UNIVERSITÀ COMMERCIALE LUIGI BOCCONI
PH.D. SCHOOL

Ph.D. Program in Economics and Finance

Cycle: 34

Disciplinary field: SECS P/01 Economia Politica

Three Essays in Public Economics

Advisor:

Prof. G. I. P. Ottaviano

Candidate:

Bence Szabó

ID number: 3081185

Year 2023

Acknowledgements

As Professor Pavoni told us in the first year of the program, it is not a sprint, it is a marathon. Accordingly, this Ph.D. thesis would not have been possible without a steady, unraveling net of support.

First and foremost, I thank my parents and my sister for their unfaltering help and love.

Second, I thank my friends and colleagues in Milan, who made these years of ups and downs so much more joyful. I would like to thank especially Francesca, Ádám, Daria, Qi, Alberto, Manuela, and Goonj, for all the memorable and lovely evenings through these difficult years of pandemic and war. I would also like to mention the support we received from our colleagues from the previous cohorts, especially Blanka, who gave me so much advice on how to navigate my academic and personal life in Italy.

Third, I thank my friends and loved ones at home in Hungary, who encouraged and motivated me, or even made great sacrifices for me throughout these years and provided a safe haven to return to.

Fourth, I thank my supervisors, Gianmarco Ottaviano, Joseph-Simon Görlach, and Francesco Billari, for all the professional and personal advice and guidance they have given me over the years. I am incredibly thankful to Simon for the attentiveness and thoroughness with which he followed my progress. Furthermore, I am also grateful to Professors Pamela Giustinelli and Guido Tabellini for all the extracurricular support throughout my years at Bocconi.

Finally, I am very thankful to all those kind people that provided useful feedback on the papers that have turned into these thesis chapters. I indicated their names below each abstract.

Contents

List of Figures

List of Tables

Introduction

This thesis contains three chapters based on my studies in the field of demographic, public, and to a smaller extent, labor economics. The chapters are connected by the common theme of studying the interaction between government interventions and household behavior, touching on three major aspects of life: fertility, charity, and housing.

My job market paper, titled "Tax Incentives and Completed Fertility", constitutes the first chapter of the thesis. It studies the effect of tax incentives on the number of children by the end of the female fertility cycle by identifying the impact of a large-scale family tax break reform in Hungary. I use linked Census data to estimate the average treatment effect on the treated, explore the heterogeneity of the effect, and build a small structural model to examine counterfactual policy scenarios. I find a positive impact of the policy on completed fertility in a relatively large magnitude compared to the previous literature.

The second chapter is the result of a joint project with a Ph.D. colleague, Francesca Leombroni, titled "Charitable Behavior and Public Intervention: a Survey Experiment". We examine how the availability of donations affects the overall welfare of the poor in a society with equilibrium tax rates using a survey experiment in the U.S., in which participants elicited their preferences over taxes, their donation expectations, and their choices conditional on tax rates. We find that the availability of donations plays a pivotal role in providing financial resources to the poor in the U.S.

Finally, the third chapter, titled "Housing and Fertility", studies the interaction between housing choices and fertility in the context of a housing subsidy conditional on the number of children introduced in Hungary. I build and estimate a life-cycle model of households, and I simulate temporal pseudo-equilibria on the housing market to examine how different policy instruments would affect fertility, location, and house size choices of families. I find that while housing policies with children-related requirements positively impact fertility for poorer households, their housing conditions would worsen in the long run.

Chapter 1

Tax Incentives and Completed Fertility

Abstract

This paper studies the effect of tax incentives on completed fertility based on the 2011 family tax break reform in Hungary, which introduced a disproportionately large increment for the third child. I build a unique, family-level linked dataset from the 2016 Microcensus and the 2011 Census of Hungary to measure changes in the number of children within families, and labor market outcomes while observing a rich set of individual parental pre-policy characteristics. I use variation along the initial number of children and prospective household income in 2011 and compare cohorts just before and after the end of the fertility cycle. I argue that the policy created a quasi-experiment along these dimensions, enabling the identification of the conditional average treatment effect on the treated. The estimates suggest that the policy increased completed fertility by around 0.8% for families with at least two children and around 5.6% for families with fewer than two. The findings are driven by religious and married couples with young children. No significant indirect short-run effects on maternal employment are detected. Furthermore, I build and estimate a simple household model to study counterfactual policy scenarios and long-run partial equilibrium effects and find qualitatively similar completed fertility effects driven by middle-income families.

Keywords: Completed Fertility, Family Tax Break

JEL Classification: J13, J18, H24

I thank my thesis supervisors, Gianmarco Ottaviano, Joseph-Simon Görlach, and Francesco Billari, for the years of excellent guidance and support. I am grateful for the comments and discussions with Jérôme Adda, Gábor Békés, Renata Bottazzi, Marco Di Maggio, Sarah Eichmeyer, Stefano Fiorin, Hedvig Horváth, Blanka Imre, Gábor Kertesi, János Köllő, Thomas Le Barbanchon, Francesca Leombroni, Jacopo Lunghi, Manuela Puente Beccar, Carlo Schwartz, Maria Sironi, Guido Tabellini, and Bálint Ván. I am also grateful for the valuable remarks of the seminar participants at the Centre for Economic and Regional Studies - Institute of Economics in Hungary, the Hungarian Ministry of Finance, and the Political Economy Reading Group at Bocconi University. Finally, I thank the Centre for Economic and Regional Studies for supporting this research, the Databank of the Centre for Economic and Regional Studies, and the Hungarian Central Statistical Office for providing access to the databases and all the technical help their colleagues have given me.

