teixeira, 2016).

The MCMV-*Sub-50* initiative targets poor households living in municipalities with

[4]Houses in the other three segments are built and sold by private construction companies with *Caixa* providing financing to the buyers.

less than 50,000 inhabitants. In this initiative, the federal government subsidizes housing units' construction in these municipalities through contracts with local governments, local companies, or self-construction efforts organized by the communities themselves. The selection of projects to be subsidized occurs through public notices issued by the federal government. It is important to note that the MCMV-*Sub-50* did not exist when the MCMV was created. It emerged later as the result of lobbying efforts of officials from municipalities of less than 50,000 inhabitants. However, from its beginning, this initiative encountered numerous problems for its implementation. Indeed, it missed the target of constructing 200,000 units until 2017, with 101,612 being delivered until this period, 11.5% of the total units built.

The contrast between the rules of the MCMV-FAR and the MCMV-*Sub-50* indicates municipalities with more than 50,000 inhabitants find it much easier to get a house from the MCMV than municipalities with less than 50,000 inhabitants. This reflects in the number of signed contracts under the different initiatives of the MCMV described above and in the amount of government funds on each modality. While the MCMV-*Sub-50* received R$ 1 billion in subsidies, the MCMV-FAR received R$ 16.5 billion in subsidies (Biderman et al., 2019). Our work explores the different intensities of the MCMV investments to obtain causal estimates of the construction of these houses. Specifically, we explore the differences in the MCMV rules that facilitated municipalities with a population above 50,000 inhabitants to obtain funds from this program. This enables us to use a Regression Discontinuity (RD) design to estimate the effects of the MCMV program on health outcomes for municipalities close to the 50,000 inhabitants cutoff.

2.2 The Roll-Out of the *Minha Casa Minha Vida*

To understand the program's roll-out, Figure 1 depicts the number of contracts of segment I of the MCMV by year. The first contracts of the program are signed in 2010. The number of contracts expands rapidly between 2010 and 2012, stabilizes between 2012-2016, and ends in 2017. A total of 886,898 contracts were signed between poor urban households and *Caixa*.[5]

As explained before, the different rule for obtaining investments from the MCMV for municipalities of different sizes implies that the program's roll-out might differ substantially between municipalities of different sizes. Figure 2 shows this is indeed the case. It plots the accumulated number of signed contracts signed in municipalities with 40,000-50,000 and municipalities with 50,000-60,000 inhabitants. Panel A reports the average number of contracts per municipality, while Panel B the number of contracts as a proportion of the housing deficit in 2010. Both panels report large differences in the program's investments in the two groups. About 170 houses were delivered for the typical municipality with a population between 40,000-50,000 between 2011-2017. This contrasts with about 520 houses delivered for the typical municipality with a population between 50,000-60,000 inhabitants. This represents about 20% of the housing deficit in these municipalities.

The differences in the MCMV investments by municipality population are driven mainly by abrupt changes in the program's investments at the 50,000 inhabitants threshold, as reported in Figure 3. This figure plots the mean number of contracts (Panel A) or the number of contracts divided by the housing deficit (Panel B) at fifteen population bins in the 20,000-80,000 inhabitants interval. There is hardly any relationship between the number of contracts or the number of contracts divided by the housing deficit below and above the 50,000 inhabitants threshold. This contrasts with the sharp change in the number of contracts at this threshold. This figure indicates the program's rules generate discontinuous changes in the MCMV investments, thereby implying it is possible to estimate the program's effects using a Regression Discontinuity (RD) design.

[5] This number considers the MCMV-FAR and MCMV-*sub50*.

2.3 House Characteristics of Low-Income Families

The poor conditions of the houses in which most of Brazil's poor population lived suggest the MCMV beneficiaries might experience significant improvements in housing conditions such as access to sanitation, presence of bathrooms with proper latrines, clean floors, and good ventilation. The program's units must have at least two bedrooms, a living room, a kitchen, and a bathroom. Its minimum surface area is 37 m^2 (roughly 400 sq^2). Besides, the project's location must follow some minimum requirements in terms of environmental planning, sewage treatment, connection to the water network, etc.[6] These characteristics contrast with the houses' characteristics in which the poor population lives in the country. According to the 2010 Population Census, 43.8% (14,588,592) of the households eligible for MCMV's segment I do not have access to proper sanitation.[7] and that 4.1% (1,374,160 households) live in houses poorly built.[8] This suggests that MCMV investments might have increased housing conditions markedly.

The MCMV beneficiaries might also have experienced significant decreases in housing costs. The government subsidizes 90% of the cost of the unit ($\approx$ R$ 50,000) and finances the rest in 120 months with no interest rates. This implies that the beneficiaries typically pay less than R$ 50 per month. According to the 2010 Population Census, the mean rent paid by households eligible for MCMV's segment 1 was R$ 252.45, a much larger number. Moreover, 3.8% (2,209,688) of the households in the segment I are considered in deficit due to excessive rent. This suggests that MCMV investments might have generated noticeable reductions in housing costs and, therefore, income increases.

Evidence for the municipality of Rio de Janeiro (RJ) presented in chapter 1 corroborates these hypotheses. Exploring the lotteries used to select the program's beneficiaries, we

[6]These requirements were instituted by the Provisional Measure 459 enacted in March 2009. This provisional measure was later converted into Law #12,424 enacted in June 2011.

[7]with no access to water and sewage network

[8]defined as improvised households or permanent households (houses or apartments) made of material other than masonry or paired wood.

document that the MCMV reduced housing costs and improved housing conditions. It is certainly not possible to extrapolate the evidence from Rio de Janeiro to our setting of mid-size municipalities. However, this evidence highlights the potential connection between MCMV investments and better housing conditions.

3 The Expected Effects of the MCMV on Infant Health

Before presenting the results, we briefly discuss the expected effects of the MCMV program on infant health. The MCMV might influence infant health through the houses themselves and their construction.

Delivering houses to poor households might influence infant health through numerous channels. First, the houses built by the MCMV program might improve access to sanitation. There is considerable evidence that access to clean water and appropriate sewage collection improve infant health substantially by reducing the incidence of communicable diseases due to oral contamination (Cutler & Miller, 2005; Hutton et al., 2004; Lilford et al., 2017). The evidence further indicates that there are complementarities between water and sewage services (e.g., Duflo et al. (2015) and Alsan & Goldin (2019)). Better sanitation might have long run consequences on health and human capital as suggested by the studies of Gould et al. (2011) and Lavy et al. (2016).

Second, the houses built by the MCMV might increase housing quality in general, thereby reducing the likelihood households live in houses without bathrooms with proper latrines, clean floors, and good ventilation. Bathrooms with proper latrines and clean floors improve child health by reducing fecal-oral transmission. For instance, Hammer & Spears (2016) finds evidence of substantial benefits in terms of infant mortality and height of a program to induce the use of latrines in India, while Cattaneo et al. (2009) finds evidence of significant decreases in the incidence of parasitic infections, diarrhea, and the prevalence of anemia of a program that installed cement floors in Mexico. Lack of ventilation might deplete health of its residents by increasing the incidence of respiratory diseases (Cappelletty, 1998). This effect might be strengthened by the prevalence of traditional cooking techniques which are a major source of indoor air pollution (Ezzati et al., 2004).

Third, the houses built by the MCMV might reduce housing costs. This is equivalent to an increase in the non-labor income. The existing evidence indicates that increases in non-labor income unambiguously influence infant health (Strully et al., 2010; Hoynes et al., 2015). This effect is typically tied to increases in maternal nutrition, reductions in maternal stress, and changes in time use of the mothers towards home production.

Fourth, the houses built by the MCMV might reduce property rights insecurity. Several studies indicate that more secure property rights increase welfare in general with positive effects on female labor force participation (e.g., Field (2007)), physical and human capital investments (e.g., Galiani & Schargrodsky (2010)), and child health (e.g., Galiani & Schargrodsky (2004); Vogl (2007)).

The construction of houses is the other mechanism through which the MCMV program might influence infant health. The construction activities promoted by the program might increase labor demand in construction, thereby increasing the employment and earnings of the households. Theoretically, the effects of improvements in labor market conditions on infant health are ambiguous as changes in time use of the mothers might offset the increases in earnings (Glick, 2002). However, most empirical studies indicate that increases in earnings improve health outcomes among poor households in developing countries (e.g., Baird et al. (2011), Rocha & Soares (2015), Adhvaryu et al. (2019)).

We expect the total effects of the MCMV program on infant health to reflect the combination of the effects of houses and their construction. These mechanisms are expected to improve infant health, thereby reducing the incidence of pre-term births, increasing birth weight, reducing child hospitalization and death rates. However, the timing of their expected impacts differs. Improvements in labor market conditions generated by the MCMV are temporary and concentrated in the period of construction of the houses. Therefore, they are expected to influence infant health mostly in the period 2011-2015 (which concentrates most of the program's investments) but not later. The changes in living conditions generated by the houses built by the MCMV are persistent and increase as more households move to these houses. Therefore, they are expected to influence infant health more as the number of houses delivered increases.[9]

[9]This is true if we consider that the house depreciation nor program exit are nor relevant to our context.

4 Data

4.1 Data Sources

We use data from multiple data sources. To obtain information on MCMV housing contracts, we use official data of the program's contracts obtained from *Caixa*.[10] To generate health outcomes on the municipality level, we use health data at birth, hospital admissions, and mortality from the Brazilian Ministry of Health (MS/DATASUS). Furthermore, to generate information on demographic and socio-economic characteristics, we use the 2010 Population Census. We describe each data source in detail below.

Contracts. *Caixa* provides information on the 886,898 mortgages signed by the beneficiaries of the MCMV program from 2010 to 2017. The data is at the individual level. We have information on the date the contract was signed and the municipality of each contract's housing project. Using this data, we construct a municipality-level panel of the number of signed contracts by year. During the period, individuals from 1,671 municipalities signed contracts to purchase subsidized houses from segment 1 of MCMV under MCMV-FAR and MCMV-*sub50*. From these municipalities, 1,174 had population below 50,000 inhabitants and 497 above this threshold.

Health Outcomes. We construct a dataset on health at birth, infant and child morbidity, and mortality outcomes, combining microdata from the Brazilian National System of Information on Birth Records (Datasus/SINASC), the Brazilian National System of Hospital Admissions (Datasus/SIH), and the Brazilian National System of Mortality Records (Datasus/SIM).

The birth records (Datasus/Sinasc) provides information on birth weight, length of gestation, and APGAR score. The database also provides the exact date of birth and the municipality of birth. This information allows us to construct a municipality-by-year of

[10]Information made available using the Access to Information Requirement number 99902.001060/2017-08.

the birth panel over the 2009-2017 period containing information on the number of births, average birth weight, and the average length of the gestational period.

Hospitalization microdata is obtained from the National System of Information on Hospitalizations (Datasus/SIH), which contains administrative information at the hospital admission level and is managed by the Health Care Agency (SAS/Ministry of Health). The system includes all hospital admissions covered by SUS, both in public facilities and private hospitals accredited by the government. It provides information on patients' age, gender, and cause of hospitalization (ICD-10).[11]

We obtain mortality microdata from the Brazilian National System of Mortality Records (Datasus/SIM), which collects every death officially registered in Brazil. It contains data on deaths by cause (also following ICD-10), birth date, municipality of birth, and residence. We select all deaths of individuals up to one year of age born between 2009 and 2017 and deaths of individuals from one to five years old in the same period. We then build a municipality-by-year death panel for the 2009-2017 period containing information on the number of infant and child deaths (total and by cause of death).

Both SIH and SIM microdata sets include patients' municipality of residence and the date of the hospital admission or death. The date of the event and the code of the municipality of residence are used to aggregate the microdata into a municipality-by-year data set and to match it with data from other sources. We follow the literature on the health impacts of houses and access to sanitation and focus specifically on infectious diseases, nutritional diseases, respiratory diseases, and diseases with perinatal origin's (these are also the main drivers of infant/child morbidity and deaths).[12]

To facilitate comparisons across municipalities and time, we compute health outcomes (such as hospitalizations and mortality) in rates per 1,000 municipality births in the last

[11] The diagnostic codes follow the International Classification of Diseases, 10th Revision (ICD-10).

[12] Infections diseases refer to events classified under ICD-10 A00-B99; digestive diseases refer to events organized under ICD-10 E00-E90); respiratory conditions refer to events classified under ICD-10 J00-J99; conditions originated in the perinatal period refers to events classified under ICD-10 P00-P96

year for infants. We approximate that accumulating the previous four years of births for children aged 1-5 years old.

Other data. We use other sources of data to conduct the analysis. We collect municipality level data on the population size in 2007 (before MCMV) from the population count conducted by the Brazilian Census Bureau (IBGE). We construct an indicator of baseline characteristics using the 2010 Brazilian Census. We construct information on socioeconomic indicators such as the shares of females, young (<18), adults (>28), old (>60), the share of households located in rural areas, the total number of households, the share of migrants, the share of workers, the average wage, the share of individuals with less than 9-11 years of education, the share with less than 12-15 years of education, the share of people with 16 or more years of schooling, the share of workers, and the average wage (in R$). We also construct information on infrastructure characteristics such as the percentage of households with access to piped water and households with access to sanitation. We get information on housing deficit at the municipality level using data from the 2010 Population Census using the methodology proposed by Furtado et al. (2013). Finally, we use information on health inputs from CNES/Datasus. We calculate the number hospital beds, the number of hospitals, and the coverage of the Family Health Program in the municipality.

Merge. The information on births and infant/child mortality and hospital admission is merged by municipality and year with the MCMV contract data and the other datasets described above. The average number of contracts per year is 14. Table B1 presents descriptive statistics of our main variables for all municipalities and for the 235 municipalities between the 40,000 and 60,000 inhabitants that are the focus of our empirical analysis. These municipalities have a total of 11,380,994 inhabitants, accounting for 6.2% of the country's population. Their average population is 48,429, higher than the 33,063 average population observed in the country as a whole. The municipalities around the threshold are similar to the country in terms of age structure, labor and schooling characteristics. They are less rural and have worse sanitation indicators. Birth characteristics and morbidity and mortality characteristics are similar to the country's statistics. The average birth weight is 3.2 kg, and 9% of pregnancies last less than 37 weeks. The infant hospitalization rate is 184 per 1000 births per year. For children, the hospitalization rate is 68 per 1000 births. Mortality under 1 year old is 14 per 1000 births, while mortality from 1 to 5 years old is 0.71 per 1000 births.

5 Empirical Strategy

Estimating the effects of large-scale government investments such as the *Minha Casa Minha Vida* is challenging because the allocation of these investments is typically correlated with factors like political favoritism or economic potential. This implies that comparisons of regions more affected by these programs ("treatment") with regions less affected by them ("control") will be biased. Moreover, because the direction of the correlation between the factors governing the investments and the outcomes of interests might be positive or negative, the direction of bias is unknown. For instance, it is unclear whether the unobserved factors which influence the investments of the MCMV are positively or negatively related to infant health.

To overcome these issues, we use a Regression Discontinuity (RD) design to obtain causal estimates of the effects of the MCMV program on health outcomes. Our empirical framework explores a program's rule that facilitated access of municipalities with more than 50,000 inhabitants to this program's funds. As detailed in Section 2, municipalities with less than 50,000 inhabitants had to submit proposals to be evaluated by the federal government before obtaining financing to built houses with funds of the MCMV program, while municipalities with more than 50,000 inhabitants could obtain these funds directly with *Caixa*. This enables us to use RD to estimate the effect of being able to get MCMV funds directly from Caixa (hereafter MCMV investments) on municipalities' health outcomes close to the 50,000 inhabitants cutoff.

We implement our RD design using a local linear regression approach following the guidelines from Imbens & Lemieux (2008), Lee & Lemieux (2010), and Gelman & Imbens (2019). Formally, we use the following model to obtain RD estimates of the effects of the MCMV on infant health:

$$Y_{its} = \beta_0 + \beta_1 T_{is} + f(P) + \beta_2 Y^I_{is} + \gamma X_{is} + \eta_s + \eta_t + \nu_{its}, t \in [2011, 2017] \tag{1}$$

in which $T_{is} = 1[P_{is} > 50,000]$, $f(P)$ is a linear spline ($f(P) = \mu P_{is} + \varphi P_{is} \times T_{is}$), and $P_{is} \in [50,000 - h, 50,000 - h]$. Y_{its} denotes an outcome of interest of municipality i, year t and state s, P_{is} is the population of the municipality measured, T_{is} is a treatment indicator which is one if the municipality population is above 50,000, Y^I_{is} is the baseline value of the outcome of interest, X_{is} is a vector of municipality controls (measured in 2010), η_s is a state-fixed effect, and h is the bandwidth chosen to select the municipalities used in the estimation. The coefficient of interest is β_1 which measures the difference in outcomes of municipalities just below and above the 50,000 inhabitants threshold.

The outcomes of interest Y_{its} are measured in 2011-2017. This is the period for which we observe most of the MCMV contracts being signed.[13] The initial conditions Y^I_{is} and the controls X_{is} are measured in 2009 and 2010, i.e., immediately before MCMV investments began. The controls included are the share of rural households, the baseline access to the water and sewage networks, and health infrastructure indicators. We include initial conditions, controls, and state fixed effects to improve our estimates' precision as it is common in the literature.[14]

We estimate equation (1) using a preferred bandwidth of 10,000 inhabitants. This results in a sample of 1,645 observations (235 municipalities per year). We fix the bandwidth (instead of choosing the bandwidth optimally as proposed by Calonico et al. (2014)) to ensure the set of municipalities choice does not drive our results estimation. However, we provide evidence that the results are robust to using these authors' optimal bandwidth. We further show that the optimal bandwidth is close to 10,000 for most of the outcomes. We use a triangular kernel to put more weight on the observations close to the discontinuity but provide evidence that results are unchanged if we use a rectangular kernel. We cluster standard errors at the municipality-level to allow for arbitrary correlation of municipalities' error term across time.

[13]There is a small number of contracts signed in 2010 but their number is negligible.

[14]See Burlig & Preonas (2016) and Asher & Novosad (2020) for recent examples of papers using eligibility rules to obtain RD estimates of large-scale government programs.

The effects of the MCMV program estimated using equation (1) pool together the effects of houses and their construction on infant health. To separate these effects, we estimate period-specific effects of the MCMV on infant health by estimating the following equation:

$$Y_{its} = \beta_0 + \sum_{k=2011}^{2017} \beta_{1k} T_{is} \times 1[\text{Year} = k] + \beta_2 f(P) + \beta_3 Y_{is}^{I} + \gamma X_{is} + \eta_s + \eta_t + \nu_{is}, \qquad (2)$$

The coefficients β_{1k} measures the difference in municipalities' outcomes just below and above the 50,000 inhabitants threshold. If the program's effects on infant health operate through houses, we expect this coefficient to increase through time. If the effects of the program on infant health operate through house construction, we expect this coefficient to die out as houses' construction diminishes.

5.1 Threats to the Validity

Our regression discontinuity estimates have causal interpretations under three assumptions. First, the RD design requires it is not possible to manipulate the running variable at the threshold. This is an important concern in our setting since it is not clear about the municipality population the government uses in the MCMV program. We opt to use the official count of the population from 2007 as our running variable. This was the most recent source of population data when the program was announced. While using the past population might add noise to our estimates, it ensures municipalities could not manipulate the running variable. Figure 4 provides evidence that the running variable's distribution is smooth around the cut-off. In Panel A, we show that the number of municipalities falls smoothly with the municipality size. In panel B, we formally test the difference in the distribution near the cut-off using the McCrary test (McCrary, 2008). This test examines whether there is a discontinuity in the distribution of the running variable around the cut-off. Its test statistic is 0.268 (s.e. = 0.246). Hence, there is no evidence that our assignment was manipulated.

Second, the RD design requires continuity of the municipality outcomes other than the number of houses built under the MCMV program at the threshold. Table 1 provides support for this hypothesis. It reports that socioeconomic characteristics (measured in 2010) and infant health (measured in 2009 and 2010) are similar for municipalities slightly below and above the 50,000 inhabitants cut-off. Not only the coefficients are insignificant, but their magnitude is typically small.

Third, the RD design required no other policies which change close to the 50,000 inhabitants threshold. We mapped two other policies using population cut-offs near this threshold to determine its investments: the sanitation investments from the PAC (*Programa de Aceleração do Crescimento*) and the transfers from the *Fundo de Participação dos Municipios (FPM)*.

The sanitation investments from the PAC prioritizes municipalities with a population below 50,000. Thus, to the extent the PAC effectively improves the sanitation of these municipalities, it might improve the infant health of the municipalities below 50,000. This implies the PAC might bias downward our estimates of the effects of the MCMV on infant health. However, we believe this effect would be relatively small as larger municipalities are more likely to have sanitation services provided by better-capitalized state companies than small municipalities. (Estache et al., 2016; Kresch, 2017). Moreover, there is a concern that PAC investments were not well executed. According to (Ceri, 2016), from 2007 to 2015, the execution of the "Sanitation for All" under the PAC program was slow – for contracts with execution duration between three and five years, the proportion of completed projects was less than 10%. In March 2016, approximately nine years after the start of the Program, 66% of the total projects were not completed.

The transfer from the FPM also prioritizes smaller municipalities. In particular, municipalities with a population below 50,940 received more transfers per capita than municipalities with a population above this population. This might increase local income (e.g., Corbi et al. (2019)) and public goods provision, thereby improving infant health and biasing our estimates downwards. However, these effects are likely to be insignificant because the change in per capita transfers at this cut-off is small and because FPM transfers do not improve public goods provision (see Gadenne (2017)). To strengthen this conclusion, we provide evidence that FPM transfers do not change discontinuously around the 50,000 inhabitants threshold (coef. = 4.76, p-value= 0.438) and that controlling for FPM transfers does not influence our results, as shown in Section 6.

6 Results

We present our results in two parts. We begin exploring MCMV's contract-level data to document the 50,000 population threshold's effects on this program's investments. We then use official birth, hospitalization, and mortality records to document the effects of the MCMV on health at birth and on morbidity and mortality of children under 5 years.

6.1 Housing Investments

Figure 5 graphically presents the regression discontinuity estimates of equation (1). It plots the residuals from a regression of the dependent variable on the controls and state fixed effects on different bins of population size and a linear fit of the relationship between the residuals and population at each side of the 50,000 inhabitants threshold. Panel A depicts the residuals from the total number of units delivered by the MCMV in 2011-2017, and Panel B the residuals from this number divided by the housing deficit. Both panels provide clear evidence that the MCMV investments increase discontinuously at the 50,000 inhabitants threshold. The jump is driven neither by the functional form nor by observations in specific parts of the distribution's support.

Table 2 reports numerical estimates of equation (1) using the number of units delivered by the MCMV in the period 2009-2017 (Panel A) and this number divided by the housing deficit (Panel B) as dependent variables. Columns 1-2 report estimates obtained using a triangular kernel and a bandwidth of 10,000 inhabitants; columns 3-4 estimates obtained using a rectangular kernel and a bandwidth of 10,000 inhabitants, columns 5-6 estimates obtained using a triangular kernel and the optimal bandwidth of Calonico et al. (2014).[15] Odd columns include state fixed effects as controls. In contrast, even columns include state fixed effects and initial municipality characteristics as controls.

We find that the number of units delivered by the MCMV program increases discon-

[15] We use the all municipalities to obtain the optimal bandwidth.

tinuously at the 50,000 inhabitants by about 250-350 units (14-18% of the housing deficit in 2010). The mean number of units below the threshold is close to 170, implying the number of units delivered by the MCMV program effectively triples at the 50,000 inhabitants threshold. This effect is robust and statistically significant at the 5% levels regardless of the specification.

6.2 Health at Birth

Main Results. Figure 6 provides graphical evidence that birth weight jumps discontinuously at the 50,000 inhabitants threshold. It plots the residuals from a regression of the birth weight on the controls and state fixed effects on different bins of population size and a linear fit of the relationship between the residuals and population at each side of the 50,000 inhabitants threshold. The discontinuity is clear and does not seem to be driven by the functional form.

Table 3 provides the corresponding numerical estimates of the relationship shown in the figure. Panel A reports estimates obtained using a triangular kernel and a bandwidth of 10,000 inhabitants, panel B estimates obtained using a uniform kernel and a bandwidth of 10,000 inhabitants, and panel C estimates obtained using a triangular kernel and the optimal bandwidth of Calonico et al. (2014). Column 1 reports estimates obtained controlling only by state and year fixed effects, column 2 adds the birth weight in the initial period as an additional control, and column 3 adds other initial municipality characteristics as controls.

The effects of the discontinuity on birth weight are imprecisely estimated in the specifications without controls (column 1). The coefficients change a lot depending on the kernel and bandwidth chosen and are not statistically significant at the usual levels. This is common in settings in which the dependent variable is measured with error (e.g., Burlig & Preonas (2016)), emphasizing the importance of controlling for initial conditions as discussed in section 5.