1.1 Introduction

Persistent low fertility is one of the central policy issues in developed countries. Since the 1970s, the total fertility rate[1] in OECD countries have declined from around 2.84 to 1.61, below the replacement level of 2.1 that would maintain the population size. As Figure 1.1a shows, below replacement level fertility occurs in many developed countries: the United States experienced a rapid decline of around 0.5 in total fertility rate in the last fifteen years to reach approximately 1.6, while during this period, the European Union stagnated around values of 1.5-1.6. In some countries of Southern, Central, and Eastern Europe, total fertility rates reached levels that are considered 'lowest-low', below the value of 1.3.[2] Countries with lowest-low fertility and no substantial migration are projected to experience worsening living standards in the long run (Lee and Mason, 2014), even if somewhat lower than replacement level fertility rates could be beneficial in terms of maximal consumption per capita. Therefore, governments might be inclined to boost fertility through policies even if the decline is due to the changing preferences of families regarding the number of children. However, evidence suggests that the realized number of children actually lags behind the preferred amount in many developed countries[3]. Figure 1.1b displays the difference between average completed fertility at age 40 and the ideal number of children for women between ages 15-39 in European countries, showing that in many countries, families end up with fewer children on average than what they would ideally prefer. So the economic interests of governments align with the potential for welfare improvements at the individual level to implement successful family policies and increase fertility levels.

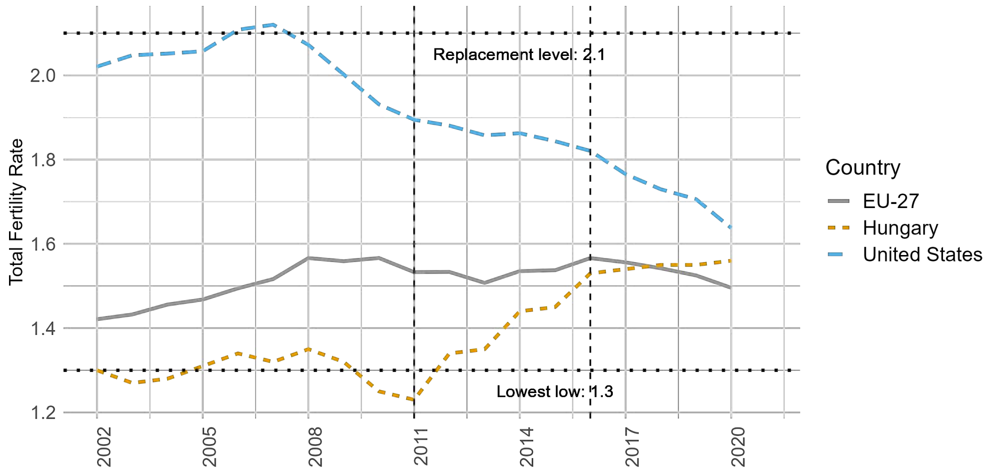
[1] Total fertility rate according to the OECD is defined as 'the total number of children that would be born to each woman if she were to live to the end of her childbearing years and give birth to children in alignment with the prevailing age-specific fertility rates. It is calculated by totaling the age-specific fertility rates defined over five-year intervals. Assuming no net migration and unchanged mortality, a total fertility rate of 2.1 children per woman ensures a broadly stable population.', source: https://data.oecd.org/pop/fertility-rates.htm

[2] At which birth cohorts would shrink to half of their initial size in 45 years (Kohler et al., 2006).

[3] Measuring fertility preferences is a complex problem, as described by (Spéder and Kapitány, 2009).

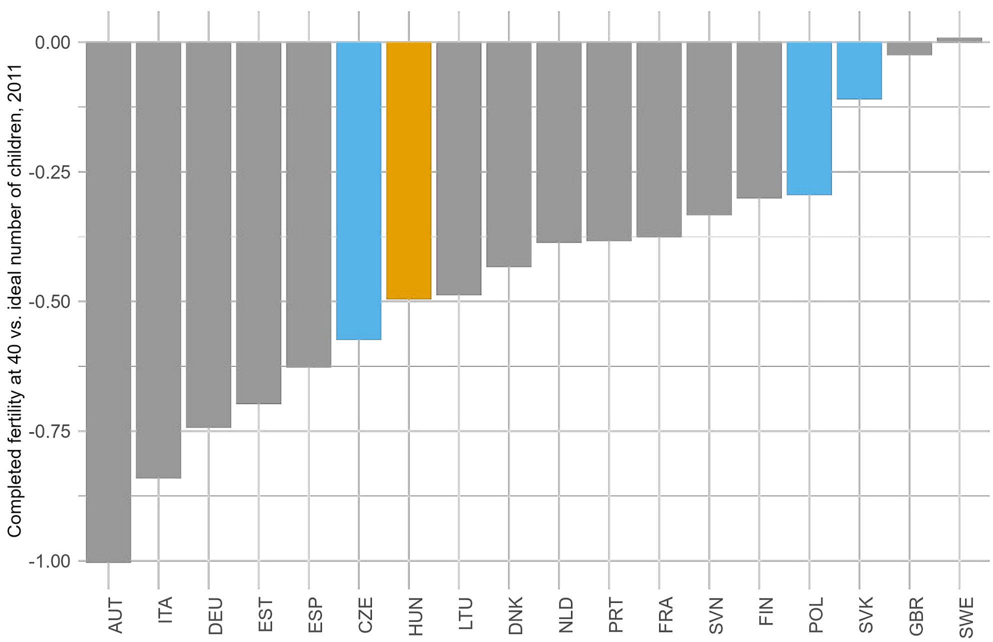
Figure 1.1: Key stylized facts of realized and ideal fertility in the developed world

(a) Evolution of total fertility rates in developed countries



Note: based on the data of the World Bank. Total fertility rate measures the total number of children that would be born to each woman if she were to live to the end of her child-bearing years and give birth to children in alignment with the prevailing age-specific fertility rates (OECD).

(b) Completed fertility vs. ideal number of children in 2011



Note: based on the data of the Eurobarometer and the Human Fertility Database. Completed fertility at 40 measures the number of children on average for women at age 40, and the average ideal number of children is provided by women aged 15-39.

In this paper, I examine the causal effect of taxation-related financial incentives on completed fertility by studying a large-scale family tax break reform in Hungary. The country stagnated at 'lowest-low' fertility levels for an extended period before 2011, rebounding to the EU average by 2016. The Hungarian Government introduced a significant expansion of the family tax break in 2011, increasing its magnitude from 0.05% of the GDP to 0.64% of the GDP, representing around 17% percent of all public spending on families (OECD, 2021). The tax break does not phase out at higher income levels, and the household's tax base for personal income taxation limits the maximum amount of money receivable. Additionally, the policy has a non-linear structure due to a large jump in deductible amounts for the third child and because of the non-refundability of the deductions. Indeed, with three children, a typical household in Hungary did not have enough personal income to claim the entirety of the tax break.

I argue that the policy change enables the identification of the effects of tax incentives on completed fertility. To achieve this, I construct a family-level linked dataset based on the 2016 Microcensus (post-policy) and the 2011 Census (pre-policy), allowing us to observe a rich of pre-policy observable characteristics. According to the labor market status of the couple, I impute the prime age prospective income in the household using conditional averages observed in National Wage Surveys. Based on prospective household income, I calculate the financial incentives introduced by the policy and use regression difference-in-differences to compare the within-cohort completed fertility differences for those cohorts that can react to the policy vs. those that cannot due to being post-fertile age. Identification of the conditional average treatment effect on the treated relies on the assumption that for families in the older cohorts, the treatment would have induced similar change on average as it did for those that could still increase their number of children. Along with completed fertility, I also examine indirect, short-run maternal employment effects while exploring the heterogeneity of the results. Finally, I build a simple partial equilibrium structural model to study counterfactual scenarios, compare long-run fertility effects to the estimates, and shed light on the long-run effects regarding women's labor supply.

The findings suggest that the policy increased completed fertility for families with at least two children by around 0.8% and for those with fewer than two children by around 5.6%, with the weighted average being approximately 2.4%. The results are driven by religious, married couples and those whose youngest child is less than three years old. Estimates regarding indirect maternal employment are not significant overall; however, there seems to be some evidence for short-run lower employment for those segments where completed fertility increased, suggesting slower transitioning back to the labor market. The estimates are robust to several changes in the specification, such as different measurements of the outcome and treatment variables or changes in the sample restrictions. I also estimate a simple partial equilibrium model of household behavior to simulate counterfactual policies, and to examine possible long-run effects regarding maternal employment, completed fertility, and income redistribution. A comparative statics exercise using the structural model indicates around 3.2% higher completed fertility after the policy driven by middle-income families, and around 4.8 percentage points lower employment driven by lower-income families. I also find that alternative family policies might have led to slightly higher fertility effects without changing the overall public expenditure and with less regressive income distribution.

Economists have been examining fertility decisions and their effects for centuries, dating back to Malthus (1798), and in modern times to the seminal works of Gary Becker and coauthors on the economics of the family (Becker, 1960; Becker and Lewis, 1973; Becker and Tomes, 1976). Fertility choices are deeply intertwined with short and long-run labor supply decisions of the household, disproportionately affecting women. The question is traditionally modeled jointly as part of the production of market and household goods while trading off the quantity and quality of children. The microeconomic literature has studied several taxation policies affecting fertility (Cigno, 1986, 2001; Fraser, 2001; Balestrino and Cigno, 2002; Cigno and Pettini, 2002; Docquier, 2004; Meier and Wrede, 2013; Baudin et al., 2015), resulting in a wide array of predictions about optimal policies with a dependence on the underlying preferences. While these theoretical papers focus on completed fertility and the change in the number of children by the end of the female fertility cycle, such effects are not easy to detect empirically.

Most policies are expected to induce two types of adaptation: a change in the timing of births and a change in completed fertility, also called *tempo* vs. *quantum* effects. It is easy to see that several policies might induce families to change the timing of births to receive the newly available benefits earlier. Still, in the end, the completed number of children might remain unaffected. For instance, papers studying a reform in Quebec found that in the short-run, fertility increased by 9% due to the policy (Milligan, 2005), follow-up investigations concluded that completed fertility was, in the end, unaffected (Parent and Wang, 2007; Kim, 2014). Adda et al. (2017) estimates a dynamic structural model using German data to study women's fertility decisions and career choices to differentiate between short-run and long-run consequences. Similarly, they find that relatively large short-run effects of a one-time birth subsidy are dominantly due to a change in the timing of births, and they are mitigated over the life-cycle of women.