Including controls stabilizes the coefficients and increases their precision (columns 2 and 3). This effect becomes extremely robust and statistically significant at the 5% levels regardless of the specification. Quantitatively, we find that birth weight increases by about 12.8 to 15.6 grams at the threshold. The mean birth weight below the threshold is about 3214 grams, implying the 50,000 inhabitants threshold increases the weight on average in 0.4-0.5%. Our effects on birth weight are slightly below the effects of fasting with Ramadan on birth weight found by Almond & Mazumder (2011) and to the effects of job losses through announced notices during pregnancy found by Carlson (2015).[16,17]

To further understand the impacts of the MCMV on birth outcomes, Table 4 reports estimates of equation (1) for other measures of health at birth. It uses our preferred specification presented in Panel A, column 2 of Table 3 – triangular kernel, 10,000 inhabitants bandwidth, and the controls discussed in section 5.

In Panel A, we examine the effects of the MCMV on other measures of birth weight. This is important for interpreting the effects discussed before because the literature emphasizes that the long-run effects of low birth weight are typically driven by events in the the lower tail of the weight distribution (Almond et al., 2018). Column 1 estimates the effect of the MCMV on the share of births below 1500 grams. The point estimate is negative but not statistically significant at the usual levels (p-value = 0.24). Column 2 examines the effect of the MCMV on the share of births below 2000 grams. The point estimate is negative and statistically significant at the 10% level (p-value = 0.06). Column 3 estimates the effect of the MCMV on the share of births below 2500 grams. The findings from columns

[16] Almond & Mazumder (2011) find that birth weight is about 18 grams lower for Arab-pregnancies that overlap with Ramadan (0.6% of the mean). Carlson (2015) find that that the effect of job losses through announced notices ranges from -15 to -20 grams (-0.4% to - 0.6%).

[17] As discussed in section 5, municipalities with a population below 50,940 received more transfers per capita from FPM than municipalities with a population above this threshold. This might increase local income (e.g., Corbi et al. (2019)) and public goods provision, thereby improving infant health and biasing our estimates downwards. Table B2 provides evidence that controlling for FPM transfers does not influence our results.

1-3 indicate that the effect documented in Table 3 is driven by changes in the incidence of low (< 2000 grams) but not very low birth weight events (< 1500 grams). Column 4 further documents a 1 p.p. reduction in children's share of births relatively small for their gestational age (SGA). This effect is significant at the 1% level. This indicates that the effects of MCMV on birth weight do not operate simply by increasing gestational age but also by increasing weight conditional on age.

In Panel B, we examine effects of the MCMV on other markers of health at birth. Columns 1 documents a significant decrease in the share of pregnancies below 32 weeks. The effect size is 0.2 p.p, which corresponds to 13% of the mean. Column 2 finds no significant effect of the MCMV on the share of pregnancies below 37 weeks. The coefficient is negative but not statistically significant. This is suggestive that reductions in the incidence of very premature births are other mechanisms linking MCMV investments and birth weight. Columns 3-4 examine the effects of the MCMV on the share of births with APGAR scores below 7. Point estimates are negative for APGAR1, and APGAR5 with magnitudes between 13-17% of the outcomes mean in the municipalities below the cutoff. However, the coefficients are not statistically significant (p-values = 0.144 and 0.175, respectively).

Figures 7 and 8 present the corresponding RD figures of the estimates presented in Table 4. The discontinuities are not so apparent for the other outcomes as for birth weight. The exception is the effect on the share of births small for their gestational age for which the discontinuity is visibly apparent.

Dynamics. The effects of the MCMV on birth weight reflect a combination of changes in houses and labor market conditions. While it is impossible to disentangle these two mechanisms, it is possible to use the timing of the effects to disentangle between the effects of houses and the effects of the construction of the houses on labor market conditions. As discussed in section 3, the program is expected to temporarily improve labor market outcomes during the construction of the houses and permanently improve living conditions after the houses are delivered. Thus, we expect the effects of the construction to be stronger in the program's early periods (when the program's investments are at their peak, but the number of units delivered is modest) and the effects of houses to be stronger in the program's final periods (when the program's investments fall but the number of units delivered is considerable).

We obtain period-specific effects of the MCMV on health at birth and infant health by estimating equation (2). Figure 9, Panel A plots the estimated coefficients β_{1k} for birth weight. It shows that MCMV effects increase weakly through time going from an statistically insignificant effect of 8.12 grams in 2011 (p-value = 0.257) to a statistically significant effect of than 20.9 grams in 2017 (p-value = 0.010). This increase in the effects over time is suggestive that the effects of the MCMV on birth weight operate primarily through increases in living conditions.

To gain further insight on the mechanisms, we decompose the total effect of the MCMV on birth weight on the effects of houses and their construction using a exercise similar to the one proposed by Dix-Carneiro et al. (2018). As detailed in Appendix A, under the hypothesis that the relationship between birth weight, house construction, and housing conditions is constant over time, it is possible to use the RD coefficients obtained in different periods to determine the role of houses and their construction. This decomposition indicates that the number of houses delivered explains between 60.4-66.5% of the mean effect of the MCMV on birth weight in the period 2011-2017. The contribution of better houses effect increases from 37.2% in the first years of the program to 81.8% in the final years of the program.

Figure 9, panels B-D plots the estimated coefficients β_{1k} of equation (2) for other birth outcomes. Panel B reports results using the share of births below 2000 grams as dependent variable. The effects on the share of births below 2000 grams do not have a clear dynamic. The effect declines between 2011-2016, but reverts in 2017. The effects are significant only in 2015 (0.23 p.p., p-value = 0.03) and 2016 (0.38 p.p., p-value = 0.001). Panel C depicts results obtained using the share of births below 32 weeks. The dynamics of the effects is similar to the effects on birth weight. The effects on the share of pre-term births increase in absolute value through time until 2016, going from statistically insignificant effects of 0.05 percentage points in 2011 (p-value = 0.62) to a statistically significant effect of 0.30 percentage points in 2016 (p-value = 0.013). Panel D reports results obtained using the share of births of children relatively small for their gestational age. The dynamic of this effect is also consistent with the effects on birth weight with the magnitude of the effect increasing over time. In 2011, this effect is -0.87 percentage points (p-value = 0.187), while in 2017 it is -1.24 percentage points (p-value = 0.009). Taken together, the timing of the effects on birth outcomes suggests the effects of the program are driven primarily by houses and not by a temporary increase in labor market income that may have occurred during the construction.

6.3 Morbidity and Mortality of Children Under 1 Year

Table 5 reports regression discontinuity estimates of equation (1) using measures of morbidity and mortality of children under 1 year as the outcomes of interest. We report results based on our preferred specification presented in column (2) of Table 3 – triangular kernel, 10,000 inhabitants bandwidth, and the controls discussed in section 5. As discussed in section 3, the MCMV investments might improve health during the childhood by increasing the share of households living in houses with proper bathrooms, tile floors, and adequate ventilation as well as with proper sanitation.

Panel A depicts the results for hospitalization rates. Column 1 uses total hospitalization rates (per 1,000) as the dependent variable. We find no effect of the MCMV on this measure. The point estimate is negative but economically small and statistically insignificant. Column 2-6 reinforces this conclusion by looking at hospitalization rates for specific causes – infectious and parasitic diseases, respiratory diseases, and perinatal conditions.[18] Column 2 focuses on these three leading causes of infant diseases and columns 3-5 focus on each of these causes separately (infectious and parasitic diseases, respiratory diseases, and perinatal conditions, respectively). Column 6 presents the estimates for the residual causes. We find no effect on these measures.

Panel B reports the results for death rates. Column 1 reports negative but statistically insignificant effects of the MCMV on children's mortality under 1 year. Columns 2-6 find this negative effect is entirely driven by a statistically significant reduction in perinatal deaths. Our coefficient indicates that the MCMV reduces perinatal deaths by -1.06 per 1,000 births. This represents 14% the mean and implies the program reduced in 0.8 the number of deaths per year due to perinatal conditions in the typical municipality to the left of the cutoff.

The reduction in perinatal deaths is consistent with the positive effects on health at birth previously documented. Indeed, the literature suggests that exposure to environmental hazards such as inadequate sanitation and nutrition (itself related to poor sanitation) constitute substantial risks to infant health, increasing the mortality rate for low-birth-weight and preterm infants (Prüss-Üstün & Corvalán, 2006; Zhang et al., 1992; Longnecker et al., 2001). Thus, our findings are suggestive that, by improving the environment in which the households live, the MCMV improved the quality of births and decreased the likelihood of deaths due to perinatal conditions.

Figures B1 and B2 from the appendix B report the corresponding RD figures of the estimates presented in Table 5. They reinforce the conclusions of this table. Hospitalizations are continuous at the 50,000 population threshold, while death rates decrease at this threshold, mostly due to the decrease in deaths due to perinatal conditions.

Figures B3 and B4 from the appendix B plot the estimated coefficients β_{1k} from equa-

[18] The theory indicates that housing conditions might be particularly affected by infectious and parasitic diseases, respiratory diseases, and perinatal conditions (Organization et al., 2019)

tion 2 using hospitalizations and deaths as dependent variables. As expected by the results on Table 5, Figure B3 shows no statistically significant effect on the hospitalization for infants over time, both overall and for specific diseases. The results of infant mortality also show no statistical effect on overall infant deaths over time. However, the estimated effect on mortality originated by perinatal origin is statistically different from zero in 2013 and again in 2016 and 2017.

6.4 Morbidity and Mortality of Children Between 1 and 5 Years

Table 6 reports regression discontinuity estimates of equation (1) using measures of morbidity and mortality of children between 1 and 5 years as the outcomes of interest. The Table is identical to Table 5, except that we combine mortality due to perinatal conditions with mortality due to other diseases. We do this because there are too few hospitalization events and deaths due to perinatal conditions in this age group (the mean of deaths by perinatal conditions is 0.003 per thousand births for this age group).

We find no effects of the MCMV on hospitalization and deaths of children between 1 and 5 years. Point estimates are economically small and statistically insignificant for all measures considered. Figures B5, B6, B7, and B8 from the appendix B report the corresponding RD figures and estimates by period of the effects presented in Table 6. They reinforce the conclusions of the table. Hospitalizations and deaths of children are continuous at the 50,000 population threshold.

7 Conclusion

In this paper, we examine the effects of housing conditions on health at birth, exploring exogenous changes induced from investments of the MCMV Program. This program built about 900,000 houses to poor households in Brazil during the period 2010-2017. We obtain causal estimates of the construction of these houses exploring differences in the MCMV rules that facilitated municipalities with a population above 50,000 inhabitants to obtain funds from this program.

Using regression discontinuity design and administrative data, we estimate the program's effects on signed contracts under the program, health at birth and infant health. We find that the number of houses delivered by the program increases by 300-350 units during the period 2011-2017 at the 50,000 inhabitants threshold. This corresponds to 14-18% of the housing deficit of the typical municipality to the left of the discontinuity. We find the increase in MCMV investments led to increases of 12-16 grams in birth weight and decreases of 1 per 1,000 live births in infant (before 1 year) mortality caused by conditions originating in children's perinatal period. We find no effect of the program in children with more than one year. Decomposition exercises indicate that most of this effect is due to improvements in houses (as opposed to improvements in labor market conditions coming from the program's investments).

Health at birth is an important determinant of physical and mental health, human capital accumulation, and income (Gluckman et al., 2005; Cunha & Heckman, 2007; Currie, 2009). Thus, understanding its determinants is fundamental to guide public policies. Nevertheless, while there is a growing body of empirical work documenting the role of shocks during fetal development on health at birth (e.g., (Almond & Currie, 2011; Almond et al., 2018)), there is much less evidence on the role of the environment on health at birth.

Our results contribute to this literature by documenting the importance of better houses to improve fetal development and, consequently, health at birth. These results imply housing policies can have important health externalities. For instance, comparable effects on birth weight increases earnings in the long run by 1.7% (Bharadwaj et al., 2014). Understanding whether these health externalities influence the optimal design of housing policies is an important agenda for future research.

Moreover, it is important to more clearly disentangle the mechanisms behind the effects of MCMV investments on health at birth. Assessing the program's effects on local income, housing quality, and housing costs as well as understanding the heterogeneity of the program's effects with respect to mother characteristics (e.g., schooling and age) are also important agendas for future research.

References

Adhvaryu, A., Fenske, J. & Nyshadham, A. (2019). Early life circumstance and adult mental health. *Journal of Political Economy*, *127*(4), 1516-1549. Retrieved from https://doi.org/10.1086/701606 doi: 10.1086/701606

Almond, D. & Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of economic perspectives*, *25*(3), 153–72.

Almond, D., Currie, J. & Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, *56*(4), 1360–1446.

Almond, D. & Mazumder, B. (2011). Health capital and the prenatal environment: the effect of ramadan observance during pregnancy. *American Economic Journal: Applied Economics*, *3*(4), 56–85.

Alsan, M. & Goldin, C. (2019). Watersheds in child mortality: the role of effective water and sewerage infrastructure, 1880–1920. *Journal of Political Economy*, *127*(2), 586–638.

Asher, S. & Novosad, P. (2020). Rural roads and local economic development. *American Economic Review*, *110*(3), 797–823.

Baird, S., Friedman, J. & Schady, N. (2011). Aggregate income shocks and infant mortality in the developing world. *The Review of Economics and Statistics*, *93*(3), 847-856. Retrieved from https://doi.org/10.1162/REST_a_00084 doi: 10.1162/REST_a_00084

Barnhardt, S., Field, E. & Pande, R. (2017). Moving to opportunity or isolation? network effects of a randomized housing lottery in urban india. *American Economic Journal: Applied Economics*, *9*(1), 1–32.

Belchior, C. A. (2019). *Two essays on housing programs and the labor market* (Unpublished doctoral dissertation).

Bharadwaj, P., Johnsen, J. V. & Løken, K. V. (2014). Smoking bans, maternal smoking and birth outcomes. *Journal of Public economics*, *115*, 72–93.

Biderman, C., Acosta, C., Ramos, F., Hiromoto, M., Terentim, G., Caldeira, J. & Pacheco, T. (2019). Morar longe: o programa minha casa minha vida e a expansão das regiões metropolitanas. *São Paulo: Instituto Escolhas. Recuperado em*, *15*.

Burlig, F. & Preonas, L. (2016). Out of the darkness and into the light? development effects of rural electrification. *Energy Institute at Haas WP*, *268*, 26.

Calonico, S., Cattaneo, M. D. & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295–2326.

Cappelletty, D. (1998). Microbiology of bacterial respiratory infections. *The Pediatric infectious disease journal*, *17*(8), S55–S61.

Carlson, K. (2015). Fear itself: The effects of distressing economic news on birth outcomes. *Journal of health economics*, *41*, 117–132.

Cattaneo, M. D., Galiani, S., Gertler, P. J., Martinez, S. & Titiunik, R. (2009). Housing, health, and happiness. *American Economic Journal: Economic Policy*, *1*(1), 75–105.

Ceri, F. (2016). Efetividade dos investimentos em saneamento no brasil. *Centro de Estudos em Regulação e Infraestrutura*, 42.

Chetty, R., Hendren, N. & Katz, L. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, *106*(4), 855–902.

Corbi, R., Papaioannou, E. & Surico, P. (2019). Regional transfer multipliers. *The Review of Economic Studies*, *86*(5), 1901–1934.

Cunha, F. & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47.

Currie, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *JoUrnal of economIc lIteratUre*, *47*(1), 87–122.

Cutler, D. & Miller, G. (2005). The role of public health improvements in health advances: the twentieth-century united states. *Demography*, *42*(1), 1–22.

Dix-Carneiro, R., Soares, R. R. & Ulyssea, G. (2018). Economic shocks and crime: Evidence from the brazilian trade liberalization. *American Economic Journal: Applied Economics*, *10*(4), 158–95.

Duflo, E., Greenstone, M., Guiteras, R. & Clasen, T. (2015). *Toilets can work: Short and medium run health impacts of addressing complementarities and externalities in water and sanitation* (Tech. Rep.). National Bureau of Economic Research.

Estache, A., Garsous, G. & da Motta, R. S. (2016). Shared mandates, moral hazard, and political (mis) alignment in a decentralized economy. *World Development*, *83*, 98–110.

Ezzati, M., Bailis, R., Kammen, D. M., Holloway, T., Price, L., Cifuentes, L. A., . . . Dhanapala, K. N. (2004). Energy management and global health. *Annu. Rev. Environ. Resour.*, *29*, 383–419.

Field, E. (2007). Entitled to work: Urban property rights and labor supply in peru. *The Quarterly Journal of Economics*, *122*(4), 1561–1602.

Franklin, S. (2019). The demand for government housing: Evidence from lotteries for 200,000 homes in ethiopia. In *Technical report.* Working Paper.

Furtado, B. A., Lima Neto, V. C. & Krause, C. (2013). Estimativas do déficit habitacional brasiliero (2007-2011) por municípios (2010).

Gadenne, L. (2017). Tax me, but spend wisely? sources of public finance and government accountability. *American Economic Journal: Applied Economics*, 274–314.

Gale, R. (2018). Housing mobility programs and health outcomes. *Health Affairs, Health Policy Brief*.

Galiani, S., Gertler, P. J., Undurraga, R., Cooper, R., Martínez, S. & Ross, A. (2017). Shelter from the storm: Upgrading housing infrastructure in latin american slums. *Journal of urban economics*, *98*, 187–213.

Galiani, S. & Schargrodsky, E. (2004). Effects of land titling on child health. *Economics & Human Biology*, *2*(3), 353 - 372. Retrieved from http://www.sciencedirect.com/science/article/pii/S1570677X04000541 (Special Issue on Child Health in Latin America) doi: https://doi.org/10.1016/j.ehb.2004.10.003

Galiani, S. & Schargrodsky, E. (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics*, *94*(9-10), 700–729.

Gelman, A. & Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, *37*(3), 447–456.

Glick, P. (2002). Women's employment and its relation to children's health and schooling in developing countries: conceptual links, empirical evidence, and policies. *Cornell Food and Nutrition Policy Program Working Paper*(131).

Gluckman, P. D., Hanson, M. A., Spencer, H. G. & Bateson, P. (2005). Environmental influences during development and their later consequences for health and disease: implications for the interpretation of empirical studies. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1564), 671–677.

Gould, E. D., Lavy, V. & Paserman, M. D. (2011). Sixty years after the magic carpet ride: The long-run effect of the early childhood environment on social and economic outcomes. *The Review of Economic Studies*, *78*(3), 938–973.

Hammer, J. & Spears, D. (2016). Village sanitation and child health: Effects and external validity in a randomized field experiment in rural india. *Journal of Health Economics*, *48*, 135 - 148. Retrieved from http://www.sciencedirect.com/science/article/pii/S0167629616300030 doi: https://doi.org/10.1016/j.jhealeco.2016.03.003

Hoynes, H., Miller, D. & Simon, D. (2015, February). Income, the earned income tax credit, and infant health. *American Economic Journal: Economic Policy*, *7*(1), 172-211. Retrieved from https://www.aeaweb.org/articles?id=10.1257/pol.20120179 doi: 10.1257/pol.20120179

Hutton, G., Haller, L., Water, S., Organization, W. H. et al. (2004). *Evaluation of the costs and benefits of water and sanitation improvements at the global level* (Tech. Rep.). World Health Organization.

Imbens, G. W. & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, *142*(2), 615–635.

Katz, L. F., Kling, J. R. & Liebman, J. B. (2001). Moving to opportunity in boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics*, *116*(2), 607–654.

Kresch, E. P. (2017). The buck stops where? federalism, uncertainty, and investment in the brazilian water and sanitation sector. *American Economic Journal: Economic Policy*.

Kumar, T. (2019). *The intergenerational effects of subsidizing homeownership: evidence from a program in mumbai* (Tech. Rep.). Working Paper.

Lavy, V., Schlosser, A. & Shany, A. (2016). *Out of africa: Human capital consequences of in utero conditions* (Tech. Rep.). National Bureau of Economic Research.

Lee, D. S. & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, *48*(2), 281–355.

Lilford, R. J., Oyebode, O., Satterthwaite, D., Melendez-Torres, G. J., Chen, Y.-F., Mberu, B., ... Ezeh, A. (2017). Improving the health and welfare of people who live in slums. *The Lancet*, *389*(10068), 559 - 570. Retrieved from http://www.sciencedirect.com/science/article/pii/S0140673616318487 doi: https://doi.org/10.1016/S0140-6736(16)31848-7

Longnecker, M. P., Klebanoff, M. A., Zhou, H. & Brock, J. W. (2001). Association between maternal serum concentration of the ddt metabolite dde and preterm and small-for-gestational-age babies at birth. *The Lancet*, *358*(9276), 110–114.

Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R. & Sanbonmatsu, L. (2013). Long-term neighborhood effects on low-income families: Evidence from moving to opportunity. *The American Economic Review*, *103*(3), 226–231.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, *142*(2), 698–714.

Organization, W. H. et al. (2019). *Progress on household drinking water, sanitation and hygiene 2000-2017: special focus on inequalities*. World Health Organization.

Pacheco, T. S. (2019). *Moradia, localização e o programa habitacional" minha casa minha vida" no município do rio de janeiro* (Unpublished doctoral dissertation).

Picarelli, N. (2019). There is no free house. *Journal of Urban Economics*, *111*, 35 - 52. Retrieved from http://www.sciencedirect.com/science/article/pii/S0094119019300245 doi: https://doi.org/10.1016/j.jue.2019.04.002

Prüss-Üstün, A. & Corvalán, C. (2006). Preventing disease through healthy environments: towards an estimate of the environmental burden of disease. *World Health Organization*.

Rocha, R. & Soares, R. R. (2015). Water scarcity and birth outcomes in the brazilian semi-arid. *Journal of Development Economics*, *112*, 72–91.

Squarize Chagas, A. L., Malvezzi Rocha, G. et al. (2019). *Housing program and social conditions impact: Evidences from "minha casa minha vida" program lotteries in brazil* (Tech. Rep.). Núcleo de Economia Regional e Urbana da Universidade de São Paulo (NEREUS).

Strully, K. W., Rehkopf, D. H. & Xuan, Z. (2010). Effects of prenatal poverty on infant health: state earned income tax credits and birth weight. *American sociological review*, *75*(4), 534–562.

Tatagiba, L. & Teixeira, A. C. C. (2016). *União nacional por moradia popular and social innovation in housing policies for the poor in brazil.* ImPRovE Case Study.

Villar, J., Ismail, L. C., Victora, C. G., Ohuma, E. O., Bertino, E., Altman, D. G., ... others (2014). International standards for newborn weight, length, and head circumference by gestational age and sex: the newborn cross-sectional study of the intergrowth-21st project. *The Lancet*, *384*(9946), 857–868.

Vogl, T. S. (2007). Urban land rights and child nutritional status in peru, 2004. *Economics & Human Biology*, *5*(2), 302–321.

Zhang, J., Cai, W.-W. & Lee, D. J. (1992). Occupational hazards and pregnancy outcomes. *American journal of industrial medicine*, *21*(3), 397–408.

Table 1: Summary Statistics and Balance

	Mean	RD	S.E	N
A. Demographics				
Sh. female	0.503	-0.000	(0.003)	235
Sh. youngs	0.355	-0.001	(0.011)	235
Sh. adults	0.539	-0.005	(0.010)	235
Sh. old	0.105	0.005	(0.007)	235
Sh. rural hh	0.237	-0.005	(0.057)	235
Sh. Migrants	0.096	-0.019	(0.013)	235
# holseholds	15038	-474.697	(362.932)	235
# housefolds in deficit	1891	200.105	(159.910)	235
B. Labor and Schooling				
Sh. workers	0.624	-0.022	(0.020)	235
Av. wage	899.6	-15.641	(53.267)	235
less than 9 years	0.652	0.009	(0.021)	235
less than 9-11 years	0.142	-0.006	(0.007)	235
less than 12-15 years	0.159	-0.004	(0.013)	235
16 or more years	0.043	0.002	(0.005)	235
C. Infraestructure				
Sh. hh with water	0.688	0.005	(0.056)	235
Sh. hh with sewage	0.381	-0.033	(0.053)	235
D. Health infraestructure				
# Hospital Beds	99.80	21.246	(30.241)	235
# Hospitals	2.034	0.480	0.480	235
presence of PSF	0.677	-0.065	(0.145)	235
E. Infant Outcomes				
Birth weight	3219	-19.174	(17.819)	235
Low birth (<2500)	0.0684	0.003	(0.005)	235
Apgar5	9.293	0.068	(0.117)	235
Total infant hosp. (up to 1 age)	205.6	0.139	(27.534)	235
infectious	38.40	7.648	(8.080)	235
respiratory	80.09	-2.725	-15.368	235
perinatal	54.15	-0.313	(9.005)	235
total infant death (up to 1 age)	15.36	-0.845	(1.663)	235
infectious	0.778	-0.104	(0.298)	235
respiratory	0.832	0.210	(0.352)	235
perinatal	9.163	-0.583	(1.285)	235

Notes: The table presents mean values for municipality characteristics, measured in the baseline period. Variables from panels A-C come from the 2010 Population Census, while the final three from panel D come from the CNES/Datasus. Panel E come from SINASC, SIH, and SIM (datasus). Column 1 shows the unconditional means for all municipalities, column 2 shows the regression discontinuity estimate, following equation 1, column 3 is the robust standard errors, and column 4 the number of observations. The bandwidth of $\pm$ 10 around the population thresholds has been used to define the sample of municipalities.