The fertility effects of economic circumstances such as unemployment or insecurity are well-studied empirically (Currie and Schwandt, 2014; Sommer, 2016; Clark and Lepinteur, 2020), and there is a separate branch of literature that specializes in examining the role of tax-related incentives. One such branch of literature focuses on the Earned Income Tax Credit in the United States, making use of low-level aggregated data in quasi-panel settings (Moffitt, 1998; Baughman and Dickert-Conlin, 2003; Gauthier, 2007; Baughman and Dickert-Conlin, 2009; Thévenon and Gauthier, 2011), Bördős and Szabó-Morvai (2021) applied this style of designed to analyze family policies in Hungary as well. The main weakness of such studies is that the examined subsets or cells within the population might alter in composition due to gradual demographic change, so the generally small positive fertility effects estimated by this literature still leave place for further investigation (Hoynes, 2019). Additionally, short-run timing adjustments and long-run completed fertility adjustments are often not treated separately.

In the last two decades, more studies have been investigating individual-level evidence of countries with fertility-related tax policies (Lalive and Zweimüller, 2009; Azmat and González, 2010; Kalwij, 2010; Cohen et al., 2013; González, 2013; Cygan-Rehm, 2016; Dahl et al., 2016; Garganta et al., 2017; Riphahn and Wiynck, 2017; Raute, 2019). The results indicate overall positive effects from such policies with varying magnitudes. Still, similarly to the quasi-panel evidence, these papers also identify short-run treatment effects, mostly with innate limitations to address timing vs. completed fertility convincingly. On the other side of the methodological spectrum, we can find structural econometric models explicitly estimating the life-cycle models for women and separate the short-run and long-run effects of policies (Moffitt, 1984; Francesconi, 2002; Keane and Wolpin, 2002, 2010; Laroque and Salanié, 2014; Adda et al., 2017). However, the conclusions regarding completed fertility rely on counterfactual simulations.

I contribute to the literature of quasi-experimental studies in demographic economics, public economics, and to a smaller extent, labor economics in the following way. I take advantage of the unique structure of a tax policy targeting fertility, enabling causal identification of the quantum effect of large-scale government intervention in a developed country. At the same time, I avoid the problem of usual short-run estimates that might also capture the timing effect. I find a magnitude larger effect on long-run, completed fertility compared to earlier estimates while staying below the short-run estimates that might jointly capture the quantum and tempo effects of the policy. The constructed new family-level linked census dataset provides individual-level evidence with rich pre- and post-policy observable characteristics, including often ignored couples that are unmarried or without children. The dataset enables the study of the heterogeneity of the treatment effects, providing a future reference for targeted policy interventions. Finally, I study the short-run and long-run changes in maternal employment and completed fertility by augmenting the reduced form estimates with a simple model comparing relevant policies, considering the redistributive impact.

The structure of the paper is as follows. I introduce the context of the family tax break extension and describe the policy in Section 1.2. In Sections 1.3 and 1.4, I detail the construction of the analysis dataset and establish the empirical strategy for identifying the reduced form results. In Sections 1.5 and 1.6, I present the results, run several robustness checks, and then continue in Section 1.7 to build and estimate a simple household model examining different counterfactual scenarios. Finally, I discuss the findings in Section 1.8 and conclude in Section 1.9.

1.2 Context: the 2011 family tax break extension in Hungary

This paper focuses on the case of Hungary, a country that had experienced a steady decline in total fertility rate since 1990, along with several other countries in the region (Billari and Kohler, 2004; Sobotka, 2011), reaching its minimum of around 1.23 in 2011. Hungary has become remarkably older in its age structure, while the population size has also shrunk, primarily due to the low fertility rate (Őri and Spéder, 2020). In this period, fertility behavior can be characterized by a postponement of the first child's birth, which would not necessarily mean a decline in completed fertility if these children are eventually born, so if a timing effect occurs only (Bongaarts and Sobotka, 2012). However, in Hungary, evidence suggests that especially at the second birth parity, completed fertility also declined (Kapitány and Spéder, 2015). At the same time, childlessness has slowly become more prevalent for younger cohorts (Spéder, 2021).

Although some smaller interventions affected fertility behavior before 2011 as well (Gábos et al., 2009; Spéder et al., 2020), the governments after 2010 made families a focal point of economic policy.[4] The Hungarian Government introduced a complete overhaul of the personal income tax system in 2011, including the large-scale extension of the family tax break (National Assembly of Hungary, 2010). Beforehand only a 4,000 HUF/child/month (around 15 EUR at 2010 exchange rates) tax allowance was available for families with at least three children (along with eligibility constraints for high-income families), and the tax break was not considered a significant part of the child subsidy policies in the country (Makay, 2015). The 2011 reforms extended the circle of eligibility and the amount of the available subsidy. Since then, families with less than three children also became eligible for a smaller degree, but the tax break includes a large jump for the third child. According to the OECD Family Database, the family tax break corresponded to around 0.64% of the GDP (compared to the 0.05% in 2010), around 17% of all public spending on families in 2011 (for the other

[4]Beside extending the family tax break, the Hungarian government also introduced the family housing allowance in 2015, coming into action in 2016, providing a large lump sum to purchase homes for those committing to having three children within a set period of time. Given the identification strategy and the time windows, that policy's effects should not bias the results significantly.

main policy instruments, see Table 1A1). The overhaul of the family tax break demonstrably played a part in a major redistribution towards higher income families with more children (Tóth G. and Virovácz, 2013; Varga et al., 2020).

The policy instrument itself is a non-refundable tax base deduction that can be exercised by either one of the parents or jointly, regardless of marital status. However, in its initial version, it is implicit that the tax base deduction cannot exceed the available taxable income of the family[5]. It results in profound consequences regarding the tax break change for an additional child based on the family's income. If the tax base deductions are great enough (e.g., the family has three or more children), high-income families have more financial incentives (lower marginal cost) than low-income families to have an additional child. Contrary to the Earned Income Tax Credit in the United States, and more similar to the Child Tax Credit, there is no phasing out at higher levels of income (Hoynes, 2019; Goldin and Michelmore, 2021), which results in regressive redistribution.

Indeed, the tax base deductions were sizable compared to average salaries in Hungary, which for reference at that time stood around 215,000 HUF$\sim$ 800 EUR per month per person[6]. If the family had fewer than two children, the deductions amounted to 62,500 HUF ($\sim$ 230 EUR) per child per month, while for families with at least three children, it rose to 208,500 HUF ($\sim$ 770 EUR) per child per month. Under the new regime, a family with parents earning the average salary would not be able to claim the full amount of the tax break with three children; a family would need 45% more than the average income.

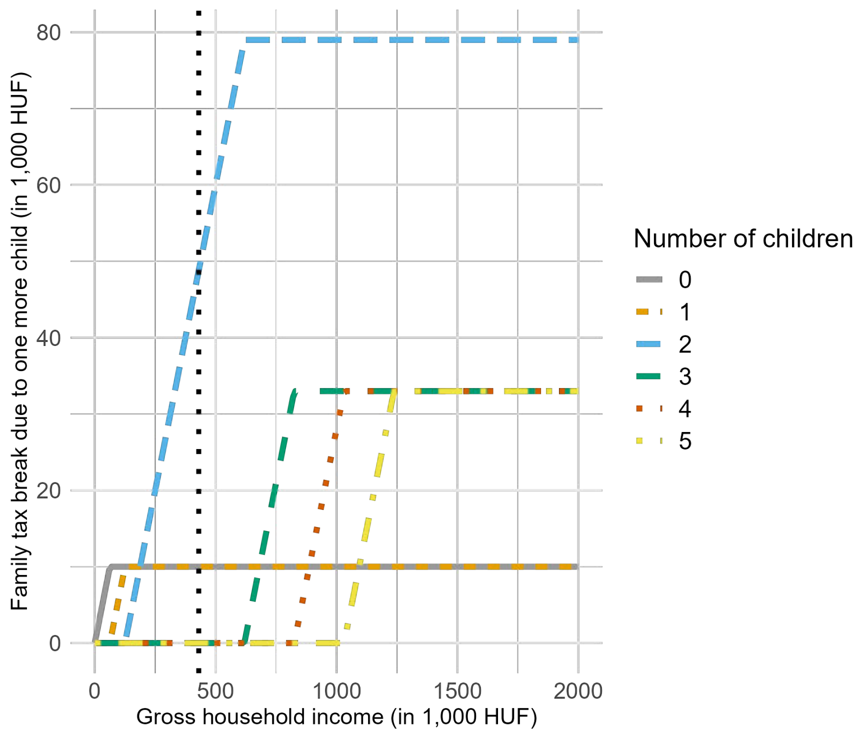
Figure 1.2 plots the additional money a family would receive after having one more child against household income by the number of children. The vertical line indicates the household income with two people earning average salaries. We can see that for households with zero or one child, the tax break increment is around 10,000 HUF per month ($\sim$ 37 EUR). However, for households with two children, one additional child could give around an 80,000 HUF monthly increment ($\sim$ 290 EUR) if their income is sufficiently high, increasing at the slope

[5]Note that since 2014, families not being able to claim the total amount based on personal income taxes are allowed to decrease their health and pensions contributions up to 15% of the remaining amount of the deduction. Even though many families are affected (Makay, 2019), in this paper, I abstract away from these changes as the family tax break is still not fully refundable, so the financial magnitudes remain unchanged.

[6]Source: Hungarian Central Statistical Office, I use a 270 HUF/EUR exchange rate which approximates well the 2011 value

of 0.16 along with income due to the flat 16% personal income tax rate. For families with three children, the tax break increases only between 620 and 825 thousand HUF of income: lower-income families will not be able to have any more deductions due to their low tax base, while higher-income families will exhaust the tax break. For four children, we see the same but at the 825,000 and 1,030,000 HUF (∼ 3,000 EUR, ∼ 3,800 EUR) household income levels.

Figure 1.2: Additional family tax break for one more child



Note: the figure shows the amount of family tax break families would get for one additional child, based on household income and the number of children. The vertical dotted line in the figure indicates the level of household income with around two average gross salaries.

1.3 Data

1.3.1 Analysis sample

To study the question, I build a unique dataset based on three databases: the 2016 Microcensus of Hungary, the 2011 Census of Hungary, and the 2010-2016 waves of the National

Wage Surveys. The Microcensus contains around 340,000 family units, a 10% anonymized sample of the entire population, representative on the household level according to counting district, age, education, employment, marital status, and residency. The 2011 Census aims to contain the entire population living in the country, also anonymized, entailing around 4,140,000 families. The National Wage Surveys are composed of random samples of employees of larger companies and public institutions and random samples of smaller firms reporting on all of their employees.

First, I describe the steps of data manipulation concerning the censuses. I use the 2016 Microcensus and the 2011 Census of Hungary to build a unique dataset by matching families across the two, utilizing the birth dates of the couple constituting the base unit of the family. Both datasets are anonymous without unique identifiers for data security reasons, so we must create a new identifier on which we can match the observations. This identifier is based on the man or woman self-identifying as the leader of the household and the spouse of the leader of the household, allowing for the inclusion of non-married couples as well. I excluded those families where the female or the male adult is missing from the analysis, so single-parent households are not considered. This choice is by necessity as I could not find these households in the 2011 Census using only one person's information. However, a rich set of control variables from 2011 are available that can be considered unaffected by the policy, which allows for mitigating some of the selection bias originating from observing only couples in the sample.