Table 2: Effect of Municipality Prioritization on MCMV Investments

	(1)	(2)	(3)	(4)	(5)	(6)
			Paneal A: Contracts			
1[pop >50,000]	251.89**	296.29**	350.32***	354.29***	266.02**	293.55**
	(126.30)	(128.66)	(122.61)	(126.88)	(131.15)	(128.41)
Mean	171.88	171.88	171.88	171.88	166.06	177.53
Observations	235	235	235	235	255	255
RD bandwidth	±10	±10	±10	±10	±10.77	±10.77
			Paneal B: Contracts/deficit			
1[pop >50,000]	0.14 *	0.16**	0.18 **	0.18 **	0.16**	0.16 **
	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)	(0.08)
Mean	0.10	0.10	0.10	0.10	0.08	0.09
Observations	235	235	235	235	314	314
RD bandwidth	±10	±10	±10	±10	±12.33	±12.33
Controls	No	Yes	No	Yes	No	Yes
Kernel	Triangular	Triangular	Uniform	Uniform	Triangular	Triangular

Notes: The table reports estimates of equation (1) using the number of units delivered by the MCMV in the period 2009-2017 (Panel A) and this number divided by the housing deficit (Panel B) as dependent variables. Columns 1-2 report estimates obtained using a triangular kernel and a bandwidth of 10,000 inhabitants; columns 3-4 estimates obtained using a uniform kernel and a bandwidth of 10,000 inhabitants, columns 5-6 estimates obtained using a triangular kernel and the optimal bandwidth of Calonico et al. (2014). Robust standard errors in parenthesis. *** p<0.01; ** p<0.05; * p<0.10.

Table 3: Effect on Weight at Birth

	(1)	(2)	(3)
		Birth Weight	
Panel A: Triangular, BW ±10			
1[pop >50,000]	-1.53	13.84**	15.58**
	(15.55)	(6.49)	(6.54)
Mean	3214.34	3214.34	3214.34
Observations	1645	1645	1645
RD bandwidth	±10	±10	±10
Kernel	Triangular	Triangular	Triangular
Baseline Control	No	Yes	Yes
Controls	No	No	Yes
Panel B: Uniform, BW ±10			
1[pop >50,000]	15.34	14.63**	15.18**
	(13.37)	(6.39)	(6.17)
Mean	3214.34	3214.34	3214.34
Observations	1645	1645	1645
RD bandwidth	±10	±10	±10
Kernel	Uniform	Uniform	Uniform
Baseline Control	No	Yes	Yes
Controls	No	No	Yes
Panel C: Triangular, BW optimal			
1[pop >50,000]	9.30	12.88**	14.71**
	(10.77)	(5.99)	(5.99)
Mean	3216.52	3216.52	3217.91
Observations	3017	1911	1960
RD bandwidth	±15.90	±15.90	±11.49
Kernel	Triangular	Triangular	Triangular
Baseline Control	No	Yes	Yes
Controls	No	No	Yes

Notes: The table presents estimates of equation (1) using the birth weight as dependent variable. Panel A reports estimates obtained using a triangular kernel and a bandwidth of 10,000 inhabitants, panel B estimates obtained using a uniform kernel and a bandwidth of 10,000 inhabitants, and panel C estimates obtained using a triangular kernel and the optimal bandwidth of Calonico et al. (2014). Columns 1 report estimates obtained controlling only for state and year fixed effects, column 2 adds the birth weight in the initial period as an additional control, and column 3 adds other initial municipality characteristics as controls. Standard errors clustered at the municipality level are reported in parenthesis. *** p<0.01; ** p<0.05; * p<0.10

Table 4: Effects on Birth Outcomes

	(1)	(2)	(3)	(4)
Panel A: Health at Birth	$< 1500g$	$< 2000g$	$< 2500g$	Small
1[pop >50,000]	-0.001	-0.002*	-0.001	-0.010***
	(0.000)	(0.001)	(0.002)	(0.004)
Observations	1645	1645	1645	1635
R-squared	0.120	0.225	0.376	0.268
Dep. Variable Mean	0.010	0.022	0.109	0.110
Panel B: Gestation and Apgar	$<$ 32 weeks	$<$ 37 weeks	Low Apgar1	Low Apgar5
1[pop >50,000]	-0.002**	-0.001	-0.016	-0.004
	(0.001)	(0.004)	(0.011)	(0.003)
Observations	1645	1645	1645	1645
R-squared	0.136	0.254	0.581	0.309
Dep. Variable Mean	0.014	0.106	0.135	0.024

Notes: The table reports estimates of equation (1) for several measures of health at birth. It uses our preferred specification – triangular kernel, 10,000 inhabitants bandwidth, and the controls. Panel A reports estimates for the share of births below 1500 grams (column 1), the share of births below 2000 grams (column 2), the share of births below 2500 grams (column 3) and SGA (column 4). Panel B reports the estimates on the share of pregnancies below 32 weeks (column 1), the share of pregnancies below 37 weeks (column 2), and the share of births with APGAR scores below 7 (columns 3 and 4). Standard errors clustered at the municipality level are reported in parenthesis. *** p<0.01; ** p<0.05; * p<0.10.

Table 5: Effects on Morbidity and Mortality of Children Under 1 Year

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Infant Hospitalization	Total	Main	Infectious	Resp.	Perinatal	Residual
1[pop >50,000]	-1.718	-5.558	0.344	-2.171	-2.246	4.914
	(12.763)	(11.364)	(3.730)	(5.364)	(5.734)	(3.167)
Observations	1645	1645	1645	1645	1645	1645
R-squared	0.479	0.470	0.505	0.518	0.520	0.425
Mean	179.268	147.180	26.288	57.414	63.478	32.087
Panel B: Infant Mortality	Total	Main	Infectious	Resp.	Perinatal	Residual
1[pop >50,000]	-0.534	-0.919	0.044	0.098	-1.062**	0.430
	(0.691)	(0.566)	(0.131)	(0.115)	(0.530)	(0.264)
Observations	1645	1645	1645	1645	1645	1645
R-squared	0.190	0.202	0.124	0.114	0.136	0.033
Mean (per 1000)	13.751	9.470	0.701	0.628	8.141	4.281

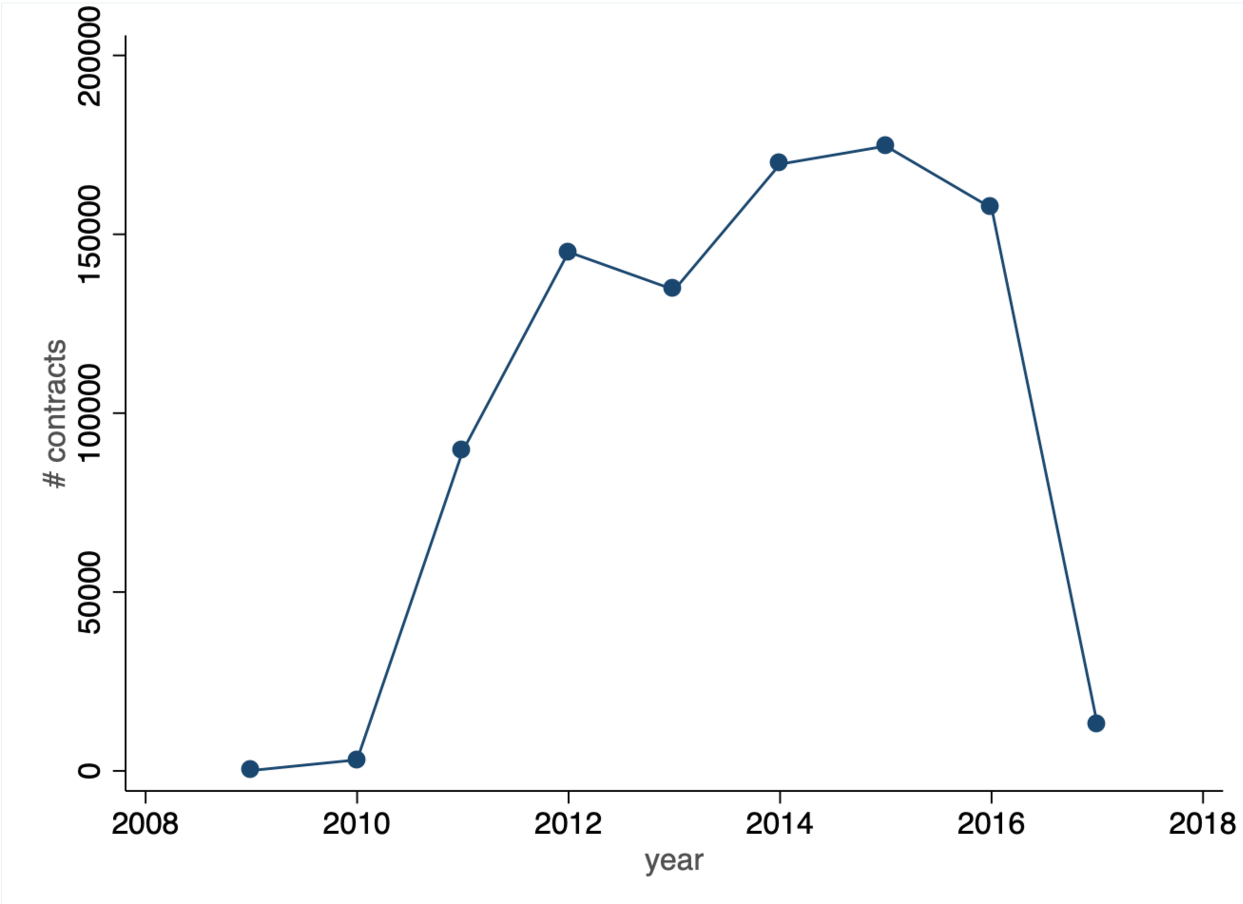
Notes: The table reports regression discontinuity estimates of equation (1) using measures of morbidity and mortality of children under 1 year as the outcomes of interest. We report results based on our preferred specification – triangular kernel, 10,000 inhabitants bandwidth and the controls. Panel A depicts the results for hospitalization rates and panel B reports the results for mortality rates. In Panel A (Panel B), Column 1 uses total hospitalization (mortality) rates per 1,000 births as the dependent variable. Column 2 reports estimates for the combined hospital admission (mortality) due to infectious and parasitic diseases, respiratory diseases, and perinatal conditions. Column 3-6 reports these estimates for hospitalization (mortality) rates for these specific causes, separately. Standard errors clustered at the municipality level are reported in parenthesis. *** p<0.01; ** p<0.05; * p<0.10.

Table 6: Effects Morbidity and Mortality of Children From 1 to 5 Years

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Child Hospitalization	Total	Main	Infectious	Resp	Perinatal	Residual
1[pop >50,000]	-1.708	-2.080	-1.925	0.439	-0.099**	-1.618
	(7.355)	(6.118)	(3.074)	(3.376)	(0.042)	(11.807)
Mean	63.386	42.197	15.653	26.474	0.069	-59.253
Observations	1645	1645	1645	1645	1645	1645
R-squared	0.567	0.604	0.605	0.563	0.251	0.680
Panel B: Child Mortality	Total	Main	Infectious	Resp.	Perinatal	Residual
1[pop >50,000]	-0.011	0.002	0.018	-0.011	-	0.009
	(0.067)	(0.031)	(0.023)	(0.020)	-	(0.084)
Mean (per 1000)	0.653	0.158	0.070	0.088	0.003	0.146
Observations	1645	1645	1645	1645		1645
R-squared	0.149	0.130	0.087	0.044	-	0.157

Notes: Table (6) reports regression discontinuity estimates of equation (1) using measures of morbidity and mortality of children between 1 and 5 years as the outcomes of interest. We report results based on our preferred specification – triangular kernel, 10,000 inhabitants bandwidth and the controls. Panel A depicts the results for hospitalization rates and panel B reports the results for mortality rates. In Panel A (Panel B), Column 1 uses total hospitalization (mortality) rates per 1,000 births as the dependent variable. Column 2 combine mortality due to perinatal conditions with mortality due to other diseases. Column 3-6 reports these estimates for hospitalization (mortality) rates for these specific causes – infectious and parasitic diseases, respiratory diseases, perinatal conditions and residual causes, respectively. Standard errors clustered at the municipality level are reported in parenthesis. *** p<0.01; ** p<0.05; * p<0.10.

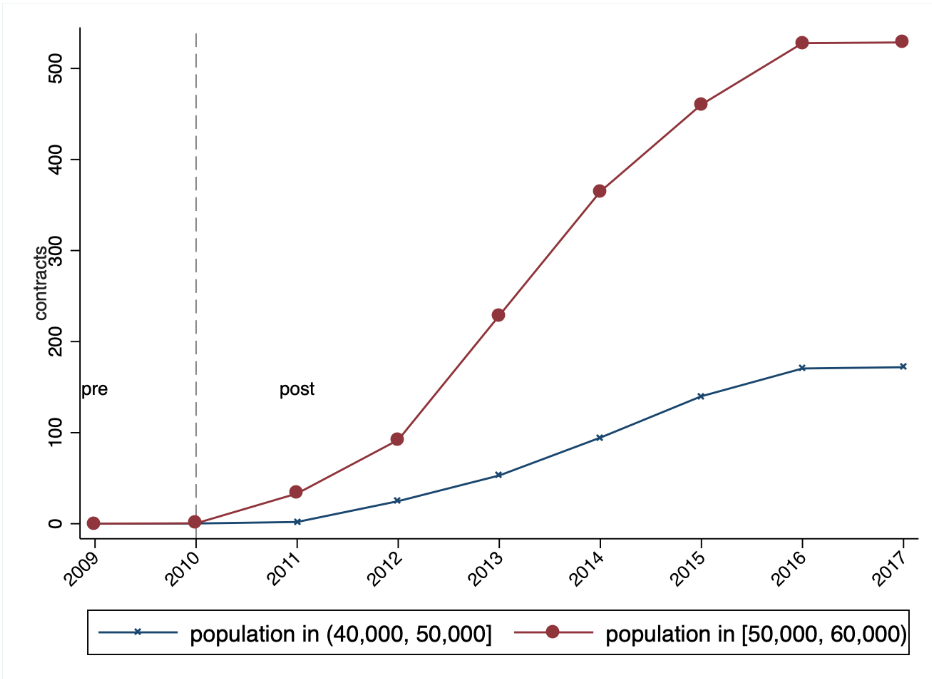
Figure 1: MCMV signed contracts by year (segment 1)



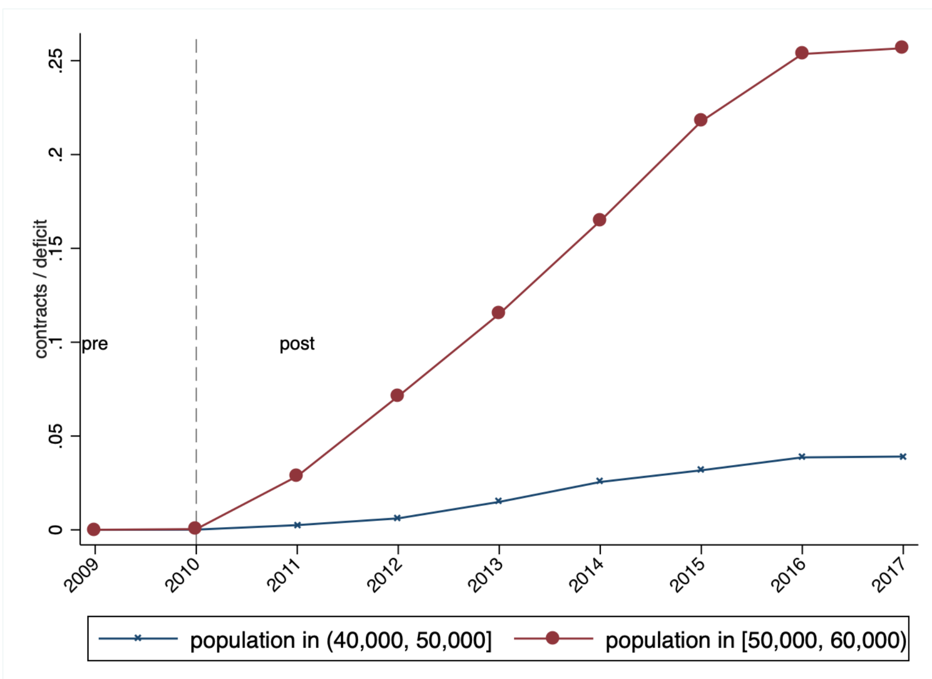
Note: The Figure shows the flow of signed contracts by year in segment 1 of MCMV (MCMV-FAR and MCMV-*Sub50*) for all municipalities in Brazil. The data was obtained from Caixa (2010-2017).

Figure 2: The Roll-Out of the MCMV

(a) Contracts, by year



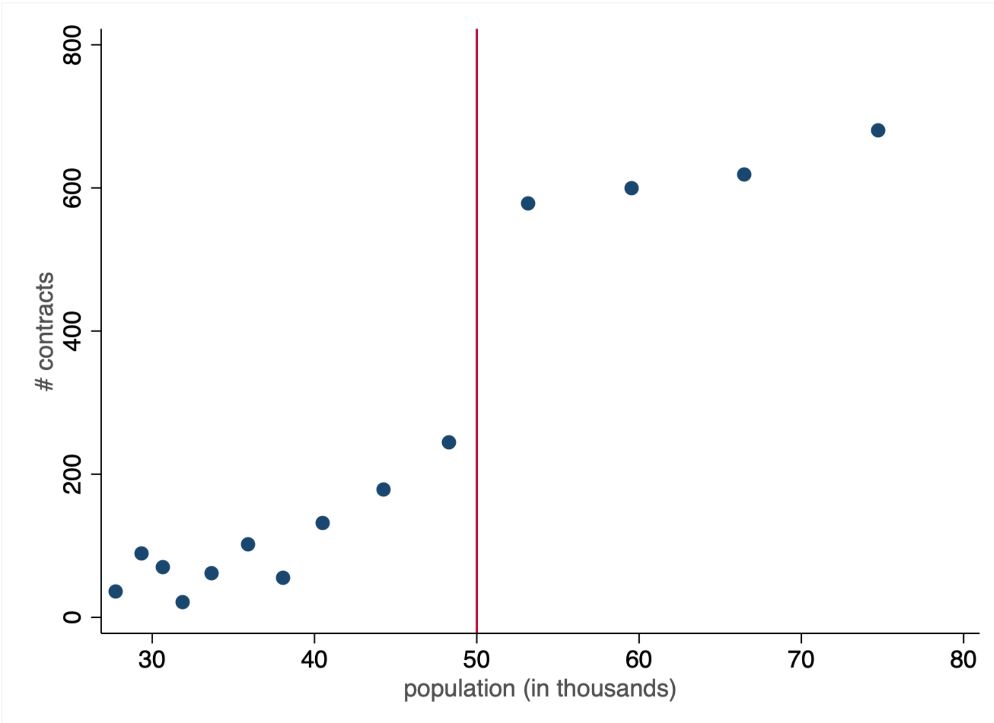
(b) Contracts as a proportion of the households in housing deficit, by year



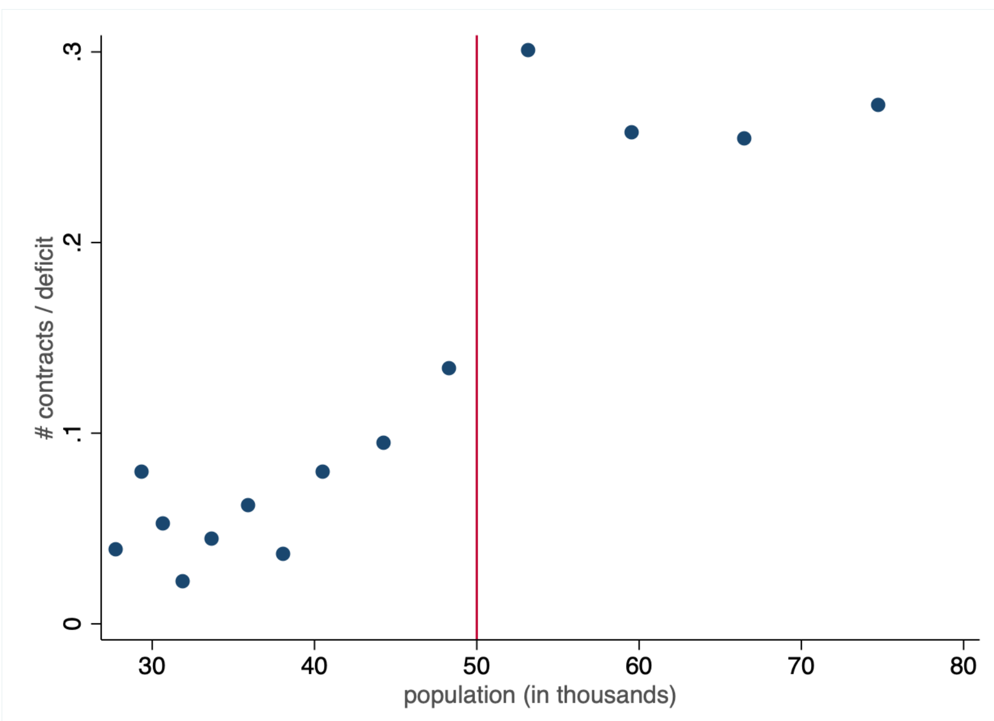
Note: The figure reports the size of the MCMV investments below and above 50,000 inhabitants. Panel A reports the average number of signed contracts by municipality until the year and Panel B reports the share of signed contracts as a proportion of the number of households in housing deficit in 2010. The sample is restricted to observations around the 50,000 population threshold.

Figure 3: Discontinuity on the 50,000 population threshold

(a) Contracts by year



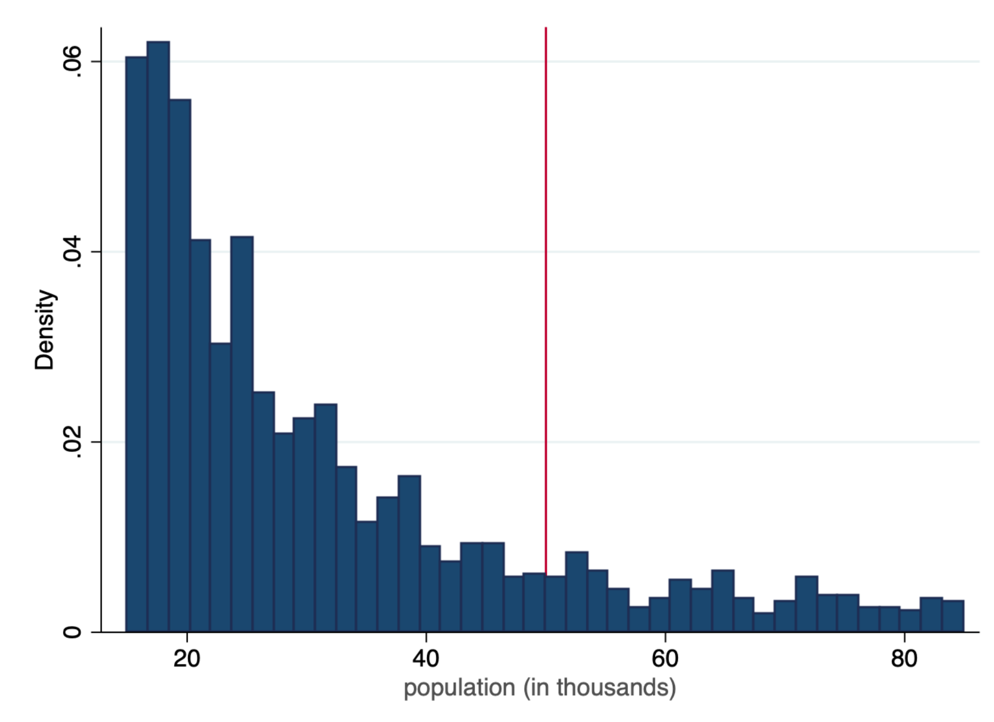
(b) Contracts/HH in deficit



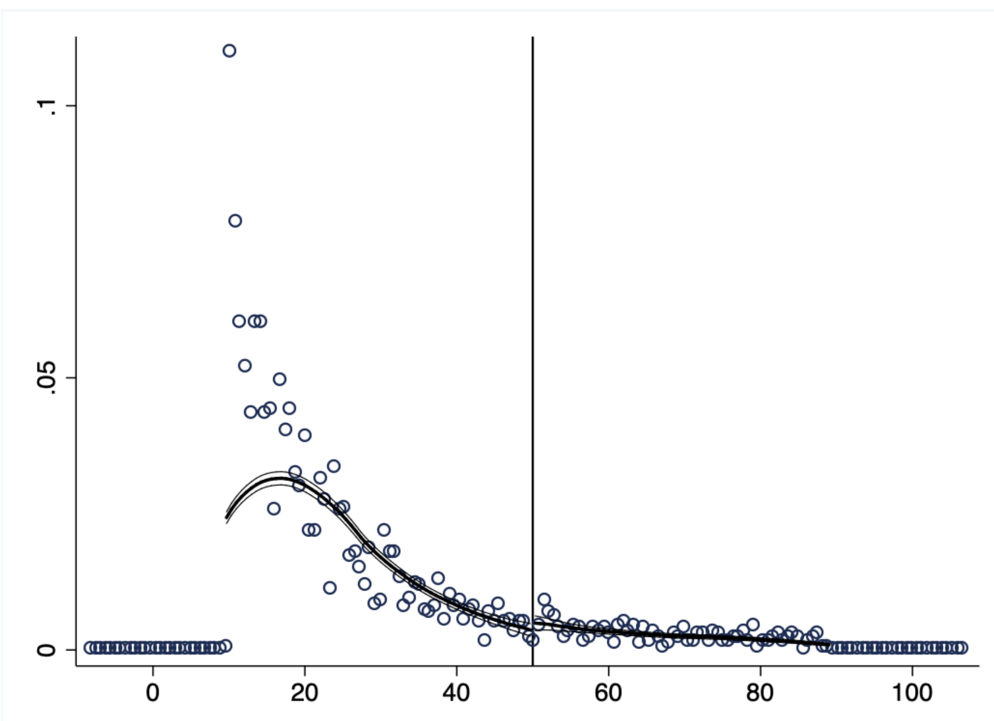
Note: The figure reports bins of the mean number of contracts (Panel A) and contracts as a share of the housing deficit in terms of the population (Panel B).

Figure 4: Histogram and McCrary Test

(a) Running-variable Histogram



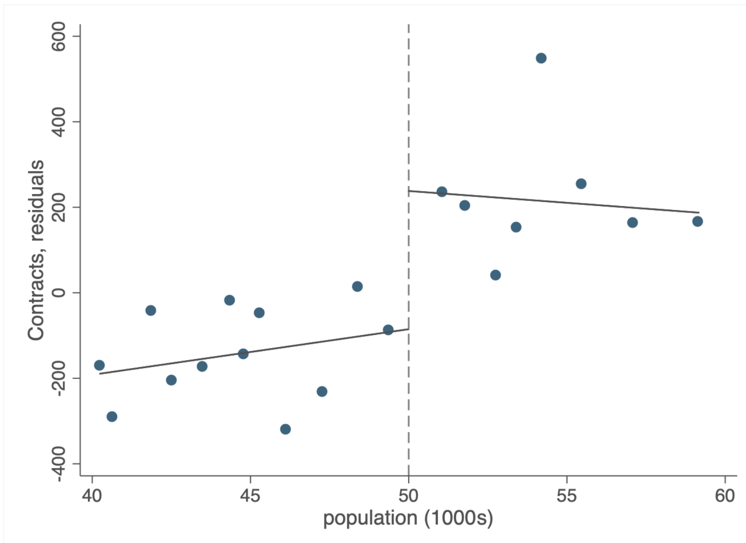
(b) McCrary Test



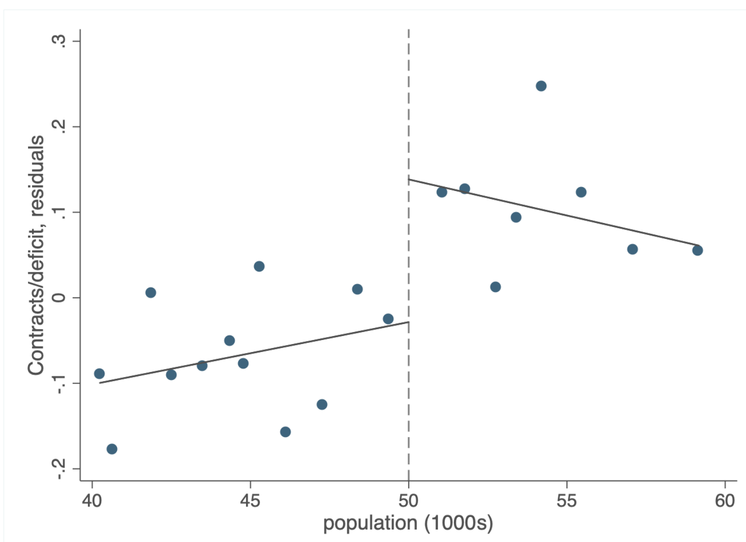
Note: Panel A shows the distribution of Brazilian population (in thousands) of municipalities in 2007. Panel B shows the figure for the McCrary test, which tests whether there is a discontinuity in the data frequency distribution around the cutoff. The McCrary test statistic is 0.268 (s.e. = 0.246).

Figure 5: RD – MCMV Signed Contracts

(a) Contracts

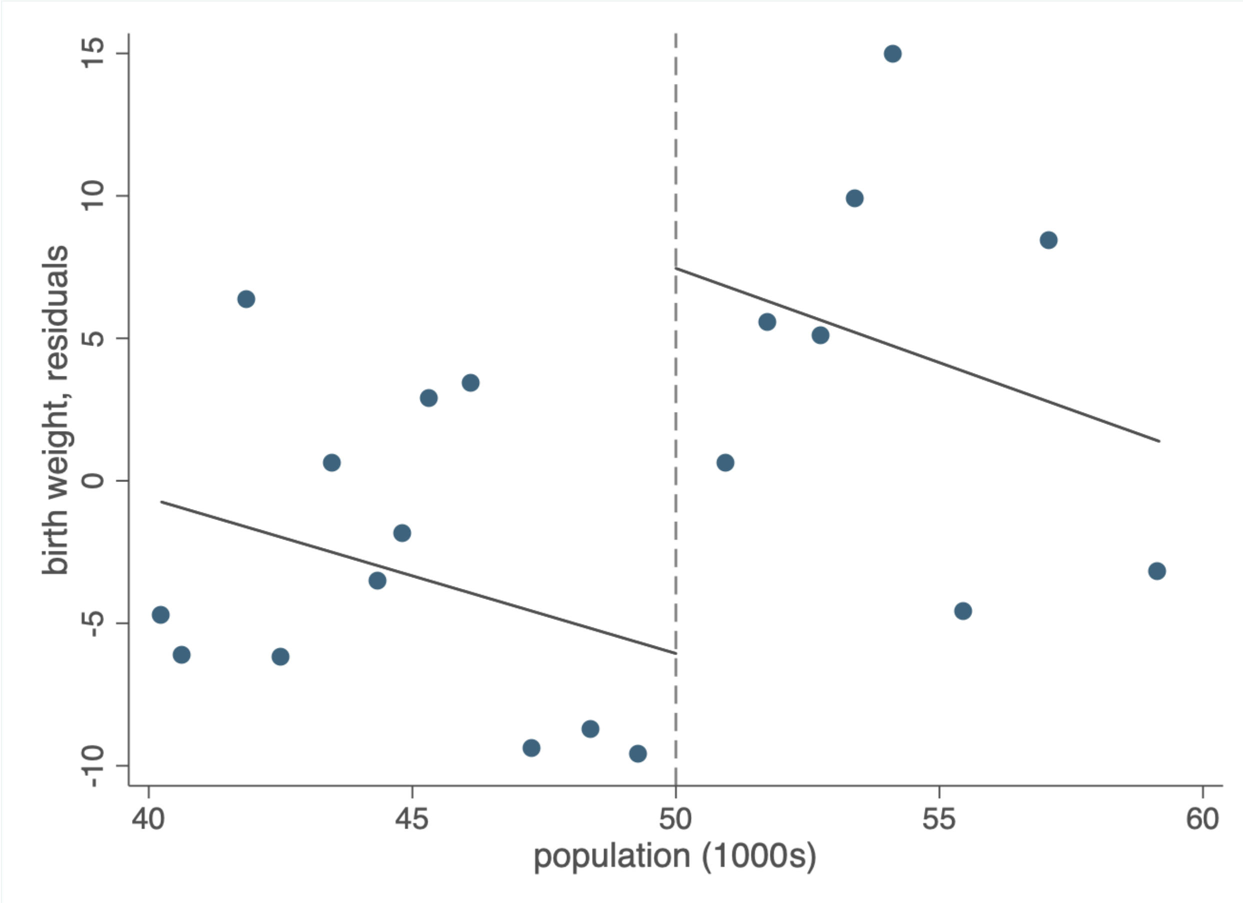


(b) Contracts/Deficit



Note: The figure presents the regression discontinuity estimates of equation (1). It plots the residuals from a regression of the dependent variable on the controls and state fixed effects on different bins of population size and a linear fit of the relationship between the residuals and population at each side of the 50,000 inhabitants threshold. Each dot contains approximately 12 municipalities, averaged in 20 bins. Panel A depicts the residuals from the number of units delivered by the MCMV in 2011-2017, and Panel B, the residuals from this number of units divided by the housing deficit.

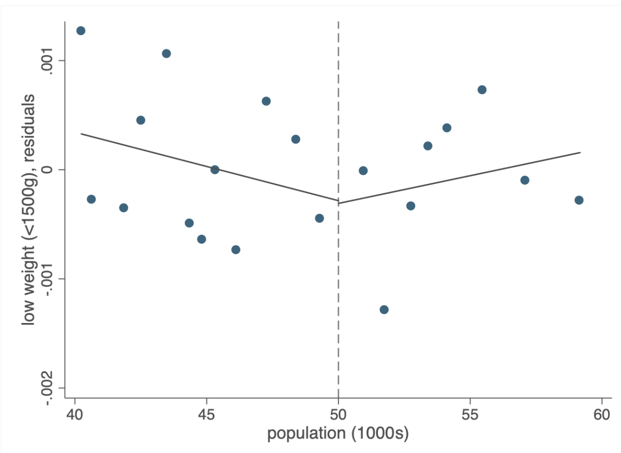
Figure 6: Effects on Birth Weight (g)



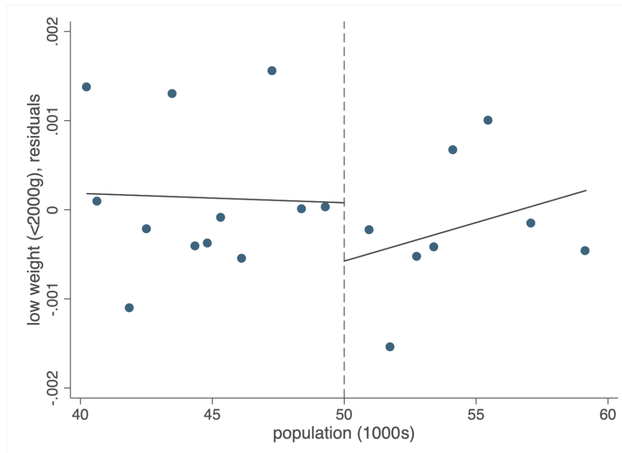
Note: The figure presents the regression discontinuity estimates of equation (1). It plots the residuals from a regression of birth weight on the controls and state and year fixed effects on different bins of population size and a linear fit of the relationship between the residuals and population at each side of the 50,000 inhabitants threshold.

Figure 7: Other Measures on Birth Weight

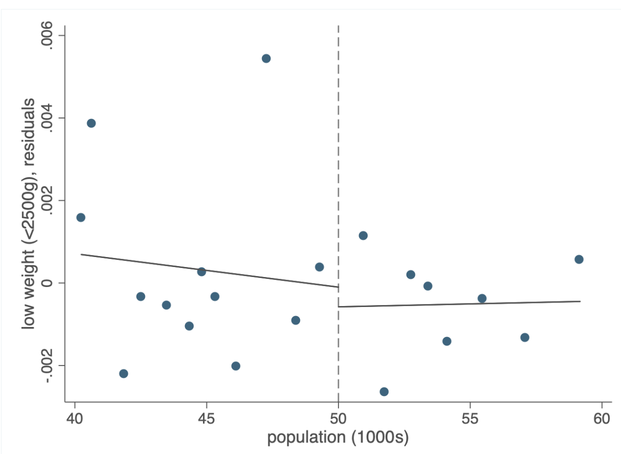
(a) Weight < 1500 (%)



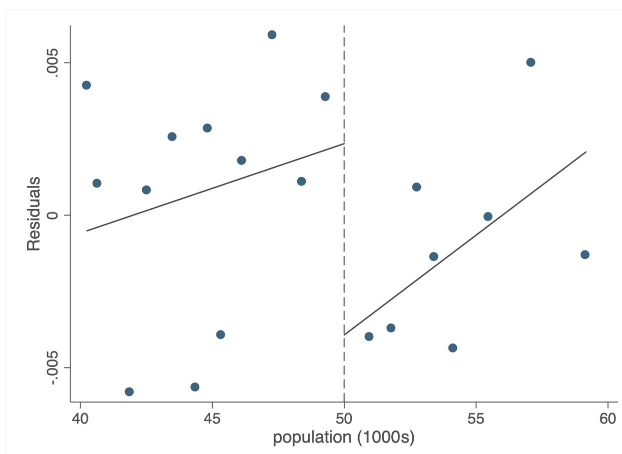
(b) Weight < 2000 (%)



(c) Weight < 2500 (%))



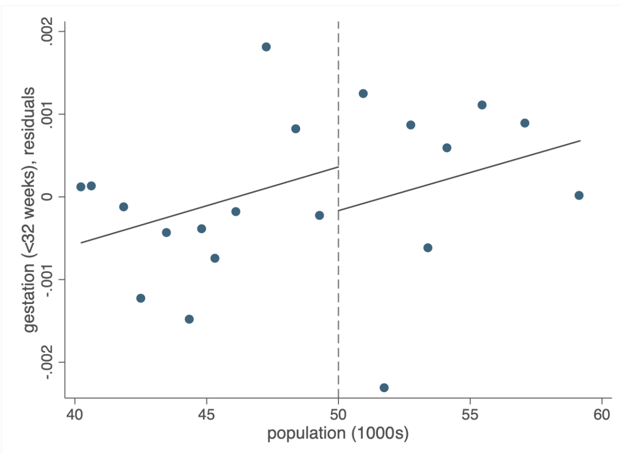
(d) Small for gest. age



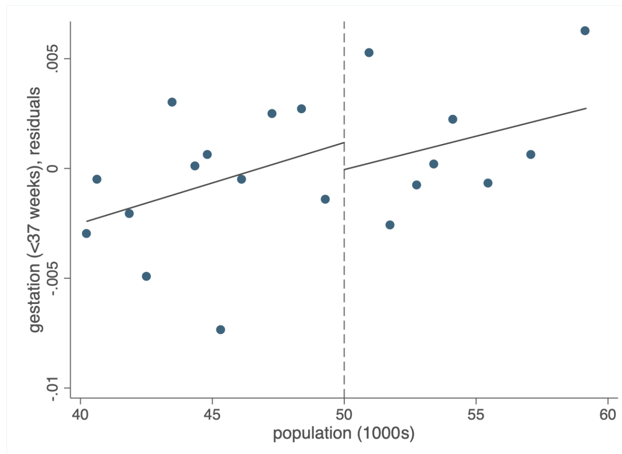
Note: The figure presents the regression discontinuity estimates of equation (1). It plots the residuals from a regression of the dependent variable on the controls and state and year fixed effects on different bins of population size and a linear fit of the relationship between the residuals and population at each side of the 50,000 inhabitants threshold. Each panel reports the results for a different dependent variable as indicated in the text.

Figure 8: Effects on Health at Birth

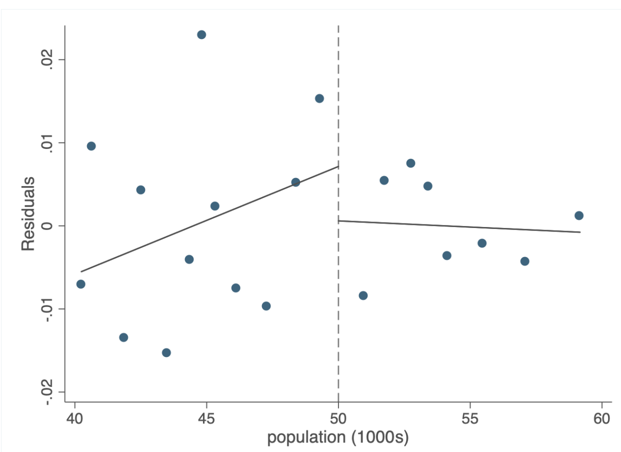
(a) $\leq$ 32 weeks of gestation (%)



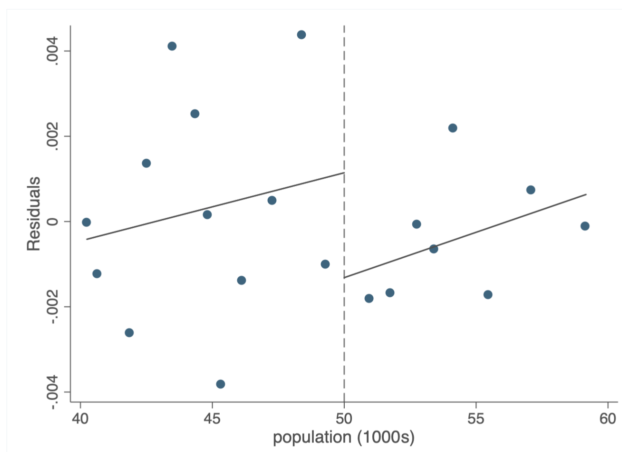
(b) $\leq$ 37 weeks of gestation (%)



(c) Low apgar 1 (%)



(d) Low apgar 5 (%)

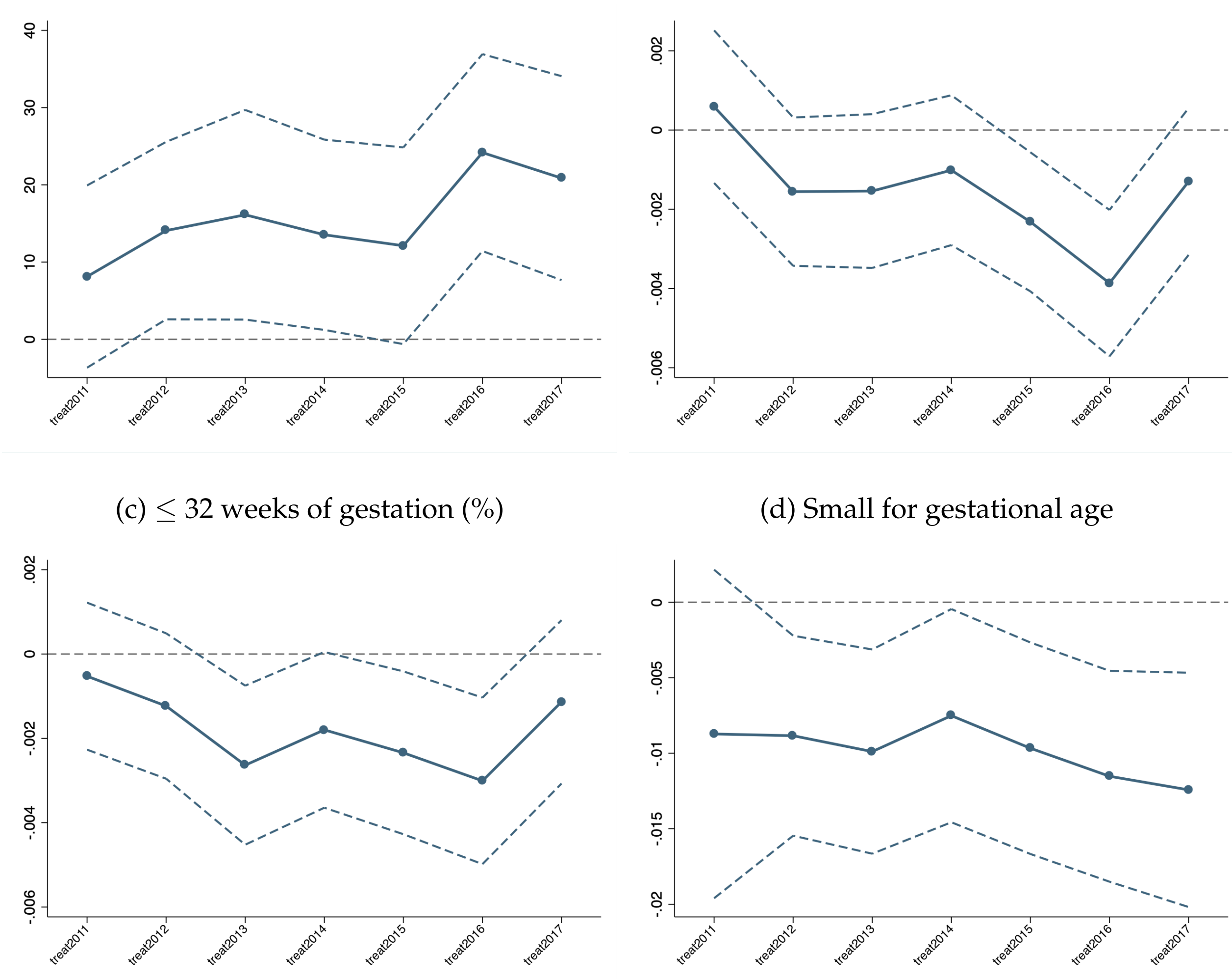


Note: The figure presents the regression discontinuity estimates of equation (1). It plots the residuals from a regression of the dependent variable on the controls and state and year fixed effects on different bins of population size and a linear fit of the relationship between the residuals and population at each side of the 50,000 inhabitants threshold. Each panel reports the results for a different dependent variable as indicated in the text.

Figure 9: Effects Over Time on Health at Birth

(a) Weight (g) (b) Weight $<$ 2000 (%)

(c) $\leq$ 32 weeks of gestation (%) (d) Small for gestational age



Note: The figure plots period-specific effects of the MCMV on indicators of health at birth (under 1 year) estimated using equation (2). The solid line reports the coefficients and the dashed line the 90% confidence interval. Panel A reports results for birth weight. Panel B reports the results for the share of births below 2,000 grams. Panel C reports the results for the share of gestations with less than 32 weeks. Panel D reports the results for small for gestational age.

Appendix to “The Effects of Better Houses on Infant Health”

A Decomposition of the Mechanisms

In this appendix, we explain in detail the procedure used to decompose the effects of the MCMV on birth weight on the effects of houses and their construction. Our decomposition exercise is inspired in the work of Dix-Carneiro et al. (2018). We assume the equilibrium relationship between the quality (and cost) of the housing stock, labor market conditions, and birth weight is constant over time and can be approximated using the following expression:

$$Y_{ist} = \beta_H H_{ist} + \beta_C C_{ist} + \gamma_t W_{is} + \eta_s + \epsilon_{ist}, \forall t \tag{A.1}$$

in which H_{is} denotes the quality of the housing stock in municipality i and state s, (proxied by the number of units of the MCMV built in the municipality), C_{is} the demand for labor in the construction sector in municipality i and state s (proxied by the the number of units of the MCMV under construction), $W_{is} = \{1, P_{is}, P_{is} \times T_{is}, Y^I_{is}, X_{is}\}$ is a vector of controls (constant, population, population interacted with dummy indicating whether the population is above the threshold, and initial municipality characteristics), η_s is a state fixed effect, and ε_{ist} an error term.

The quality of the housing stock and the labor market conditions are influenced by the rules of the MCMV. Specifically, we have:

$$H_{ist} = b^H_t T_{ist} + \gamma^H_t W_{is} + \eta_s + \epsilon^H_{ist}, \forall t \tag{A.2}$$

$$C_{ist} = b^C_t T_{ist} + \gamma^C_t W_{is} + \eta_s + \epsilon^C_{ist}, \forall t \tag{A.3}$$

in which $T_{is} = 1[P_{is} > 50,000]$.

Substituting equations (A.2) and (A.3) on equation (A.1), we obtain the following expression:

$$Y_{ist} = (\beta_H b_t^H + \beta_C b_t^C) T_{ist} + (\gamma_t + \beta^H \gamma_t^H + \beta^C \gamma_t^C) W_{is} + \eta_s + (\beta^H \epsilon_{ist}^H + \beta^C \epsilon_{ist}^C + \epsilon_{ist}), \forall t \tag{A.4}$$

Equation (A.4) shows that the RD coefficient of birth weight are the sum of the effects of houses and their construction weighted by the RD coefficients on houses and construction. Because this equation holds for all periods, it is possible to compute the effects of houses and their construction. To see this formally, define $\theta_t = \beta_H b_t^H + \beta_C b_t^C$ and suppose there are two periods (1 and 2). Then,

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \beta_H \begin{pmatrix} b_1^H \\ b_2^H \end{pmatrix} + \beta_C \begin{pmatrix} b_1^C \\ b_2^C \end{pmatrix} \tag{A.5}$$

Equation (A.5) expression demonstrates it is possible to obtain the coefficients β_H and β_C using the RD coefficients θ_1 and θ_2, b_1^H, b_2^H, b_1^C, and b_2^C and solving the system of linear equations it defines. The key hypothesis for this to be possible is that the coefficients β_H and β_C are stable over time. This might not be true, for instance, if the quality of the houses built changes over time.[19]

In our setting, we observe Y_{ist}, H_{ist}, C_{ist}, and T_{ist} for more than two periods. This implies we have an over-identified system with seven equations and two unknowns. However, to improve precision, we opt to perform the decomposition aggregating our data in two periods: initial years (2011-2014) and final years (2015-2017). The first period corresponds to the years in which the construction of houses was more intense but the changes

[19] Dix-Carneiro et al. (2018) have a system of two equations and five unknowns, implying they need to impose further restrictions and are just able to identify bounds on the parameters. They further show these bounds can be obtained using a procedure similar to a 2SLS.

in the housing stock were minor and the second period to the years in which the opposite occurs.