Furthermore, I restricted the sample to those families where the mother's birth year was between 1966 and 1978. These cohorts include families where the mother is past fertile age (forty-five at the introduction of the policy) to be able to check for pre-trends. At the same time, it also includes additional families who are not yet at the end of fecundity in 2016 (1977 and 1978). The target cohort in this analysis is the cohort born in 1975 and 1976, so forty years old at the end of our observation window. I will identify the treatment effects regarding completed fertility at forty for this cohort as they were young enough at the start of the policy to adjust their completed fertility (thirty-five) but also observed at the end of their cycle.[7]

[7]Within all births, the share of children born to mothers aged 40 or more was around 2.6% in 2010, and

Finally, I had to eliminate families which are not uniquely identified by the birth dates of the member of the couple. Note that this matching strategy is conservative as it does not utilize stochastic or more advanced algorithms such as Abramitzky et al. (2020), keeping the data linking as exogenous in that regard as possible. However, other than data errors, we can be quite confident that the families found in both Census datasets are the same. Table 1.1 shows the number of family units after the different steps of sample restrictions: we end up with around 54,000 families (representing about 680,000 families) in the 2016 Microcensus that I searched for in the set of about 597,000 families in the 2011 Census.

Table 1.1: Analysis sample restrictions in the 2016 Microcensus

Subsample	Families in 2016 Microcensus	In 2011 Census
Total	340,390	4,137,439
Mother born between 1966-1978	67,318	826,552
Both adults present	54,398	664,330
Without duplicates according to birth dates	54,157	596,820
Matched within 2016 Microcensus	34,409	
Without matching errors	31,112	
Without outliers (7+ children in 2016)	31,040	

Note: The table reports the sample sizes after excluding observations not relevant to the study and the additional sample restrictions of the 2016 Microcensus based on the success of matching to the 2011 Census.

Figure 1.3 displays the matching success rates by cohorts (left panel), along with the estimated statistically significant coefficients from regressing successful matching against relevant explanatory variables in the 2016 sample (right panel). We can see that of the 2016 sample, 58% could be matched successfully, ranging from 56-59% in different cohorts, so the variance is relatively small across cohorts. Table 1.1 shows the raw number of observations participating in the matching process and the post-matching data cleaning, which eliminates potentially erroneous observations and outliers, after which we end up with a final number of 31,040 families.[8]

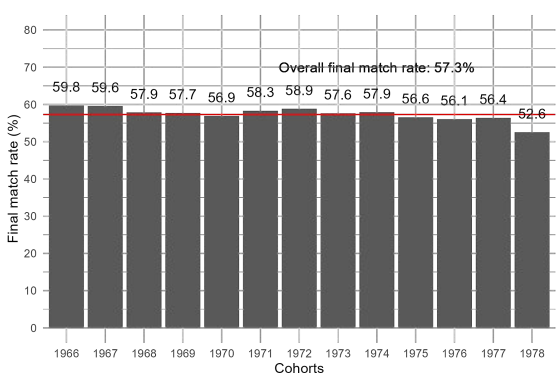
within third births or higher around 7.2%, according to the Central Statistical Office in Hungary.

[8]For reference, this match rate is higher than what was reported using stochastic matching on individuals across historical sources by Abramitzky et al. (2020) with 15% for the United States 1850-1880, and 24% for Norway 1865-1900. However, it is important to stress that those setups aim to connect older sources at the individual level, while in my case, I match on couples in recent sources, which is expected to yield much better results.

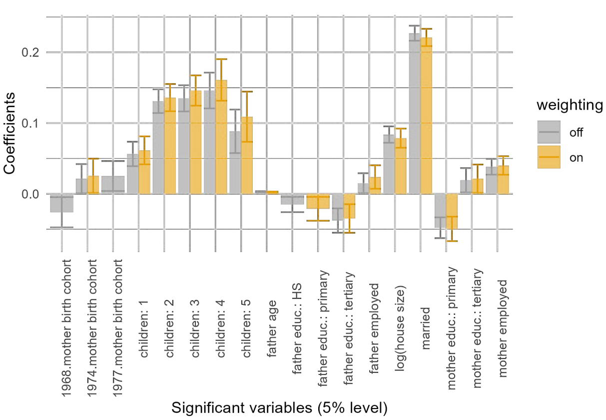
Linking families existing in 2016 to themselves in 2011 requires that the family unit was observed in both censuses. It is impossible if the family unit did not exist earlier, or it was not enumerated because of not being present in the country due to emigration, or they were not cohabiting. The right panel of Figure 1.3 sheds light on which families were matched more successfully, reporting the coefficients of regressing matching success against key observable characteristics.[9] The figure shows the statistically significant predictors at the 5% level, suggesting that the most crucial factor in finding a family is marital status: married couples given other predictors are around 25% points easier to be matched. Higher maternal education and lower paternal education (ref. level: vocational), at least two children, maternal and paternal employment, and younger fathers correlate positively with successful matching. This should come as no surprise: the families described by this set of conditions are the ones that might be more stable. This exercise suggests that when we interpret the findings, we should keep in mind that validity might be limited to those couples that are intrinsically longer-lasting.