We estimate $b_1^H = 89.62$, $b_2^H = 289.92$, $b_1^C = 56.18$, $b_2^C = 23.86$, $\theta_1 = 12.97$, and $\theta_2 = 19.05$. Using these values to solve equation (A.5), we obtain $\beta_H = 0.054$ and $\beta_C = 0.145$. The parameters imply that improvements in housing quality and decreases in housing costs improve birth weight from 9.77-10.20 grams in the period 2011-2017. This corresponds to 60.4-66.5% of the mean effect of the MCMV on birth weight in this period. Improvements in labor market conditions due to the construction of the houses correspond to the rest 33.5-39.6% of the effect. The effect of houses increases over time as the changes in the housing stock become more important and construction activities end. The effect of houses is 4.81 grams in the first period and 15.58 grams in the second period. This corresponds to 37.2% and 81.8% of the total effects in these periods, respectively.

B Additional Tables and Figures

Table B1: Summary Statistics

	All			[40,000-60,000]		
	Mean	S.E	N	Mean	S.E	N
A. Demographics						
Population (2007)	33063	(197768)	5,565	48429	(5528)	235
Sh. young	0.347	(0.059)	5,565	0.355	(0.011)	235
Sh. old	0.121	(0.033)	5,565	0.105	(0.007)	235
Sh. rural hh	0.362	(0.220)	5,565	0.237	(0.057)	235
Sh. Migrants	0.106	(0.055)	5,565	0.0963	(0.013)	235
# households in deficit	1194	(7295.167)	5,565	1891	(159.910)	235
B. Labor and Schooling						
Sh. workers	0.612	(0.134)	5,565	0.624	(0.020)	235
Av. wage	808.5	(301.220)	5,565	899.6	(53.267)	235
less than 9 years	0.689	(0.090)	5,565	0.652	(0.021)	235
less than 9-11 years	0.132	(0.029)	5,565	0.142	(0.007)	235
less than 12-15 years	0.140	(0.050)	5,565	0.159	(0.013)	235
16 or more years	0.036	(0.0231)	5,565	0.0431	(0.005)	235
C. Infraestructure						
Sh. hh with water	0.645	(0.213)	5,565	0.688	(0.056)	235
Sh. hh with sewage	0.289	(0.312)	5,565	0.381	(0.053)	235
D. Health infraestructure						
# Hospitals	1.389	(6.995)	5,565	2.034	0.480	235
Presence of PSF	0.442	(0.497)	5,565	0.677	(0.145)	235
E. Outcomes						
Birth weight	3219	(90.473)	5,565	3219	(17.819)	235
Low birth ($<$ 2500)	0.067	(0.031)	5,565	0.0684	(0.005)	235
Apgar5	9.332	(0.370)	5,565	9.293	(0.117)	235
Total infant hosp. (up to 1 age)	192.492	(110.107)	5,565	205.6	(27.534)	235
infectious	33.97	(34.293)	5,565	38.40	(8.080)	235
respiratory	76.38	(66.71)	5,565	80.09	-15.368	235
perinatal	50.31	(41.90)	5,565	54.15	(9.005)	235
total infant death (up to 1 age)	(15.39)	14.43	5,565	15.36	(1.663)	235
infectious	0.77	(2.783)	5,565	0.778	(0.298)	235
respiratory	0.84	(3.396)	5,565	0.832	(0.352)	235
perinatal	9.068	(10.813)	5,565	9.163	(1.285)	235

Notes: The table presents mean values for municipality characteristics, measured in the baseline period. Population from 2007 comes from IBGE. The remaining variables from panels A-C come from the 2010 Population Census. The variable from panel D come from CNES/Datasus and the ones from Panel E come from SINASC, SIH, and SIM.

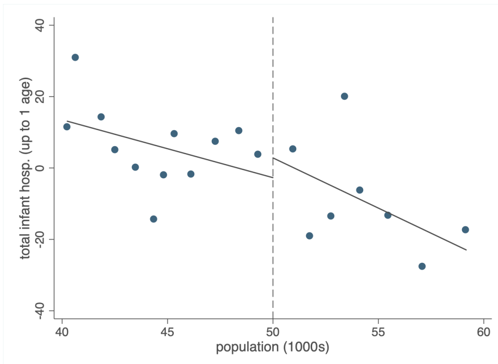
Table B2: **Effect on Weight at Birth controlling for percapita FPM**

	(1)	(2)	(3)	(4)	(5)	(6)
			Birth Weight			
1[pop >50,000]	14.47**	16.17**	14.80**	15.55**	14.29**	15.59**
	(6.52)	(6.40)	(6.11)	(6.55)	(6.20)	(6.08)
Mean	3214.34	3214.34	3214.34	3214.34	3217.34	3217.51
RD bandwidth	$\pm$10	$\pm$10	$\pm$10	$\pm$10	$\pm$11.10	$\pm$11.17
Kernel	Triangular	Triangular	Uniform	Uniform	Triangular	Triangular
Observations	1645	1645	1645	1645	1862	1883
Controls	FPM	All	FPM	All	FPM	All

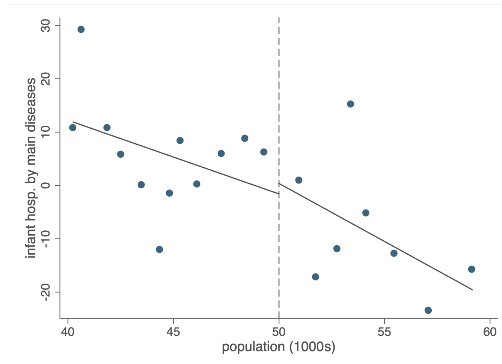
Notes: The table reports regression discontinuity estimates of the effects of the MCMV on measures of birth weight as dependent variable. Column 1 reports estimates obtained using a triangular kernel and a bandwidth of 10,000 inhabitants and including controls for the the weight birth at the baseline and control for per capita FPM. Column 2 adds the other controls for socioeconomic characteristics and health infraestructure. Column 3 reports estimates obtained using an uniform kernel and a bandwidth of 10,000 inhabitants and including controls for the the weight birth at the baseline and control for percapita FPM. Column 4 adds the other controls. Column 5 present the estimates obtained using a triangular kernel and the optimal bandwidth of Calonico et al. (2014) controlling for initial condition and per capita FPM, while column 6 adds all the controls. Standard errors clustered at the municipality level are reported in parenthesis. *** p<0.01; ** p<0.05; * p<0.10

Figure B1: Effects on infant hospitalization

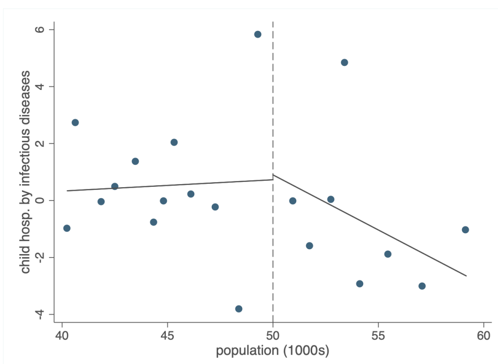
(a) Infant hosp. (overall)



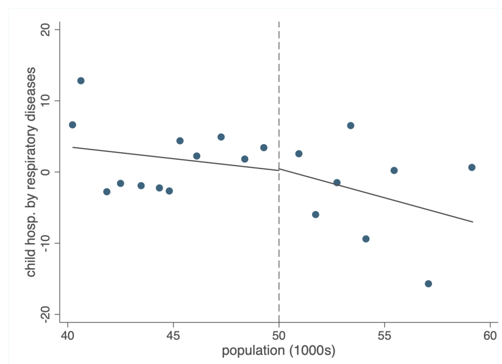
(b) Main causes



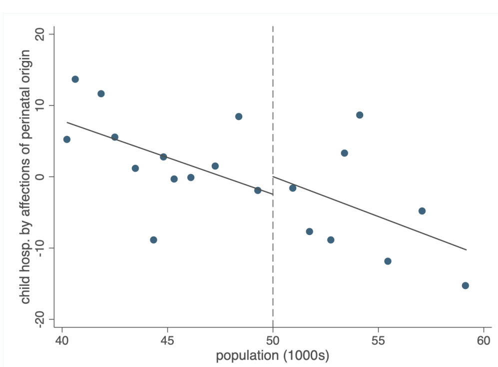
(c) Infectious diseases



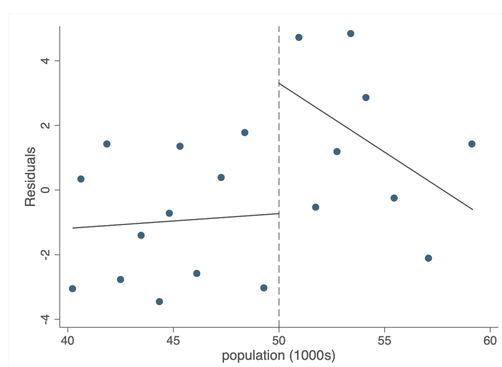
(d) Respiratory diseases



(e) Perinatal origin



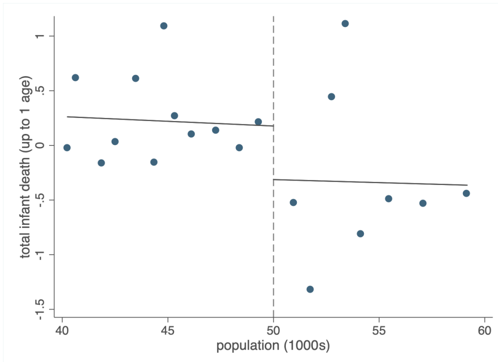
(f) Residual



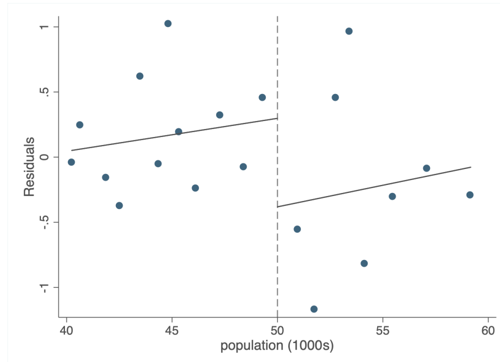
Note: The reports regression discontinuity estimates of equation (1) using measures of morbidity of children under 1 year as the outcomes of interest. We report results based on our preferred specification –triangular kernel, 10,000 inhabitants bandwidth, and the controls. Panel A reports total chil death, panel B aggregates the main causes considered (infectious and parasitic diseases, respiratory diseases, and perinatal conditions) and Panels C-D focus on each of these causes separately (infectious and parasitic diseases, respiratory diseases, and perinatal conditions, respectively). Panel E presents the estimates for the residual causes.

Figure B2: Effects on infant deaths

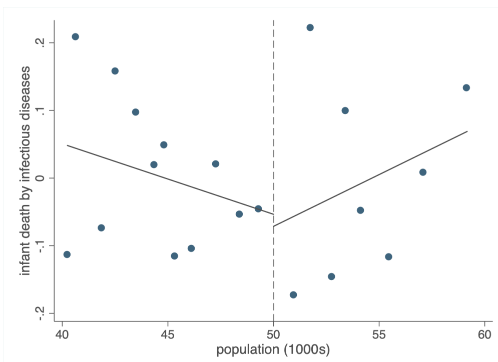
(a) Infant death (overall)



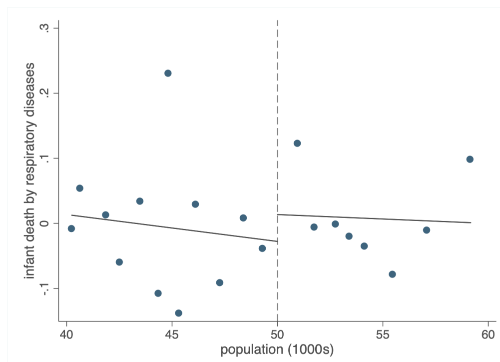
(b) Main causes



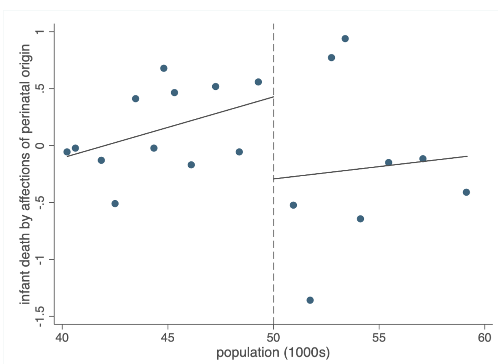
(c) Infectious diseases



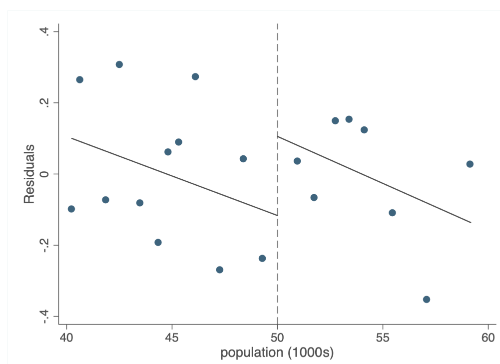
(d) Respiratory diseases



(e) Perinatal origin



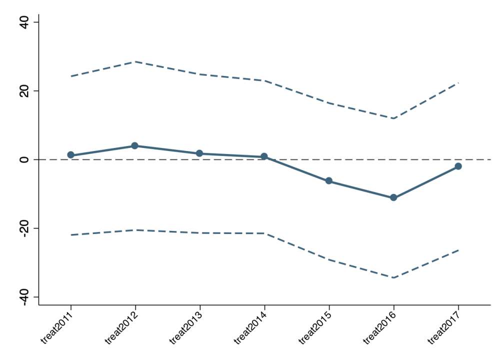
(f) Resisual



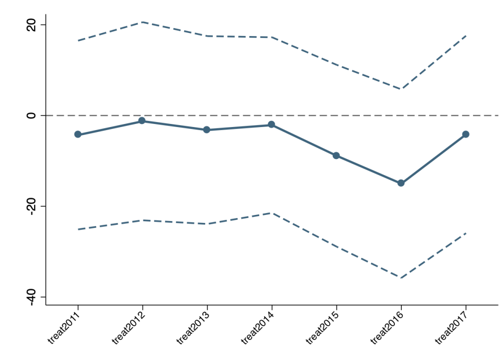
Note: The reports regression discontinuity estimates of equation (1) using measures of mortality of children under 1 year as the outcomes of interest. We report results based on our preferred specification –triangular kernel, 10,000 inhabitants bandwidth, and the controls. Panel A reports total chil death, panel B aggregates the main causes considered (infectious and parasitic diseases, respiratory diseases, and perinatal conditions) and Panels C-D focus on each of these causes separately (infectious and parasitic diseases, respiratory diseases, and perinatal conditions, respectively). Panel E presents the estimates for the residual causes.

Figure B3: Effects Over Time on Infant Hospitalization

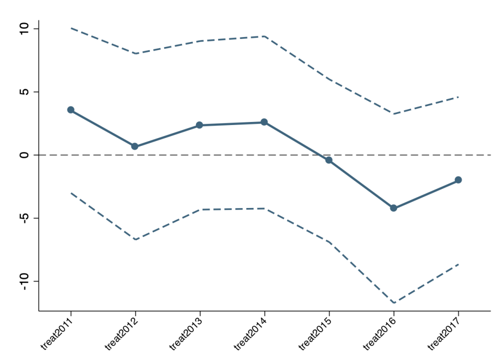
(a) Infant hosp. (overall)



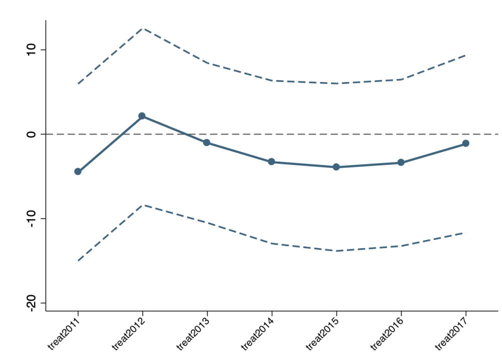
(b) Main causes



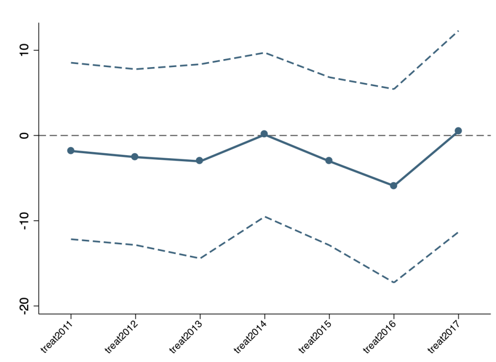
(c) Infectious diseases



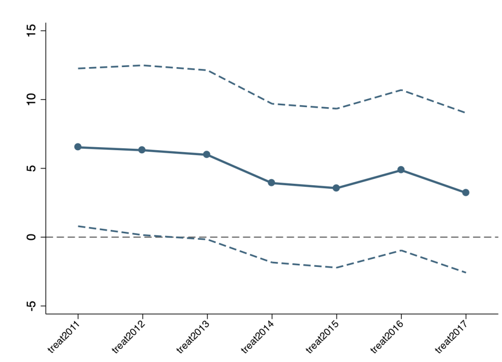
(d) Respiratory diseases



(e) Perinatal origin



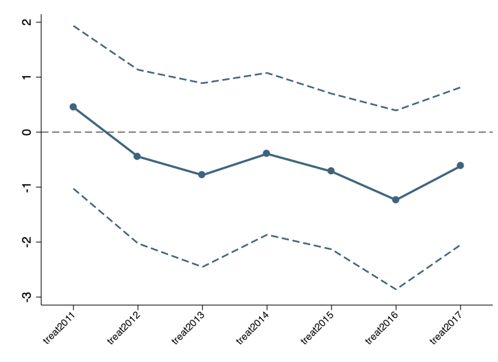
(f) Residual



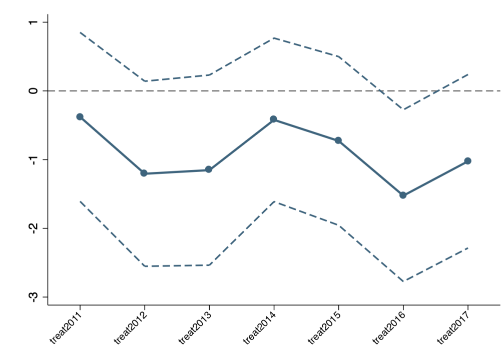
Note: The figure plots period-specific effects of the MCMV on infant hospitalization (under 1 year) estimated using equation (2). The solid line reports the coefficients and the dashed line the 90% confidence interval. Panel A reports results for hospitalizations in general. Panel B reports the results aggregating the main causes considered (infectious and parasitic diseases, respiratory diseases, and perinatal conditions). Panels C-E reports the results for each of these causes separately (infectious and parasitic diseases, respiratory diseases, and perinatal conditions, respectively). Panel F presents the estimates for the residual causes.

Figure B4: Effects Over Time on Infant Deaths

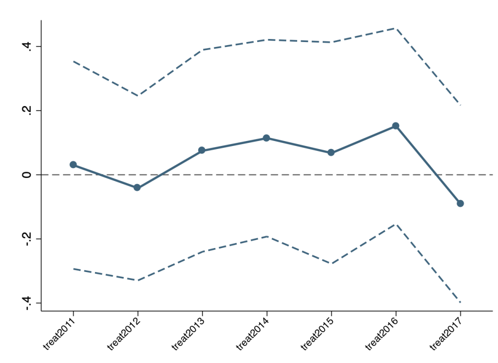
(a) Infant deaths (overall)



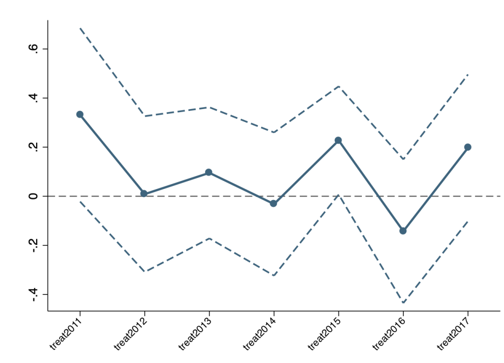
(b) Main causes



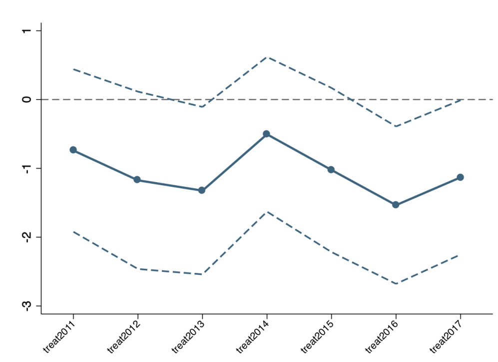
(c) Infectious diseases



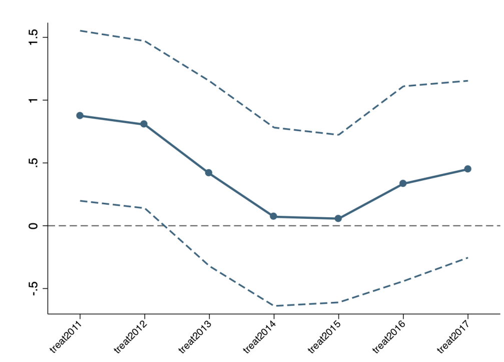
(d) Respiratory diseases



(e) Perinatal origin



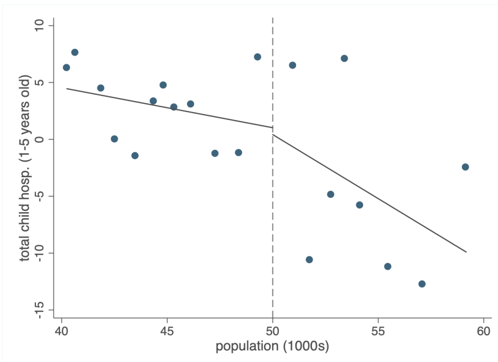
(f) Residual



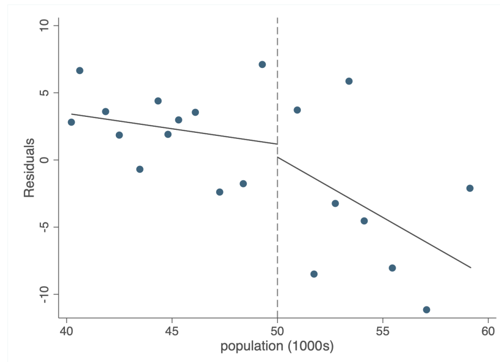
Note: The figure plots period-specific effects of the MCMV on infant mortality (under 1 year) estimated using equation (2). The solid line reports the coefficients and the dashed line the 90% confidence interval. Panel A reports results for hospitalizations in general. Panel B reports the results aggregating the main causes considered (infectious and parasitic diseases, respiratory diseases, and perinatal conditions). Panels C-E reports the results for each of these causes separately (infectious and parasitic diseases, respiratory diseases, and perinatal conditions, respectively). Panel F presents the estimates for the residual causes.

Figure B5: Effects on child hospitalization

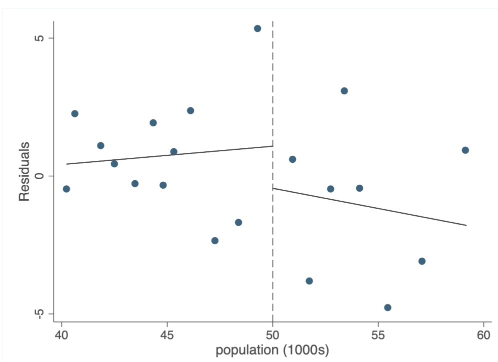
(a) Child hosp. (overall)



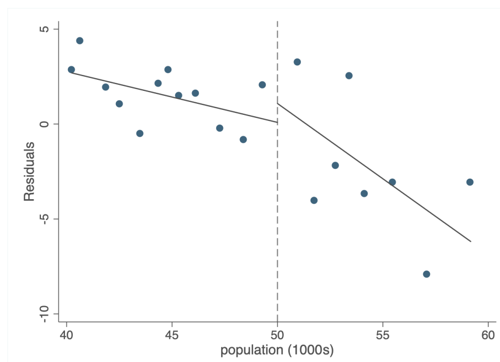
(b) Main causes



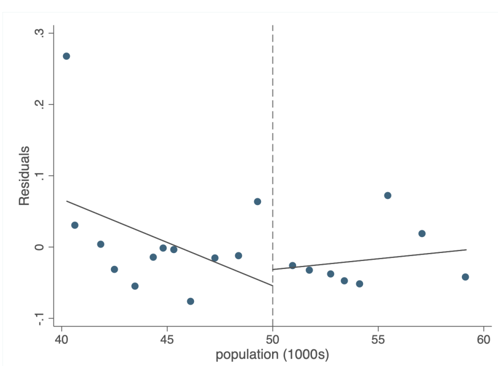
(c) Infectious diseases



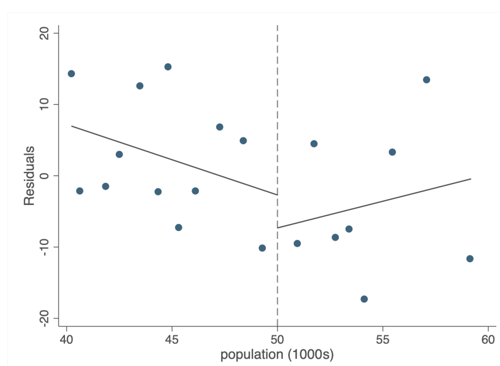
(d) Respiratory diseases



(e) Perinatal origin



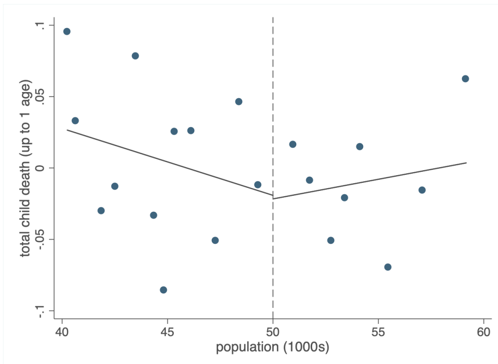
(f) Residual



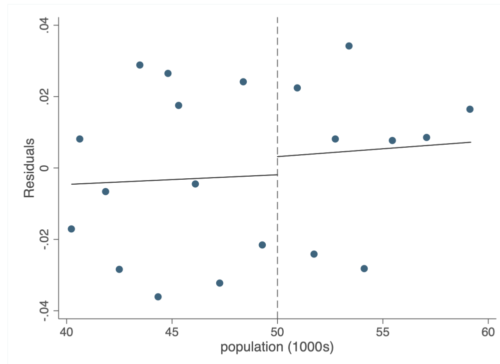
Note: The figure plots period-specific effects of the MCMV on infant hospitalizations (1 to 5 years) estimated using equation (2). Panel A reports results for hospitalizations in general. Panel B reports the results aggregating the main causes considered (infectious and parasitic diseases and respiratory diseases). Panels C-E reports the results for each of these causes separately. Panel F presents the estimates for the residual causes.

Figure B6: Effects on child deaths

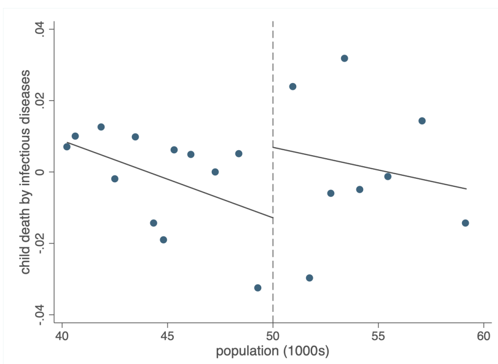
(a) Child death (overall)



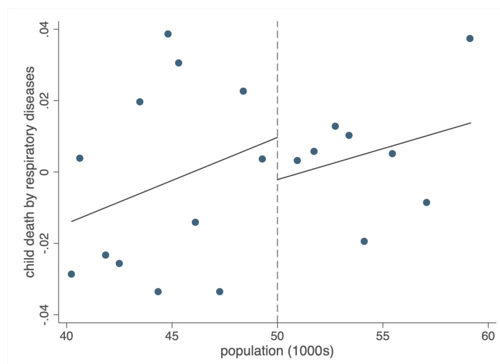
(b) Main causes



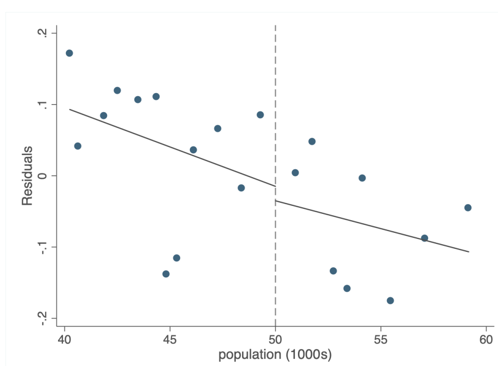
(c) Infectious diseases



(d) Respiratory diseases

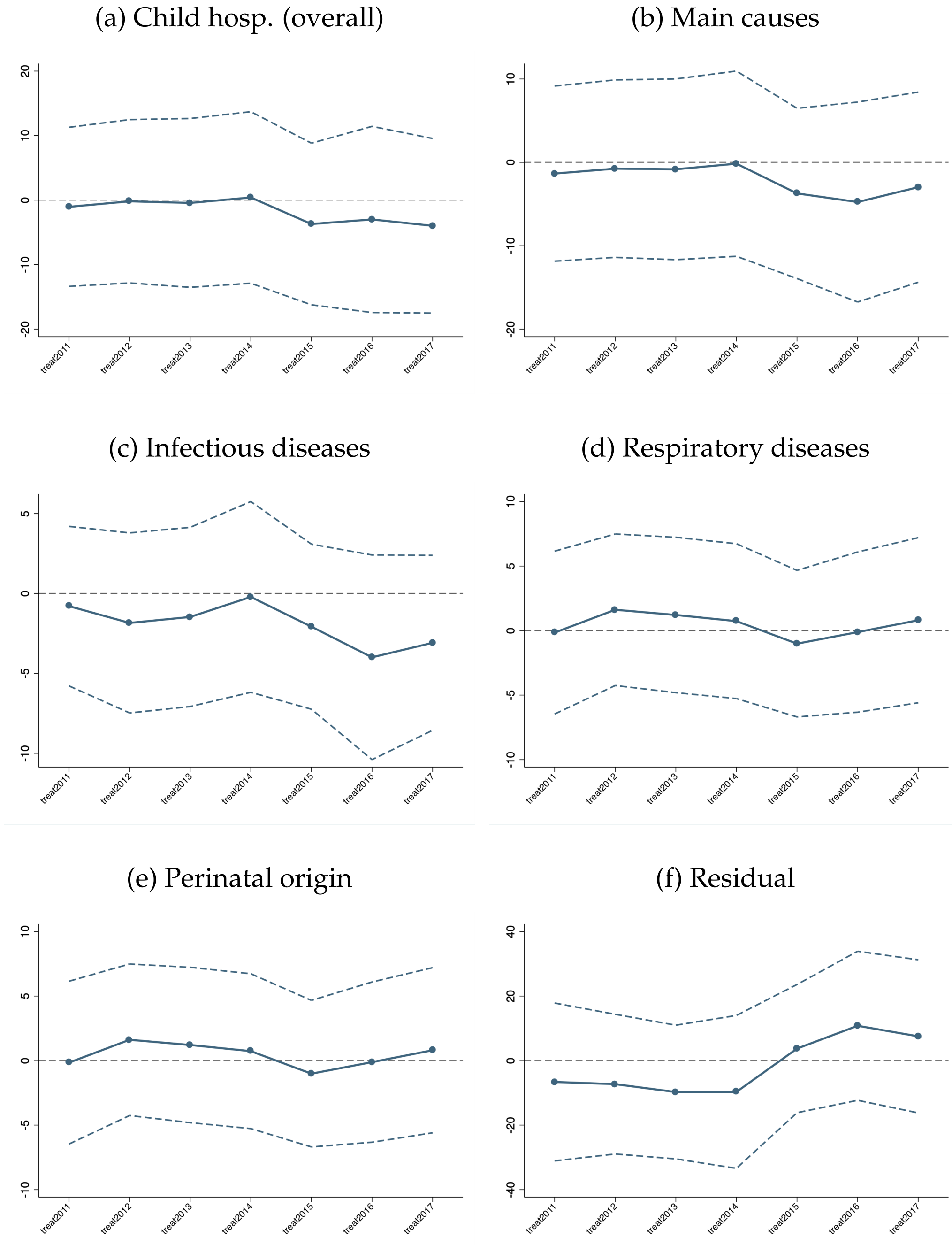


(e) Residual



Note: The figure plots period-specific effects of the MCMV on infant mortality (1 to 5 years) estimated using equation (2). Panel A reports results for hospitalizations in general. Panel B reports the results aggregating the main causes considered (infectious and parasitic diseases and respiratory diseases). Panels C-D reports the results for each of these causes separately. Panel E presents the estimates for the residual causes.

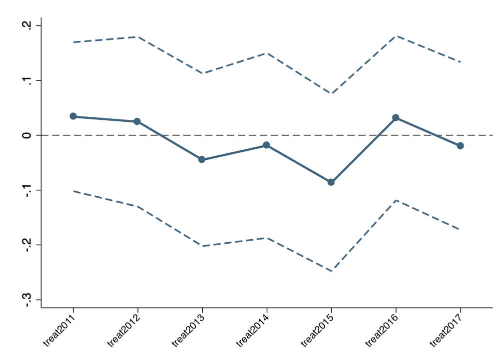
Figure B7: Effects Over time on Child Hospitalization



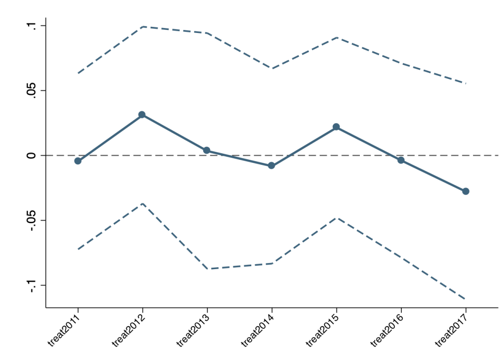
Note: The figure plots the estimated coefficients of equation (2) to obtain period-specific effects of the MCMV on hospital admission for children more than 1 and less than 5 years old. The solid line reports the coefficients and the dashed line the 90% confidence interval. Panel A aggregates the main causes considered (infectious and parasitic diseases, respiratory diseases, and perinatal conditions) and Panels C-E focus on each of these causes separately (infectious and parasitic diseases, respiratory diseases, and perinatal conditions, respectively). Panel F presents the estimates for the residual causes.

Figure B8: Effects Over Time on Child Deaths

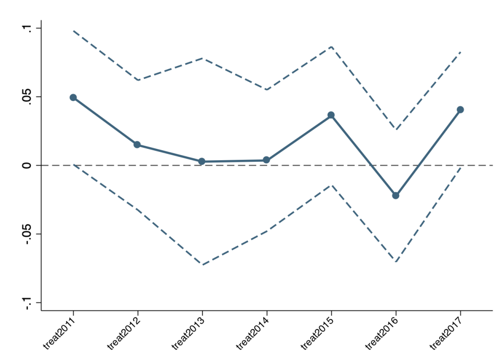
(a) Child death (overall)



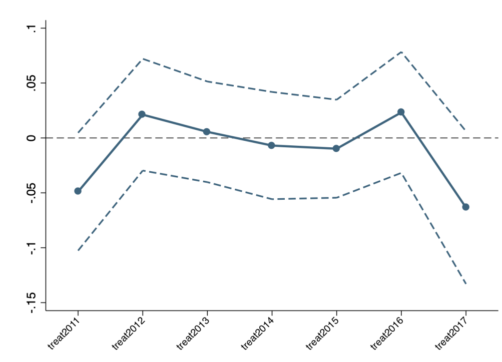
(b) Main causes



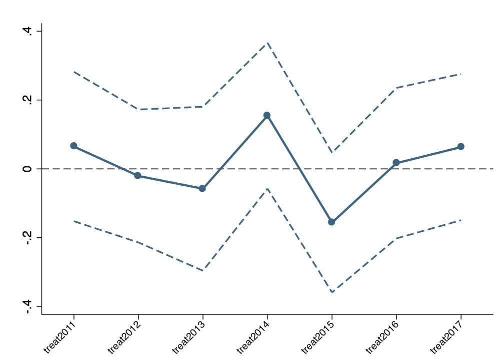
(c) Infectious diseases



(d) Respiratory diseases



(e) Residual



Note: The figure plots the estimated coefficients of equation (2) to obtain period-specific effects of the MCMV on child mortality rate. The solid line reports the coefficients and the dashed line the 90% confidence interval. Panel A aggregates the main causes considered (infectious and parasitic diseases and respiratory diseases) and Panels C-D focus on each of these causes separately. Panel E presents the estimates for the residual causes.

Non-Durable Consumption and Real-Estate Prices in Brazil: Panel-Data Analysis at the State Level*

Abstract

This paper investigates the effect of real-estate prices on non-durable consumption in Brazil. For that, we build a state-level panel of the determinants of non-durable consumption growth during the period 2008-2017. This period covers both a "boom" in real-estate prices and consumption (2008-2014) as well as a "bust" in them (2014-2017). We estimate the effect of house prices on consumption combining the techniques and ideas from Campbell & Cocco (2007) and Case et al. (2005). In particular, we estimated the same reduced-form equation proposed by Campbell & Cocco (2007), which is derived from simulating a theoretical model of housing and consumption choice under debt constraints. Due to Brazilian-data limitations, we were unable to run panel-data regressions at the cohort level (aggregation across households on different surveys) as in Campbell & Cocco (2007). Indeed, we had to resort to data aggregated at the state level to estimate our panel-data regressions as did Case et al. (2005). Our results suggest that changes in house prices significantly affect non-durable consumption in Brazil. The magnitudes are quantitatively close to the effects found for the U.K. by Campbell & Cocco (2007). Furthermore, we document that the effect of house prices on non-durable consumption is asymmetric, stronger in the "bust" than in the "boom" phase of the business cycle. This difference in the effects during different phases of the business cycle suggests that borrowing constraints might explain the effects of house prices on non-durable consumption.
JEL: *R30, E21, C23*
Keywords: *Real Estate, Non-durable Consumption; Wealth; Panel Data*

*This paper is co-authored with Victor Pina Dias (BNDES), Érica Diniz (INSPER), and João Victor Issler (FGV EPGE). The published version can be found at `http://bibliotecadigital.fgv.br/ojs/index.php/rbe/article/view/42554`. We are grateful to the editor, Carlos da Costa (FGV EPGE), and one anonymous referee for their comments and suggestions. We thank INCT, FAPERJ, CNPq for financial support. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

1 Introduction

Housing is a very important component of wealth of a household, especially when we consider the middle-class of income for any society. In the U.S., there is research indicating that a significant portion of wealth of a family is allocated to buy real estate. Bertaut & Starr-McCluer (2002) show that, in the late 1990's, residential property corresponded to about one quarter of aggregate wealth of a family living in the U.S. The official statistics (U.S. Census Bureau, 2012) show that this proportion has remained roughly stable through time, despite the recent effects of the global recession: in 2010, residential structures corresponded to 24.8% of household's net worth.

The fact that the global recession had its roots on the U.S. housing market collapse had spurred a number of studies trying to understand the links between housing prices and household welfare, or, similarly, between housing prices and household consumption; see, *inter alia*, Gan (2010), Luengo-Prado et al. (2009), and Ren & Yuan (2014). Even before the real estate market collapse, some authors recognized the importance of this issue, e.g., Case et al. (2005) with data from the U.S. and other developed countries and Campbell & Cocco (2007) with data from the U.K. Most of these studies resorted to household data to investigate the links between the housing market and consumption.

Unfortunately, in Brazil, our best household survey – *PNAD – Pesquisa Nacional por Amostra de Domicílios* – is very incomplete regarding wealth data and has no data on consumption. Perhaps this is a consequence of the fact that income inequality has dominated the welfare debate in Brazil, but one can only conjecture why our most prominent survey has neglected consumption and welfare statistics.

Previous studies have shown that real estate also represents an important portion of household wealth in Brazil, with obvious consequences to welfare. For example, Marquetti (2009) estimates wealth in Brazil between 1950 and 1998 using the perpetual inventory method and finds that residential structures amount to about a third of the net stock of fixed capital. Moreover, its average annual growth was 8.7% between 1981 and 1998. Hofman (1992) estimates the capital stock for six Latin American countries (including Brazil) between 1950 to 1989, finding that residential construction represents more than 20% of the net capital stock. Table 1 summarizes these findings. Finally, Morandi (1998) estimates that household real estate as a proportion of gross private wealth has remained roughly constant $(1/3)$ between 1970 and 1995. Compared to the importance of real estate to *net wealth* in the U.S. $(1/4)$, the results for Brazil are striking and point to the importance of the real-estate market for welfare in Brazil.

The goal of this paper is to investigate the effect of real-estate price variation on non-durable consumption and welfare, trying to close a gap between the consumption literature in Brazil and compare it with the U.S., the U.K., and other developed countries. Our first motivation relates to the fact that real estate is probably more important here than elsewhere as a proportion of wealth, which potentially makes the impact of a price change here bigger. Our second motivation is the recent boom of the real-estate prices in Brazil during several years (2008-2014), followed by a bust in these same prices (from 2015 onward). During the boom, prime real estate in Rio de Janeiro and São Paulo have tripled in value, and a somewhat smaller but generalized increase has been observed throughout the country. These changes are unusual, since the last major real-estate price boom in Brazil occurred in the late 1960's and early 1970's. Third, we have also seen a consumption boom in Brazil from 2008 to 2014 and a bust from 2015 onward.

Because our goal is to investigate the relationship between fluctuations of house prices and consumption (welfare) in Brazil, we follow the well-known studies of Case et al. (2005) and Campbell & Cocco (2007). Case et al. (2005) use panel data for 14 developed countries between the late 1970's and 1990's and find a strong correlation between house prices and the aggregate consumption of households. They also repeat this exercise using U.S. state data with similar results. Campbell & Cocco (2007) investigate the response of household non-durable consumption to house price changes using micro panel data for the U.K. They estimate the price elasticity of consumption for different cohorts, finding a positive responses of household consumption to an increase in house prices. This effect is bigger for older cohorts, and not significant for younger renters, showing a heterogeneous effect across groups.

The interesting feature of Campbell & Cocco (2007) is that they used a structural equilibrium model to understand how these fluctuations in house prices affected households' consumption decisions, identifying important channels that could explain changes in the latter. With simulated data from the theoretical model calibrated to the U.K. economy, they fit a reduced-form regression of changes in consumption on changes in housing prices, income changes, real-interest rates and additional controls, finding a positive marginal effect of changes in housing prices on changes in consumption after controlling for additional important variables such as demographics, real interest rates, income, loan conditions, etc. Moreover, their approach allows quantifying these effects.

Regarding the important channels of transmission from housing prices to consumption, they first conjecture that a possible reason for the existence of a positive correlation is a wealth effect: increasing real estate prices increases the perceived value of household wealth for home owners. However, they recall that housing is a commodity and its higher price is simply a compensation for higher implicit cost of housing – its imputed rent. So, if we rule out any substitution effect from housing services to non-durable consumption, the increase in the price of real estate must be exactly offset by the expected present discounted value of rent. Hence, in expected present value terms, there is no change in the budget constraint for the household, leaving non-durable consumption unchanged.

Campbell & Cocco (2007) also mention that rising house prices may stimulate consumption by relaxing borrowing constraints. This happens because a house is an asset that can be used as a collateral in a loan. Thus, an increase in house prices could increase consumption not by a direct wealth effect, but because a consumer may then increase borrowing to smooth consumption over the life cycle once the price of the house has increased – re-financing, for example. They also argue that this effect is heterogeneous: young renters are "short" on housing (want to buy) whereas old owners are "long," since they want to move from a larger house to a smaller one. This idea is also present in Lustig & Van Nieuwerburgh (2010).

There are other papers that investigate optimal durable versus non-durable consumption decisions with obvious relevance to the issue we want to address here; see, for instance, Bernanke (1985), Ogaki & Reinhart (1998), and Yogo (2006). Usually, they have a representative consumer who derives utility from consumption of non-durables and from the services provided by the current stock of durable goods. Given that real estate is a major component of these services, they provide an integrated framework to deal with this issue.

Campbell & Cocco (2007) go one step beyond this literature, trying to address what reduced-form equation one should expect from this basic theoretical setup, quantifying the marginal effect of a change in prices in non-durable consumption. Moreover, their simulations confirmed the empirical findings of the elasticities found in reduced-form estimation. This offers an useful guideline for investigating whether fluctuations of house prices affect consumption in Brazil, being the reason why we chose to follow their theoretical and empirical implementation. Hence, this paper will use this reduced form equation as our benchmark equation to estimate the correlation between house prices and consumption.

Although we follow Campbell & Cocco (2007), there are some limitations in our study arising from the lack of identical micro data in Brazil and in the U.K. As we stressed above, PNAD does not have consumption data for households.[1] Thus, we had to resort

[1] Another Brazilian survey, POF – *Pesquisa de Orçamentos Familiares*, has household consumption data, but it is not collected in every year, but every 7 or 8 years apart. Older POF surveys have a specific serious problem due to high inflation, in which all price data is collected in nominal terms but inflation prior to 1995 has reached up to 80% a *month* in some cases.

to state-level data on consumption. Indeed, Brazil has an index of monthly consumption in another survey, PMC – *Pesquisa Mensal de Comércio*, from February 2008 through December 2017, for all Brazilian States. In Particular, we are intereseted in the states of Rio de Janeiro, São Paulo, Minas Gerais, Bahia, Pernambuco, Ceará, Distrito Federal, Espírito Santo, Goiás, Paraná and Rio Grande do Sul. With that in hand, we also obtained real-estate price data from FipeZap on the capital of the states mentioned above [2]. Thus, we were able to find Brazilian data for the dependent variable and the main regressor in Campbell & Cocco (2007)'s reduced-form regression. We were also able to find data on other control variables used in their study as well.

Our cross-sectional units are represented by Brazilian states. On that dimension, our setup gets closer to that of Case et al. (2005) than to Campbell & Cocco (2007), although we will use the same reduced-form equation that Campbell & Cocco (2007) estimate in their paper. In adapting the latter framework to state cross-sectional units, we need to employ state-level demographic controls.

One interesting aspect of the behaviour of the recent Brazilian house-pricing boom of 2009-2014 is how wide it has been, both geographically and across different real-estate units. This point can be illustrated by comparing the monthly growth rate of nominal house prices in Brazil (Figure 1, panel A) and in the two largest cities in Brazil: São Paulo and Rio de Janeiro (Figure 1, panel B).

First, in the boom period, the increase in monthly prices reaches more than 2.5% in some months and nowhere we observe an actual decrease in the level of real-estate prices. In Rio de Janeiro it reaches more than 3%. Second, it seems that price increases follow a similar cyclical pattern across cities[3]

There are several factors that could explain this sharp increase in real-estate prices in

[2] Although PMC is available for all Brazilian states, Fipzap is available for selected cities so we restricted the analysis to the data availability of FipZap data

[3] Appendix A presents the evolution of the house prices for each state considered in this study.

Brazil from 2008 to 2014. The first is the decrease in real interest rates. The Brazilian basic interest rate (Selic) was set as 17.25% per annum by the Central Bank of Brazil in early 2006 and had decreased to 8% per annum in the middle of 2012, reaching 7.5% in 2017. As a consequence, we observed a sharp increase in real-estate credit for this period. The second is an increase in the purchasing power of the Brazilian middle class: minimum wage has increased above inflation in the recent past and the Brazilian government adopted a myriad of social programs, all of which transferred income to poor and middle-income families. Third, government revenues, private-firm and individual income all increased due the global commodity-price boom experienced until 2015.

On the other hand, the bust in real-estate prices observed recently has its roots on Brazil's worst recession ever, where real GDP decreased by 8.9% from 2014.1 through 2016.12, with a very mild recovery from then on. Real-Estate prices in Rio de Janeiro were hit hardest.

Our empirical results are as follows. First, as in Campbell & Cocco (2007) we find a positive effect of house-price growth on non-durable consumption growth in Brazil after appropriate controls are accounted for. Second, this effect is quantitatively similar to the one found in the U.K. by the latter. In Brazil, house-price elasticity estimates are in the range 0.28 to 1.58, depending on whether we employ regional or national house price variation as an explanatory variable. These elasticities are comparable to the ones found by Campbell & Cocco (2007), which range from 0.57 to 1.59. Finally, going beyond Campbell & Cocco (2007), we document an asymmetric effect of house prices on non-durable consumption, which is stronger in the "bust" than in the "boom" phase of the business cycle.

The remainder of the paper is organized as follows. Section 2 describes the model and the data considered. Section 3 presents the estimation methodology and the results. Finally, Section 4 concludes the paper.

2 The Model and the Data

2.1 Model

We motivate the empirical investigation using the theoretical model of housing and consumption choice introduced by Campbell & Cocco (2007). These authors find qualitatively identical relationships between the growth in non-durable consumption, house prices, interest rates, and income using both real data and data generated from a calibrated version of this model. This implies that this model provides a useful benchmark to investigate the relationship between house prices and non-durable consumption.

The theoretical model considers that households (indexed by i) derive utility during in each period (indexed by t) from housing services, H_{it}, and non-durable goods, C_{it}. It assumes households have time additive preferences that are separable between housing and non-durable goods consumption:

$$u(C_{it}, H_{it}) = \frac{C_{it}^{1-\gamma}}{1-\gamma} + \theta \frac{H_{it}^{1-\gamma}}{1-\gamma}. \tag{1}$$

Separability in preferences eliminates possible substitution effects coming from increases in the price of housing services. This is an important feature of this setup.

In each period, the household decides not only on H_{it} and C_{it}, but also if it is optimal to rent or to buy real estate. Let small-cap letters denote variables in logs. (Logged) real labor income is given by:

$$y_{it} = f(t, Z_{it}) + v_{it} + w_{it}, \tag{2}$$

where $f(t, Z_{it})$ is a function of time (also interpreted as age) and household characteristics Z_{it}. The components v_{it} and w_{it} are two stochastic components. One is transitory and the other persistent. The transitory component is captured by the shock w_{it} – i.i.d., normal, with mean zero and variance σ_w^2. The persistent component follows a random walk: $v_{it} = v_{it-1} + \eta_{it}$, where η_{it} is i.i.d., normal, with mean zero and variance σ_η^2.