Figure 1.3: Matching the 2016 Microcensus and the 2011 Census

(a) Matching success rate by maternal birth cohorts



(b) Predictors of matching success



Note: based on the matched dataset of the 2016 Microcensus and the 2011 Census of Hungary, weighted by the 2016 Microcensus weights. On the right-hand side, we can see linking probability in the 2016 sample regressed against important explanatory variables: mother's birth cohort interacted with the number of children, marital status, parental employment status, imputed parental log(salaries), maternal education, and father's birth year. Variables displayed are significant at the 5% level.

[9] In the regression, I included the following predictors: mother's birth cohort, number of children, marital status, parental employment status, imputed parental log(salaries), parental education, parental ethnicity, father's birth year, and home size.

1.3.2 Prospective salaries

To create the treatment variables, I attached to each member of the couple the prospective gross salary estimated using the National Wage Survey data of Hungary, which collects salary information on the private and public sectors every year since the 1990s. Most individuals in the 2011 Census have a rich set of work-related observable characteristics available, enabling a fine distribution of imputed salary income at the household level. Prospective income proxies the life-cycle earnings of the household members, which in the fertility decisions related to the tax break should play a more instrumental role than current earnings. The tax benefit provides additional income for a large portion of the life cycle, as long as the family receives family allowance (baseline child benefit) for the child, who can be 20 years of age as a maximum (Makay, 2015). Fertility choices should depend on future expected labor market choices Adda et al. (2017) or uncertainty Sommer (2016). This imputation assumes that households perceive the average salaries for the prime earners around them as informative of their own future, meaning that those internalize the career consequences of children as well. Another advantage is that personal unobserved earning capabilities (potentially correlating with fertility decisions) or temporarily higher individual-level salaries within the same subset of people cannot play a role in inducing bias in the estimates. Additionally, currently-unemployed individuals also receive a prospective salary based on their observable characteristics. I imputed salaries for 2011 by pooling the 2010 and 2011 National Wage Survey data; similarly, for 2016, I pooled data from 2015 and 2016. I use the subsample of prime-earning age individuals (30-49 year-olds) with three sets of variables:

1. Gender, region, municipality type, education (4 levels), occupation (10 categories), industry (1-digit NACE categories)

2. Gender, region, municipality type, education (4 levels), occupation (10 categories)

3. Gender, region, municipality type, education (4 levels)

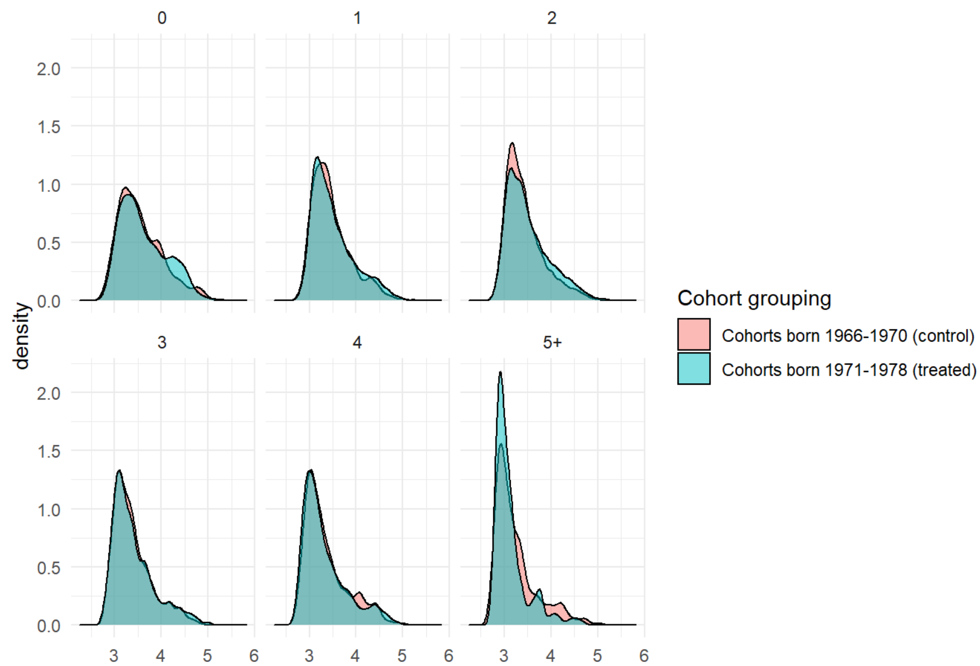
The salary match rates corresponding to the 2016 Microcensus are displayed in Table 1.2 (the rates for the 2011 Census are similar). We can see that more than 95% of the relevant sample could be matched at least using occupation information. In contrast, gender, region, municipality type, and education information are available for everyone in the sample. This ensures that the imputed values for the individuals and the households are sufficiently fine and varied in their values. Figure 1.4 illustrates that the imputation of prospective incomes in 2011 for the male and female adults in the household results in smooth household-level gross income densities. Additionally, these distributions do not differ drastically between the cohorts at least 40 years old in 2011 and those not yet 40 in 2011 (incomes are somewhat higher for younger cohorts). We can also spot that this holds even after split by the number of children and that households with more children are, on average poorer.

Table 1.2: Salary imputation for the 2016 Microcensus sample, number of families

Imputation variables	Mother	Mother%	Father	Father%
Full set of variables	40,785	75.31%	44,796	82.72%
Without industry, but with occupation	11,148	20.58%	7,894	14.58%
Without industry and occupation	2,224	4.10%	1,467	2.70%
Total	54,157	100.00%	54,157	100.00%

Note: The table reports the success rate of joining average salaries along different sets of imputation variables for the mothers, and the fathers, for the 2016 Microcensus.