The model assumes that house prices fluctuate over time. The real house price growth rate is given by:

$$\Delta p_{it} = g + \delta_{it}, \tag{3}$$

where g is a constant and δ_{it} is a zero-mean normally distributed shock.

On the financial side, Campbell & Cocco (2007) assume that "there is a single financial asset with risk-free interest rate R_t, in which households may save. Homeowners may also borrow at this rate, up to the current value of the house minus a down payment." Thus, households face a borrowing constraint given by:

$$D_{it} \leq (1 - d) P_{it} H_{it}, \tag{4}$$

where D_{it} is household's outstanding debt, d is the down payment proportion and P_{it} is the house price.

It is important to note that, if house prices go up (down), this relaxes (tightens) the borrowing constraint of the household, allowing consumption to increase (decrease) beyond what we would normally have under no price increase (decrease). This leads to a positive *partial* correlation between non-durable consumption and house prices, a channel that could be identified by estimating a reduced form as shown below.

The authors allow home owners to borrow against the value of their house at the risk free rate. Because of this they also rule out default:

$$D_{it} (1 + R) \leq (1 - \lambda) \underline{P_{it+1}} H_{it} + \underline{Y_{it+1}}, \tag{5}$$

where $\underline{P_{it+1}}$ and $\underline{Y_{it+1}}$ are the lower bounds in house prices and labor income in period $t+1$, respectively, and λ represents transaction costs in buying and selling houses.

Campbell & Cocco (2007) solve the model with parameters calibrated to represent the U.K. economy at the household level. Then, using data generated by the model, they estimate a reduced-form regression where the change in consumption is the dependent variable explained by changes in housing prices (assumed strictly exogenous in the model), interest rates, changes in income, loan conditions, and additional demographic controls. Estimation results show a positive relationships between changes in non-durable consumption and interest rates, changes in income and house prices in their pseudo-panel for different cohorts. The authors explore the results of this reduced-form estimation in their paper, where the same estimation using actual data generated qualitatively similar results.

Our main assumption here is that it is possible to analyze the effects of house prices on non-durable consumption exploring state-level data using the same reduced-form approach backed up by their structural model. Indeed, we are performing the same type of aggregation they perform, but on a larger scale: We aggregate consumption and other variables at the state level data while they aggregate by cohort. Their estimated reduced-form is the following:

$$\Delta c_{s,t} = \beta_0 + \beta_1 r_t + \beta_2 \Delta y_{s,t} + \beta_3 \Delta p_{s,t} + \beta_4 \Delta m_{s,t} + \beta_5 \Delta Z_{s,t} + \epsilon_{s,t}, \tag{6}$$

in which $(\Delta c_{s,t})$ is the growth rate of non-durable consumption goods in state s and period t, r_t is the interest rate between periods t and $t-1$, $(\Delta y_{s,t})$ is the real growth rate of income in state s and period t, $(\Delta p_{s,t})$ is the real growth rate of house prices in state s and period t, $\Delta m_{s,t}$ is the growth rate in mortgage payments in state s and period t, $Z_{s,t}$ is a vector of demographic controls, and $\epsilon_{s,t}$ is a stochastic term. The theoretical model suggests the following expected signs for the regression coefficients:

1. $\beta_1 > 0$: since there the standard positive inter-temporal substitution effect for non-durable consumption.

2. $\beta_2 > 0$: since there is a positive effect of income innovations on non-durable consumption.

3. $\beta_3 > 0$: since there is a positive effect of house prices on non-durable consumption coming from the fact that an increase in house prices will relax the borrowing constraint of the agent.

2.2 Data

We explore state-level consumption data which is available from PMC – *Pesquisa Mensal do Comércio*. This is a monthly dataset collected by *Instituto Brasileiro de Geografia e Estatística*. It is the best source of high frequency consumption data in Brazil since the household-level consumption data from the POF – *Pesquisa de Orçamentos Familiares* – is only collected at 7- or 8-year intervals.

A monthly index of disaggregated consumption was obtained for the period February 2008 to December 2017 (119 months). From it, we constructed the growth rate of total non-durable consumption for the states of Rio de Janeiro (RJ), São Paulo (SP), Minas Gerais (MG), Bahia (BA), Pernambuco (PE), Ceará (CE), Distrito Federal (DF), Espírito Santo (ES), Goiás (GO), Paraná (PR) e Rio Grande do Sul (RS) – a total of 11 states. For every state, we defined total non-durable consumption as the sum of the following consumer-good categories (weights in parenthesis): fuels and lubricants (8%), hypermarkets, supermarkets, food products, beverages and tobacco (65%); clothing and shoes (10%), pharmaceutical articles, medical, orthopedic, perfumery and cosmetics (12%); books, newspapers, magazines and stationery (2%); and other personal articles and of domestic use (3%).[4]

[4]PMC series which we did not consider fell on the following categories: hypermarkets (other), furniture

These weights were obtained from the POF survey of 2008-2009. From these weights and the growth rates of the indexes in each category, we are able to compute the monthly growth rate of total non-durable consumption in every state – the dependent variable $(\Delta c_{s,t})$ in equation (6).

The explanatory variables in equation (6) were obtained from various sources. The risk-free interest rate considered here is Selic, the basic interest rate on loans from the Central Bank of Brazil to the financial sector.[5]. Selic was used as follows: $r_t = \ln(1 + R_t)$, where R_t is Selic in real terms – deflated using the National Consumer Price Index (IPCA).

State income growth rates $(\Delta y_{s,t})$ used the regional data from IBC-Br the Regional Economic Activity Index – constructed by the Central Bank of Brazil. The only state in our sample for which IBC-Br is not available is Distrito Federal (DF). We used as a proxy the income growth rate for the Midwest region as a whole which includes Distrito Federal (DF). An alternative series for $(\Delta y_{s,t})$ was constructed using the wages and employment in the formal sector available in RAIS (Relação Anual de Informações Sociais).

The growth rate in house prices $(\Delta p_{s,t})$ was computed using FipeZap. In particular, we used the growth rates of the "Índice FipeZap de Preços de Imóveis Anunciados". It does not contain actual transaction prices, but list prices on advertised real-estate properties. As is well known, list prices are a good proxy for transaction prices.

Data are available for the cities of Rio de Janeiro (state of RJ), São Paulo (state of SP), Belo Horizonte (state of MG), Salvador (state of BA), Recife (state of PE), Fortaleza (state of CE), Brasilia (Distrito Federal - DF), Vitória (state of ES), Goiania (state of Goiás) and Porto Alegre (state of RS). Here, we were forced to use real estate price data for the state capital in each state, since state-wide data were not available. We should note that São Paulo and Rio de Janeiro have a longer time span on real-estate price data (starts in February 2008)

---

and household appliances, office equipment and supplies, computer and communication.

[5]The Interbank Certificate of Deposit rate (CDI) was also used as a robustness check. The results (not shown) are very similar

vis-à-vis other state capitals (data from 2009 or 2010). Thus, we have an unbalanced panel. Table 2 shows the sample size available for each of them. There is also a national index of real-estate prices but it is only available from 2010 onward.

There is no direct information on the growth rate of mortgage payments ($\Delta m_{s,t}$). Thus, we used proxies that control for indebtedness of Brazilian families: default rate for loans in the financial system, household debts and demand deposits. All these variables are collected from the Central Bank of Brazil at the state level.

The vector of control variables ($\Delta Z_{s,t}$) encompasses a myriad of different series: employment in the formal sector (from RAIS) and share of people in the working age in each state (from PNAD). Campbell & Cocco (2007) highlight the importance of demographic variables for the response of consumption to house prices. We build the working age variable interpolating the quarterly data from PNAD to create the share os people between 18 and 64 years old in each state. This quarterly data is available for 2012 onwards. Following Campbell & Cocco (2007), seasonal growth-rate dummies are also included in $\Delta Z_{i,t}$, since consumption growth has a clear seasonal pattern.

All nominal series were deflated by the Broad National Consumer Price Index (IPCA). For robustness sake, the same exercise was done with the National Consumer Price Index (INPC), but the results are almost identical [6].

Table 3 shows descriptive statistics for the main variables in this paper. $\Delta c_{s,t}$ is the average non-durable consumption growth per month (logged differences); $\Delta p_{s,t}$ the real monthly (log) changes in house prices, Δpnac_t is the real growth in house prices of the national index, and $\text{Diff}p_{s,t} = \Delta p_{s,t} - \Delta \text{pnac}_t$, deviations from national prices growth rates; r interest rate, $r_t = \ln(1 + R_t)$, where R_t is the Selic rate in real terms (deflated using IPCA); $\Delta \text{Inad}_{s,t}$ is the default rate of credit operations of the National Financial System; $g(\text{Wage})_{s,t}$ is the real wage growth rate in the formal sector of the economy; $\Delta y_{s,t}$ is the

[6] Available under request

Regional Economic Activity Index constructed by the Central Bank (IBC-Br); $\Delta\text{Loans}_{s,t}$ refers to the growth in loans and discounted securities; $\Delta\text{Depvista}_{s,t}$ measures the growth of demand deposits; $\Delta\text{Ocup}_{s,t}$ is the growth in the share of employment in the formal sector; and $\text{WorkingAge}_{s,t}$ is the share of people aged 25-64 years old.

As shown in Table 3, the average consumption growth per month, $\Delta c_{s,t}$, is 1.1% per month. The house price growth rate remains around 0.6% per month – higher than that of IPCA – which average monthly growth rate was 0.49%.

It is important to stress that Campbell & Cocco (2007) also had to aggregate household data forming a synthetic panel, where *synthetic* individuals, aggregated across cohorts, were followed through time. State-level aggregation, although similar in spirit, is done on a much larger scale considering the number of households in each state. Both techniques rely on the law-of-large numbers to clean up idiosyncratic measurement errors at the household level.

3 Results

3.1 Fixed Effects

To estimate equation (6), we impose the following structure for the error term:

$$\epsilon_{s,t} = a_s + u_{s,t},$$

where a_s is the fixed effect (constant across time periods) and $u_{s,t}$ is an idiosyncratic error term. We allow $u_{s,t}$ to be dependent across time and cross-sectional units. This requires for proper inference using some type of robust correction in constructing estimates for the standard errors.

Table 4 presents estimation results of equation (6) in five different specifications. The dependent variable is $\Delta c_{s,t}$ and the main explanatory variable of interest is $\Delta p_{s,t}$. Column 1 controls for the interest rate. Column 2 adds proxies of indebtedness. Column 3 adds control by the IBC-Br index $\Delta y_{s,t}$. Column 4 replaces the control $\Delta y_{s,t}$ by employment rates in the state,$\Delta Ocup_{s,t}$, and $\Delta Wage_{s,t}$ to check whether the results are sensitive to different measures of fluctuations in economic activity. Column 5 includes the share of people aged 25-64 to control for demographic changes within the states. The standard errors are clustered at the state-level, thereby allowing for unrestricted residual correlation within states. All columns include state-fixed effects and controls for seasonality. We impose *strict exogeneity* of the regressors conditional on the unobserved effect a_i. Thus, estimation of the marginal effects is performed using the so called *fixed-effects estimator*.

In line with the theoretical prediction of the simulations in Campbell & Cocco (2007), the results show that changes in the growth rate of house prices ($\Delta p_{s,t}$) are positively correlated with changes in growth rates of non-durable consumption ($\Delta c_{s,t}$) using proper controls. This effect is economically and statistically significant. One percent increase in house prices is associated with an increase in 0.285-0.413 percent in non-durable consumption. The effect is robust across specifications. The evidence further indicates that Δr_t positively influences non-durable consumption. This indicates there is a standard inter-temporal substitution effect operating for non-durable goods consumption. Income growth $\Delta \text{Wage}_{s,t}$ and $\Delta y_{s,t}$ are also positively correlated with consumption as predicted by the theoretical model. However, the effect is statistically significant only for $\Delta y_{s,t}$.

Table 5 analyzes whether the relationship between house prices and non-durable consumption is driven by national or regional trends. Odd columns repeat the specification from Table 4, column 4 using the real growth in national house price Δpnac_t and $\text{Diff}p_{s,t} = \Delta p_{s,t} - \Delta \text{pnac}_t$ as measures of house prices. Even columns repeat the specification from Table 4, column 6 using these measures of house prices. The results point out to a strong and statistically significant effect of national prices on non-durable consumption and a weak and non-significant effect of the incremental price changes observed in the states on non-durable consumption. The effect of national prices is typically more than three times larger than the effect of regional prices.

3.2 Instrumental Variables

One potential problem of the results presented in the previous tables is that the house-price series are constructed using list prices observed only in the state capitals and do not have statewide coverage. This introduces measurement error which might potentially attenuate the results obtained. Table 6 uses an instrumental-variable approach to deal with this issue. We instrument the growth of housing price with the lag of the growth of mortgages in each state. The different specifications mimic the ones used in Table 4.

The evidence from Table 6 suggests that the OLS estimates understate the importance of housing wealth on consumption. The instrumental variables coefficients of the effect of house prices on non-durable consumption are above unity while the OLS coefficients presented in Table 4 range between 0.2 and 0.4. These instrumental variable estimates are statistically significant in all but one specification. The point estimates are close to the ones reported by Campbell & Cocco (2007) for the U.K. These authors find elasticities ranging from 0.57 to 1.58 while here we find elasticities ranging from 1.04 to 1.59 for Brazil. Table 7 replicates this exercise using three lags of mortgage as instruments. The results do not change.

3.3 Heterogeneity during the Business Cycle

We now turn into interpreting the evidence presented linking house prices and non-durable consumption. Campbell & Cocco (2007) describe three potential explanations for the existence of a positive correlation between house prices and non-durable consumption. First, changes in house prices might be simply proxying changes in expectations regarding economic growth. Second, changes in house prices generate a direct wealth effect. Third, changes in house prices might relax or tighten borrowing constraints the household is subject to.

Regarding alternative reasons to find a positive correlation between house prices and non-durable consumption, we must stress that, in Tables 4 through 7, we are controlling for either income or labor income growth, as well as for employment rates and credit-market conditions in each state. Controlling for income (labor income) growth is equivalent to control for the sum of expected income (labor income) growth and an unexpected shock. Because income (labor income) has positive serial correlation, it also controls in part for future income (labor income) growth. Credit-market conditions are captured by changes in demand deposits and the growth rate of loans and discounted securities.

When analyzing potential wealth effects, Campbell & Cocco (2007) note that if a home owner lives in the house, then welfare gains will be exactly offset by the present expected value of imputed rents. Indeed, in their structural model, house prices affect non-durable consumption by relaxing the agent's borrowing constraints once a price increase is observed. To try to disentangle these mechanisms, we test whether the effects of house prices on consumption are different in the "boom" and "bust" phases of the business cycle. Borrowing constraints are typically tighter in recessions (Bernanke & Gertler, 1989; Kashyap & Stein, 2000). Thus, we expect this test to indicate the importance of this mechanism.

We implement this test by estimating equation (6) including an interaction between $\Delta p_{s,t}$ and a dummy variable which is one when the economy is in recession (Bust$_t$) and zero when it is not (Boom$_t$). Based on reports from the *CODACE – Comitê de Datação de Ciclos Econômicos* on the state of the economy, we define the period between July 2014 and the December 2016 as the "bust" period. All other months are "boom" periods.

Table 8 reports the results. The different specifications mimic the ones used in Table 4. The effect of house prices on non-durable consumption is statistically significant during "booms" and "busts". However, the evidence indicates the effect of house prices on non-durable consumption is bigger during "busts" than during "booms". The point estimates range between 0.56-0.64 during the crisis and between 0.21-0.32 outside the crisis. This asymmetric effect during "booms" and "busts" is consistent with the idea that borrowing constraints are tighter during recessions and drive the relationship between house prices and consumption observed in the data.

We use the elasticities reported in Table 8 to gauge the importance of changes in house prices in explaining the behavior of consumption during the crisis. For example, in the state of Rio de Janeiro, house prices fell 7.68% from July 2014 to December 2016. This implies a reduction in non-durable consumption of 4.31%-4.91% coming from the reduction of house prices, *ceteris paribus*. Since non-durable consumption in Rio de Janeiro fell by 12.62% in this period, this suggests that non-durable consumption would have fallen just between 7.70% to 8.32% in the absence of the changes in house prices. In the state of São Paulo, in turn, house prices rose 4.37% in the period, implying an increase in non-durable consumption from 2.44% to 2.80% coming from changes in house prices. Since non-durable consumption in the state fell 1.47%, this suggests that non-durable consumption would have fallen between 3.92% to 4.27% in the period in the absence of changes in house prices. These counterfactual exercises indicate that regional heterogeneity in the behavior of house prices is important in explaining the behavior of non-durable consumption across states during the last recession.

4 Conclusion

This paper examines the impact of changes in house prices on the growth rate of non-durable consumption expenditures in Brazil using the framework proposed in Campbell & Cocco (2007) with the data structure consistent with that of Case et al. (2005). We use the theoretical model developed by Campbell & Cocco (2007) to motivate the use of a reduced-form regression approach implemented in their paper. This reduced-form regression arises from simulating their structural model of consumption and housing choice under a debt constraint related to housing values, where parameters are calibrated to fit the U.K. environment in a synthetic panel of households.

Since Brazil does not possess any database with panel data on consumption, but one can create a synthetic panel using data on consumption for Brazilian states, we chose to use the same data setup employed by Case et al. (2005) on another well-known study of housing. Using state-level data on house prices and non-durable consumption, we estimated the reduced-form regression proposed in Campbell & Cocco (2007) with a monthly state-level unbalanced panel to examine the relationship between house prices and non-durable consumption using a myriad of appropriate controls.

We find a positive and significant effect of house prices on non-durable consumption in Brazil. The magnitude of the effect we document is close to the magnitude of the effect documented by Campbell & Cocco (2007) using data from the U.K. Taking to heart the theoretical model proposed by Campbell and Cocco, a potential explanation for this positive and significant elasticity is the fact that house-price increases relax the borrowing constraint faced by the representative agent, allowing an increase in non-durable consumption in turn (and vice-versa). For the sake of completeness, we have also discussed why alternative explanations are not as plausible as relaxing the agent's borrowing constraint.

We also go one step further than Campbell and Cocco and document that the effect of house prices on non-durable consumption is asymmetric – stronger in "bust" periods than in "boom" periods of the business cycle. This is also consistent with that idea that borrowing constraints drive the results, since the latter are typically tighter during recessions.

It is possible to offer an alternative explanation for our econometric results coming from an omitted variable in our regressions. We find unlikely the existence of a plausible alternative explanation, since our marginal effects were obtained employing a variety of important controls that are potential omitted-variable candidates: demographics, real interest rates, income (labor income), loan conditions, etc.

References

Bernanke, B. (1985). Adjustment costs, durables, and aggregate consumption. *Journal of Monetary Economics*, *15*(1), 41–68.

Bernanke, B. & Gertler, M. (1989). Agency costs, net worth, and business fluctuations. *The American Economic Review*, *79*(1), 14–31.

Bertaut, C. & Starr-McCluer, M. (2002). Household portfolios in the united states. *Household portfolios*, 181–217.

Campbell, J. Y. & Cocco, J. F. (2007). How do house prices affect consumption? evidence from micro data. *Journal of monetary Economics*, *54*(3), 591–621.

Case, K. E., Quigley, J. M. & Shiller, R. J. (2005). Comparing wealth effects: the stock market versus the housing market. *Advances in macroeconomics*, *5*(1).

Gan, J. (2010). Housing wealth and consumption growth: Evidence from a large panel of households. *The review of financial studies*, *23*(6), 2229–2267.

Hofman, A. A. (1992). Capital accumulation in latin america: A six country comparison for 1950–89. *Review of Income and Wealth*, *38*(4), 365–401.

Kashyap, A. K. & Stein, J. C. (2000). What do a million observations on banks say about the transmission of monetary policy? *American Economic Review*, *90*(3), 407–428.

Luengo-Prado, M., Sorensen, B. & Hryshko, D. (2009). *House prices and risk sharing* (Tech. Rep.). Society for Economic Dynamics.

Lustig, H. & Van Nieuwerburgh, S. (2010). How much does household collateral constrain regional risk sharing? *Review of Economic Dynamics*, *13*(2), 265-294.

Marquetti, A. A. (2009). Estimativa do estoque de riqueza tangível no brasil, 1950-1998. *Nova Economia*, *10*(2).

Morandi, L. (1998). Estoque de riqueza e a poupança do setor privado no brasil: 1970-95. *Texto para Discussão, 572 - Instituto de Pesquisa Econômica Aplicada (Ipea)*.

Ogaki, M. & Reinhart, C. M. (1998). Measuring intertemporal substitution: The role of durable goods. *Journal of political Economy*, *106*(5), 1078–1098.

Ren, Y. & Yuan, Y. (2014). Why the housing sector leads the whole economy: the importance of collateral constraints and news shocks. *The Journal of Real Estate Finance and Economics*, *48*(2), 323-341.

Yogo, M. (2006). A consumption-based explanation of expected stock returns. *The Journal of Finance*, *61*(2), 539–580.

Table 1: Stock Composition of Net Capital in Brazil (%), 1950- 1994

years	Hoffman (1992 and 2000)			Marquetti (2000)		
	Building		Machinery/	Building		Machinery/
	Residential	Nonresidential	Equipment	Residential	Nonresidential	Equipment
1950	36	21	44	51	31	18
1973	29	37	34	34	47	19
1980	26	39	35	30	49	21
1989	28	44	28	33	53	14
1994	22	61	17	34	54	12

Source: Hofman (1992) and Marquetti (2009).

Table 2: Sample

State	**Initial Month**	**End Month**
Rio de Janeiro (RJ)	Feb/08	Dec/17
São Paulo (SP)	Feb/08	Dec/17
Minas Gerais (MG)	May/09	Dec/17
Bahia (BA)	Sep/10	Dec/17
Pernambuco (PE)	Jul/10	Dec/17
Ceará (CE)	Apr/10	Dec/17
Distrito Federal (DF)	Sep/10	Dec/17
Espírito Santo (ES)	Jun/12	Dec/17
Goiás (GO)	Jun/12	Dec/17
Paraná (PR)	Jun/12	Dec/17
Rio Grande do Sul (RS)	Jun/12	Dec/17

Note: Sample from Fipzap.

Table 3: Descriptive Statistics

Variables	**N**	**Mean**	**Min**	**Max**
$\Delta c_{s,t}$	1,309	0.0108	-0.310	0.419
$\Delta p_{s,t}$	958	0.0067	-0.0356	0.0460
$\text{Diff}p_{s,t}$	866	-0.0002	-0.0456	0.0362
Δpnac_t	866	0.0062	-0.00223	0.0269
r_t	1,320	0.0037	-0.0040	0.0103
$\Delta \text{Inad}_{s,t}$	1,309	0.0034	-0.300	0.331
$\Delta \text{Wage}_{s,t}$	1,320	-0.0010	-0.0178	0.0137
$\Delta y_{s,t}$	1,309	0.0031	-0.170	0.248
$\Delta \text{Loans}_{s,t}$	1,309	0.0079	-0.181	0.248
$\Delta \text{Depvista}_{s,t}$	1,309	0.0050	-0.334	0.441
$\Delta \text{Ocup}_{s,t}$	781	0.0006	-0.0455	0.0523
$\Delta \text{WorkingAge}_{s,t}$	792	0.5350	0.479	0.562

Note: Table 3 shows descriptive statistics for the main variables in this paper. $\Delta c_{s,t}$ is the average non-durable consumption growth per month (logged differences); $\Delta p_{s,t}$ the real monthly (log) changes in house prices, Δpnac_t is the real growth in house prices of the national index, and $\text{Diff}p_{s,t} = \Delta p_{s,t} - \Delta \text{pnac}_t$, deviations from national prices growth rates; r interest rate, $r_t = \ln(1 + R_t)$, where R_t is the Selic rate in real terms (deflated using IPCA); $\Delta \text{Inad}_{s,t}$ is the default rate of credit operations of the National Financial System; $\Delta \text{Wage}_{s,t}$ is the wage growth in the formal sector of the economy; $\Delta y_{s,t}$ is the Regional Economic Activity Index constructed by the Central Bank (IBC-Br); $\Delta \text{Loans}_{s,t}$ refers to the growth in loans and discounted securities; $\Delta \text{Depvista}_{s,t}$ measures the growth of demand deposits; $\Delta \text{Ocup}_{s,t}$ is the growth in the share of employment in the formal sector; and $\text{WorkingAge}_{s,t}$ is the share of people aged 25-64 years old.

Table 4: Main Results

	(1)	(2)	(3)	(4)	(5)
Dependent Variable			$\Delta c_{s,t}$		
$\Delta p_{s,t}$	0.359	0.371	0.354	0.413	0.285
	(0.065)***	(0.086)***	(0.085)***	(0.167)**	(0.134)*
r_t	1.679	1.654	1.615	1.715	2.153
	(0.131)***	(0.150)***	(0.144)***	(0.224)***	(0.227)***
$\Delta\text{Inad}_{s,t}$		-0.031	-0.031	-0.015	-0.017
		(0.019)	(0.018)	(0.020)	(0.019)
$\Delta\text{Loans}_{s,t}$		-0.031	-0.029	-0.059	-0.088
		(0.082)	(0.078)	(0.070)	(0.065)
$\Delta\text{Depvista}_{s,t}$		-0.026	-0.026	-0.005	-0.003
		(0.018)	(0.017)	(0.022)	(0.022)
$\Delta y_{s,t}$			0.086		
			(0.016)***		
$\Delta\text{Ocup}_{s,t}$				0.041	0.055
				(0.122)	(0.113)
$\Delta\text{Wage}_{s,t}$				0.879	0.434
				(0.742)	(0.830)
$\text{WorkingAge}_{s,t}$					-0.592
					(0.169)***
Constant	-0.237	-0.239	-0.236	-0.228	0.090
	(0.007)***	(0.006)***	(0.007)***	(0.006)***	(0.089)
Observations	956	956	956	747	747
R^2	0.920	0.920	0.922	0.913	0.914
Number of States	11	11	11	11	11
Month FE	Y	Y	Y	Y	Y

Note: Table 4 presents estimation results of equation 6 in five different specifications. The dependent variable is $\Delta c_{s,t}$ and the main explanatory variable of interest is $\Delta p_{s,t}$. The first column includes a control for the interest rate growth (r). Column 2 adds bank information controls. The third column adds the IBC-Br index ($\Delta y_{s,t}$), while the fourth specification replace the later by controls for the the salary growth ($g_w age$) and employment rates in the state $\Delta\text{Ocup}_{s,t}$. In column 5 we add the share of people aged 25-64 to control for demographic changes within the states. The standard errors are clustered at the state level, allowing for unrestricted residual correlation within states. All columns include state-fixed effects and controls for seasonality.

Table 5: National versus Regional Price Variation

	(1)	(2)	(3)	(4)
Dependent Variable		$\Delta c_{s,t}$		
Δpnac_t	1.376	1.287	1.383	1.271
	(0.284)***	(0.368)***	(0.285)***	(0.355)***
r_t	2.230	2.252	2.217	2.244
	(0.210)***	(0.232)***	(0.211)***	(0.230)***
$\Delta \text{Inad}_{s,t}$	-0.013	-0.014	-0.014	-0.014
	(0.020)	(0.020)	(0.020)	(0.020)
$\Delta \text{Loans}_{s,t}$	-0.089	-0.091	-0.089	-0.091
	(0.058)	(0.058)	(0.058)	(0.059)
$\Delta \text{Depvista}_{s,t}$	-0.003	-0.002	-0.003	-0.002
	(0.022)	(0.022)	(0.022)	(0.022)
$\Delta \text{Ocup}_{s,t}$	0.029	0.031	0.029	0.032
	(0.117)	(0.114)	(0.117)	(0.114)
$\Delta \text{Wage}_{s,t}$	-0.050	-0.040	-0.061	-0.049
	(0.839)	(0.846)	(0.845)	(0.851)
$\text{Diff}p_{s,t}$			0.109	0.116
			(0.134)	(0.132)
$\text{WorkingAge}_{s,t}$		-0.079		-0.100
		(0.197)		(0.193)
Constant	-0.233	-0.191	-0.233	-0.179
	(0.006)***	(0.105)	(0.006)***	(0.103)
Observations	747	747	747	747
R^2	0.914	0.914	0.914	0.914
Number of States	11	11	11	11
Month FE	Y	Y	Y	Y

Note: Table 5 presents estimation results of equation 6 in four different specifications. The dependent variable is $\Delta c_{i,t}$ and the main explanatory variable of interest is Δpnac_t. Column 1 includes all the controls included in the last column of Table 4 except for the control on $\text{WorkingAge}_{s,t}$, which is included in column 2. Columns 3 and 4 follow the same structure of the previous columns but we are also interested in understand the relation between consumption abut add $\text{Diff}p_{s,t}$. The standard errors are clustered at the state level, allowing for unrestricted residual correlation within states. All columns include state-fixed effects and controls for seasonality. We impose *strict exogeneity* of the regressors, conditional on the unobserved effect a_s. Thus, estimation of the β's is performed using the so called *fixed-effects estimator*, which is the pooled OLS estimator on time-demeaned data. The latter eliminates a_i from the system. Since the error term is dynamically incomplete and possibly heteroscedastic, robust inference has to be conducted to account for time-depedence and heteroskedasticity of unknown form.

Table 6: Instrumental Variables (1)

	(1)	(2)	(3)	(4)
Dependent Variable		$\Delta c_{s,t}$		
$\Delta p_{s,t}$	1.037	1.117	1.588	1.318
	(0.324)***	(0.349)***	(0.749)**	(1.235)
r_t	1.922	1.851	2.084	2.178
	(0.422)***	(0.422)***	(0.514)***	(0.484)***
$\Delta\text{Inad}_{s,t}$		-0.030	-0.018	-0.018
		(0.027)	(0.031)	(0.031)
$\Delta\text{Loans}_{s,t}$		-0.099	-0.091	-0.095
		(0.067)	(0.078)	(0.076)
$\Delta\text{Depvista}_{s,t}$		-0.028	-0.003	-0.003
		(0.019)	(0.023)	(0.022)
$\Delta\text{Ocup}_{s,t}$			0.031	0.038
			(0.172)	(0.171)
$\Delta\text{Wage}_{s,t}$			-0.310	-0.250
			(1.043)	(1.101)
$\text{WorkingAge}_{s,t}$				-0.221
				(0.495)
Constant	-0.242	-0.243	-0.234	-0.115
	(0.004)***	(0.005)***	(0.007)***	(0.271)
Observations	954	954	747	747
Number of States	11	11	11	11
Month FE	Y	Y	Y	Y

Note: Table 6 presents a robustness check using an instrumental-variable approach. We instrument the growth of housing price with the lag of the growth of mortgages . The first column includes a control for the interest rate growth. Column 2 adds bank information controls. The third column adds control for the growth in the employment rates and wages in the formal sector. Column 4 adds control for demographic changes within the states. The standard errors are clustered at the state level, allowing for unrestricted residual correlation within states as in Table 4 and all columns include state-fixed effects and controls for seasonality.

Table 7: Instrumental Variables (2)

	(1)	(2)	(3)	(4)
Dependent Variable		$\Delta c_{s,t}$		
$\Delta p_{s,t}$	0.902	0.998	1.496	1.088
	(0.254)***	(0.273)***	(0.529)***	(0.842)
r_t	1.873	1.811	2.055	2.172
	(0.413)***	(0.415)***	(0.484)***	(0.479)***
$\Delta\text{Inad}_{s,t}$		-0.028	-0.018	-0.018
		(0.026)	(0.031)	(0.031)
$\Delta\text{Loans}_{s,t}$		-0.094	-0.088	-0.094
		(0.064)	(0.076)	(0.075)
$\Delta\text{Depvista}_{s,t}$		-0.028	-0.003	-0.003
		(0.019)	(0.022)	(0.022)
$\Delta\text{Ocup}_{s,t}$			0.032	0.042
			(0.171)	(0.169)
$\Delta\text{Wage}_{s,t}$			-0.217	-0.098
			(0.891)	(0.918)
$\text{WorkingAge}_{s,t}$				-0.303
				(0.371)
Constant	-0.241	-0.243	-0.234	-0.069
	(0.004)***	(0.005)***	(0.006)***	(0.203)
Observations	950	950	747	747
Number of States	11	11	11	11
Month FE	Y	Y	Y	Y

Note: Table 7 presents a robustness check using an instrumental-variable approach. We instrument the growth of housing price with the lags of the growth of mortgages (3 lags). The first column includes a control for the interest rate growth. Column 2 adds bank information controls. The third column adds control for the growth in the employment rates and wages in the formal sector. Column 4 adds control for demographic changes within the states. The standard errors are clustered at the state level, allowing for unrestricted residual correlation within states as in Table 4 and all columns include state-fixed effects and controls for seasonality.

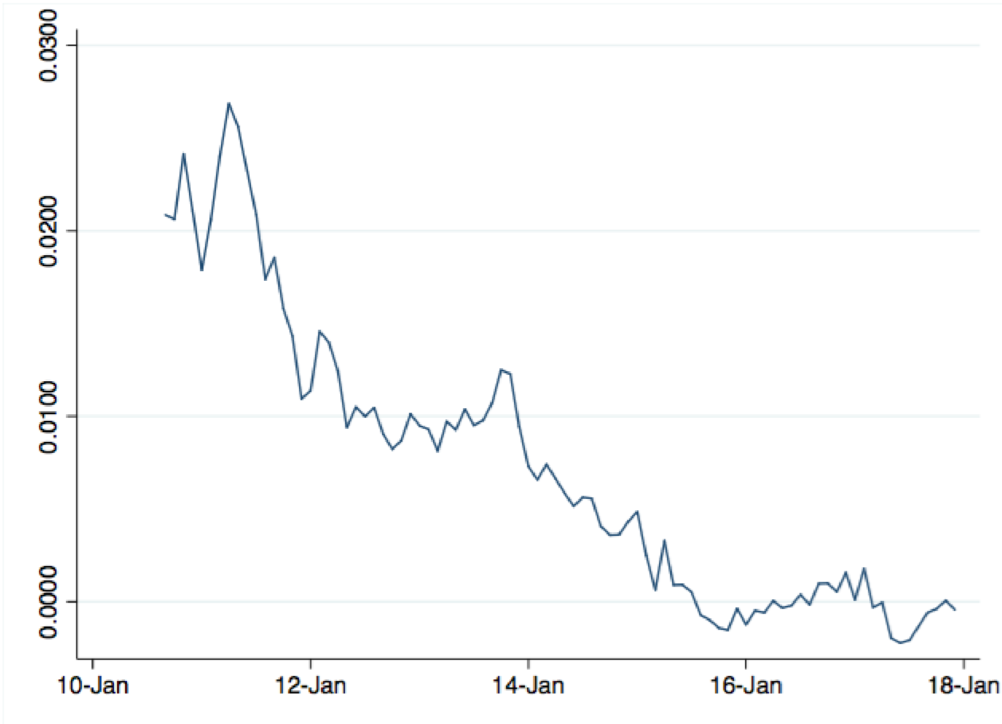
Table 8: Heterogeneity during the Business Cycle

	(1)	(2)	(3)	(4)
Dependent Variable	Δc	Δc	Δc	Δc
$\Delta p_{s,t} \times \text{Boom}_t$	0.216	0.231	0.321	0.215
	(0.060)***	(0.079)**	(0.141)**	(0.119)
$\Delta p_{s,t} \times \text{Bust}_t$	0.628	0.644	0.625	0.560
	(0.317)*	(0.315)*	(0.357)	(0.347)
Bust_t	-0.007	-0.007	-0.007	-0.005
	(0.001)***	(0.002)***	(0.002)**	(0.002)*
r_t	1.805	1.764	1.821	2.163
	(0.146)***	(0.142)***	(0.190)***	(0.217)***
$\Delta\text{Inad}_{s,t}$		-0.026	-0.010	-0.014
		(0.018)	(0.019)	(0.019)
$\Delta\text{Loans}_{s,t}$		-0.051	-0.063	-0.086
		(0.077)	(0.066)	(0.064)
$\Delta\text{Depvista}_{s,t}$		-0.029	-0.006	-0.004
		(0.020)	(0.023)	(0.022)
$\Delta\text{Ocup}_{s,t}$			0.004	0.029
			(0.126)	(0.118)
$\Delta\text{Wage}_{s,t}$			0.292	0.101
			(0.892)	(0.928)
$\text{WorkingAge}_{s,t}$				-0.520
				(0.141)***
Constant	-0.234	-0.235	-0.225	0.053
	(0.007)***	(0.006)***	(0.006)***	(0.073)
Observations	956	956	747	747
R^2	0.921	0.921	0.913	0.914
Number of States	11	11	11	11
Month FE	Y	Y	Y	Y

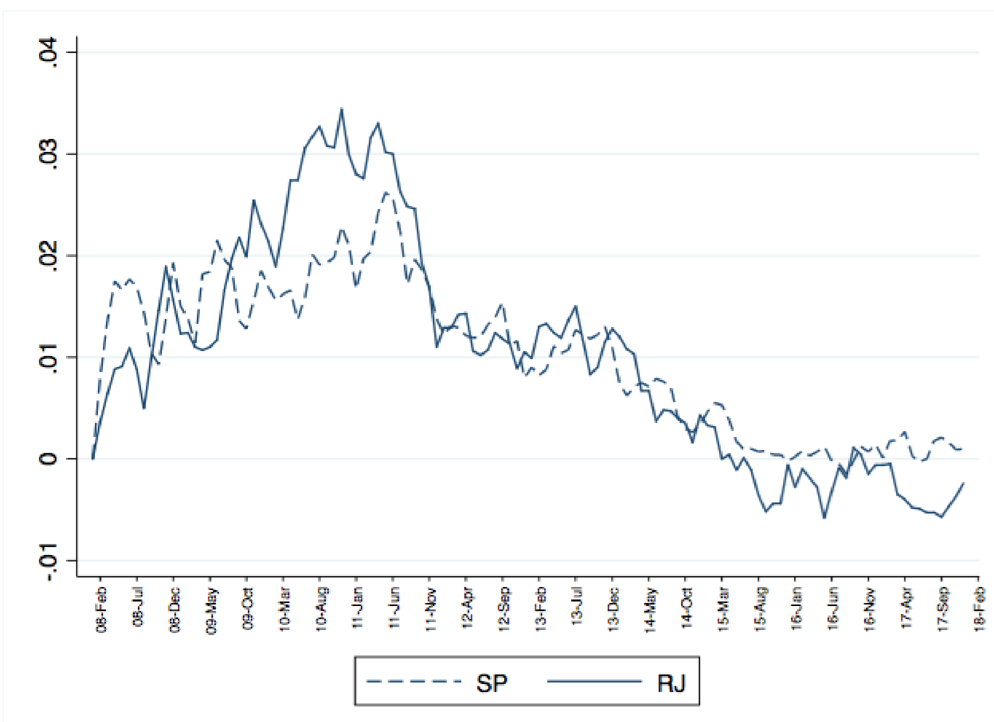
Note: Table 8 follows the same structure of Table 6 but explore the time variation to estimate equation 6. $\Delta p * Bust_t$ allows to estimate differential effects after and before the economic crisis faced by Brazil in 2014 (before and after july 2014).

Figure 1: Real Growth Rate - House Prices

(a) Brazil



(b) Rio de Janeiro and São Paulo

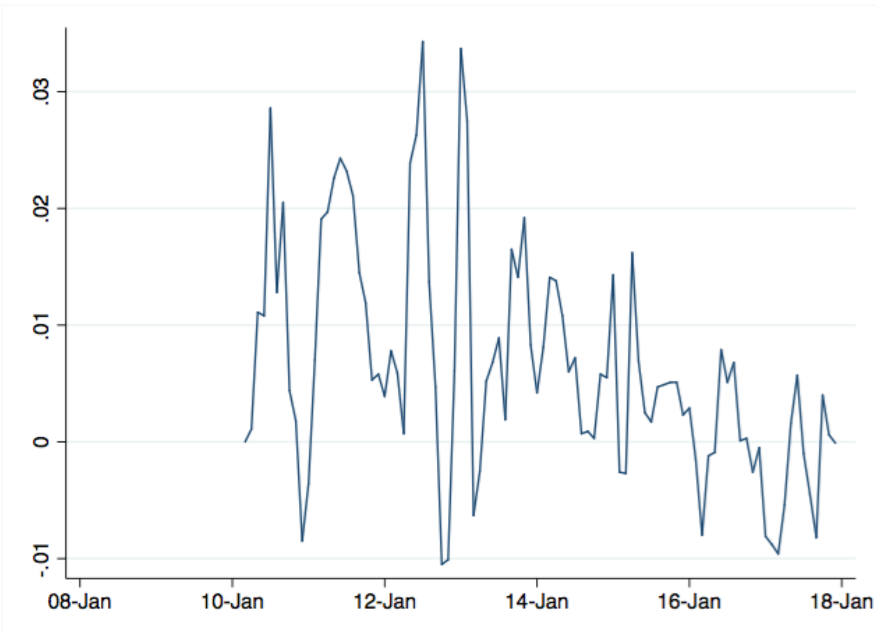


Note: Panel A of the figure reports monthly real growth rates of house prices from February 2008 to December 2017 for Brazil. Panel B reports monthly real growth rates of house prices in São Paulo and Rio de Janeiro from February 2008 to December 2017.

Appendix to "Non-Durable Consumption and Real-EstatePrices in Brazil"

A Appendix

Figure A1: Real Growth Rate - House Prices - Ceará

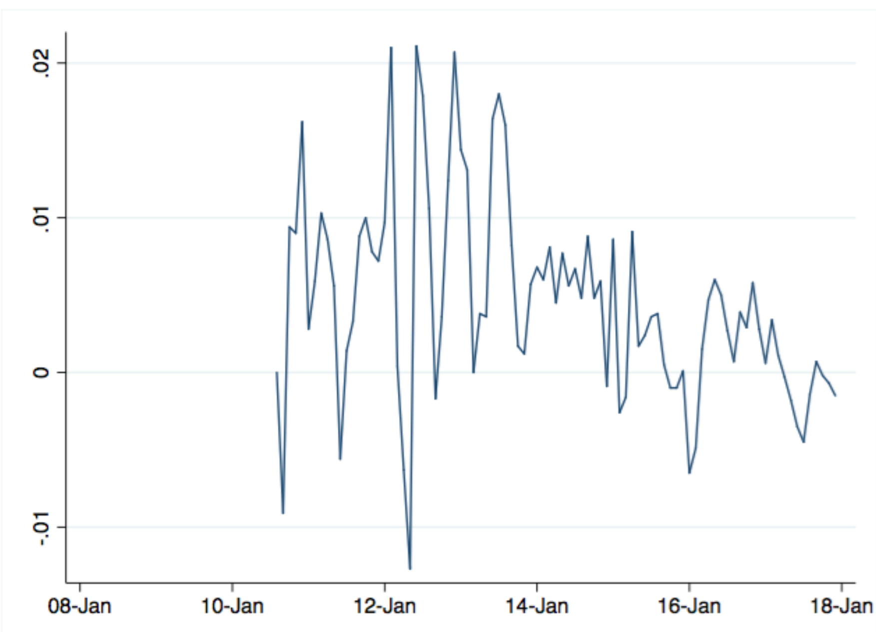


Notes: The figure reports monthly real growth rates in house prices from February 2008 to January 2018.

Figure A2: Real Growth Rate - House Prices - Pernambuco

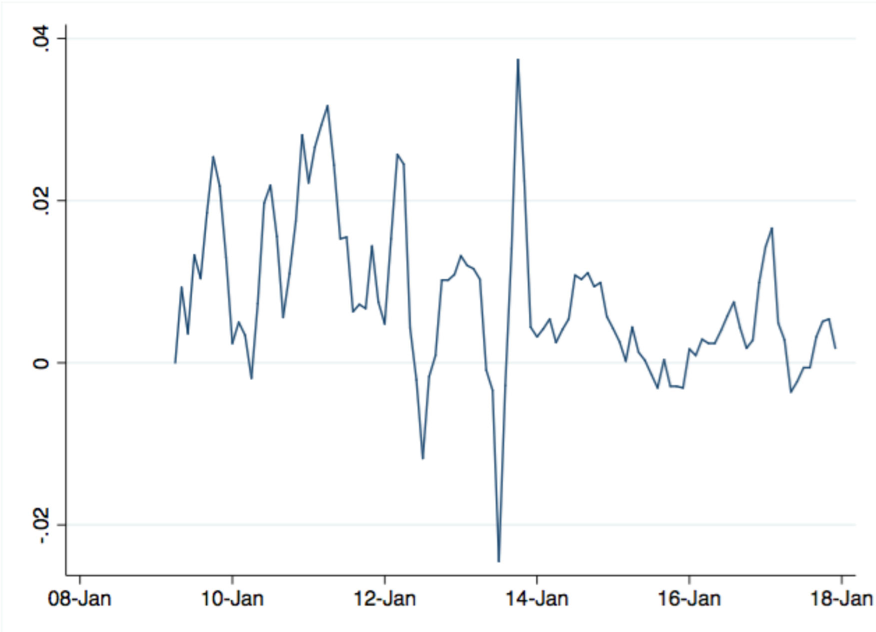
Notes: The figure reports monthly real growth rates in house prices from February 2008 to January 2018.

Figure A3: Real Growth Rate - House Prices - Bahia



Notes: The figure reports monthly real growth rates in house prices from February 2008 to January 2018.

Figure A4: Real Growth Rate - House Prices - Minas Gerais

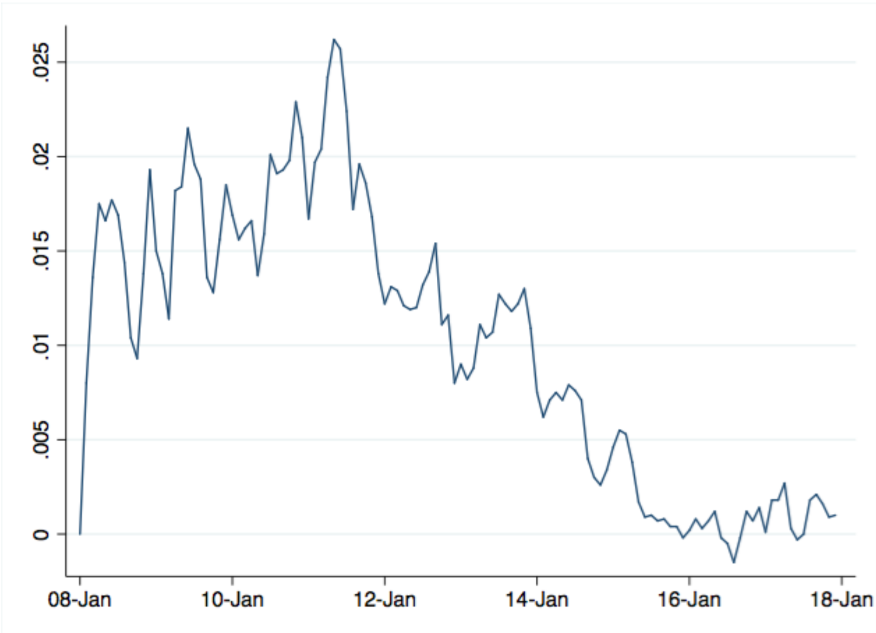


Notes: The figure reports monthly real growth rates in house prices from February 2008 to January 2018.

Figure A5: Real Growth Rate - House Prices - Espírito Santo

Notes: The figure reports monthly real growth rates in house prices from February 2008 to January 2018.

Figure A6: Real Growth Rate - House Prices - São Paulo



The figure reports monthly real growth rates in house prices from February 2008 to January 2018.

Figure A7: Real Growth Rate - House Prices - Paraná

The figure reports monthly real growth rates in house prices from February 2008 to January 2018.

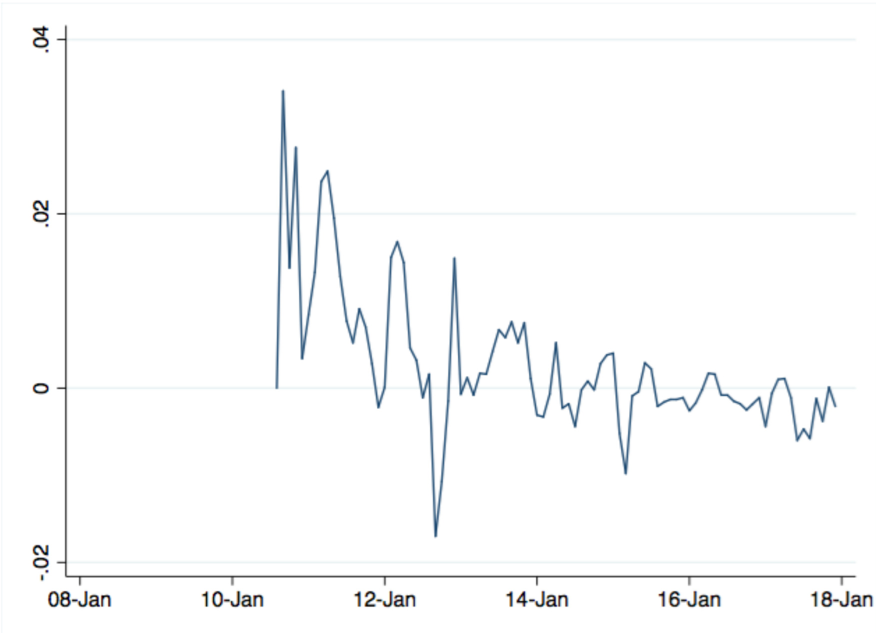
Figure A8: Real Growth Rate - House Prices - Rio Grande do Sul

The figure reports monthly real growth rates in house prices from February 2008 to January 2018.

Figure A9: Real Growth Rate - House Prices - Goiás

The figure reports monthly real growth rates in house prices from February 2008 to January 2018.

Figure A10: Real Growth Rate - House Prices - Distrito Federal



The figure reports monthly real growth rates in house prices from February 2008 to January 2018.