Figure 1.4: Kernel density of the log of imputed prospective household income by cohort groups and the number of children in 2011



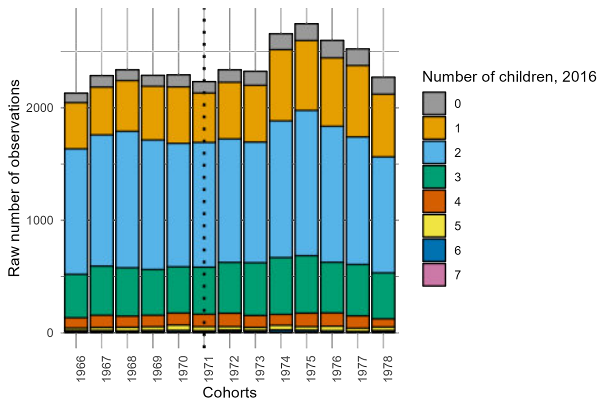
Note: the figure shows the estimated kernel densities of log household incomes (measured in 10,000 HUF units) by cohort groups and the number of children, based on the matched dataset of the 2016 Microcensus and the 2011 Census of Hungary.

1.3.3 Descriptive statistics

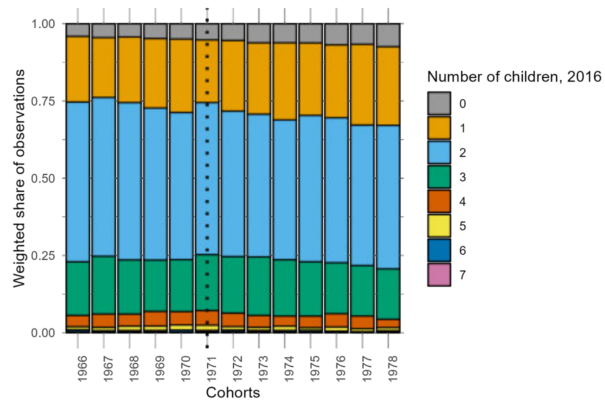
This section gives an overview of the final analysis sample. Figure 1.5 displays the unweighted and weighted number of observations and their distribution by the number of children in 2016. The observation window spans 38-50 years for the relevant cohorts, suggesting that we can treat all except for the 1977 and 1978 cohorts at the end of their fertility cycle in 2016. We can see that the weights do not substantially alter the sample's composition. There is also a steady increase in the share of childless couples, which has continued for younger cohorts not studied in this paper as well, as documented by Spéder (2021). Otherwise, the distribution seems to be similar for the other birth ranks.

Figure 1.5: Observations in the analysis sample by the number of children in 2016

(a) Unweighted number of observations



(b) Shares using the 2016 Microcensus weights

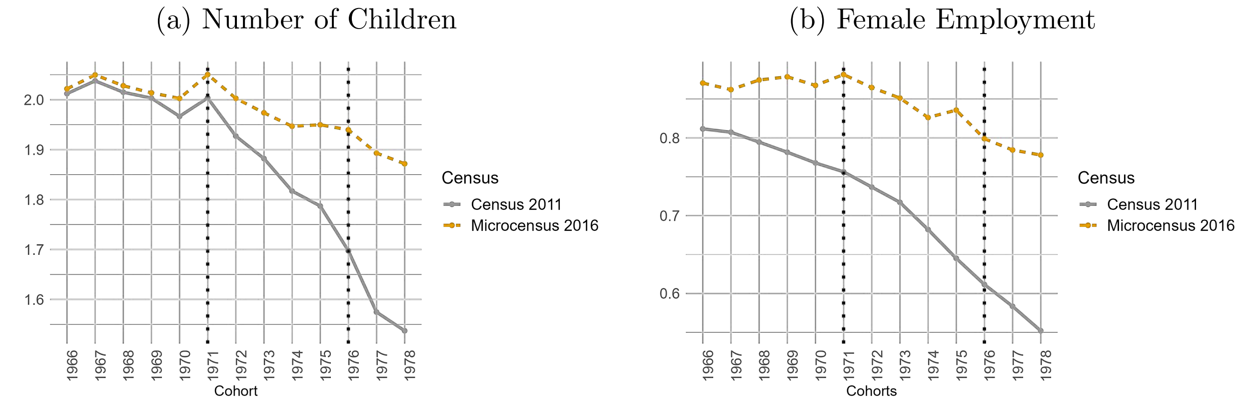


Note: the number of families in the sample, based on the matched dataset of the 2016 Microcensus and the 2011 Census of Hungary.

I am targeting two outcome variables in this paper: primarily completed fertility and secondarily maternal employment. Our focus is on the fertility effects of the policy; however, as fertility is jointly determined by labor supply and housing choices, we might be able to see indirect effects on employment as well. I also inspect the impact on paternal employment, imputed salaries, and house size as validation exercises, which are displayed in the Appendix.

Figure 1.6 shows the averages of the key variables in the weighted matched sample by maternal birth cohorts in 2011 and 2016. We can see that even for older cohorts, there is a slight increase in the number of children between the two Censuses, but it is clear that younger cohorts are not yet at their completed fertility. On average, around 0.25 additional children were born during these five years, around 13% of the total. Maternal employment rates also suggest that the 1971 cohort already transitioned back to the labor market in 2016, as their employment rate matches the ones of older cohorts, while in 2011, this was not the case. These figures support that the 1966-1971 birth cohorts can be regarded as similar in terms of completed fertility behavior and return to the labor market.

Figure 1.6: Outcome variables in the analysis sample

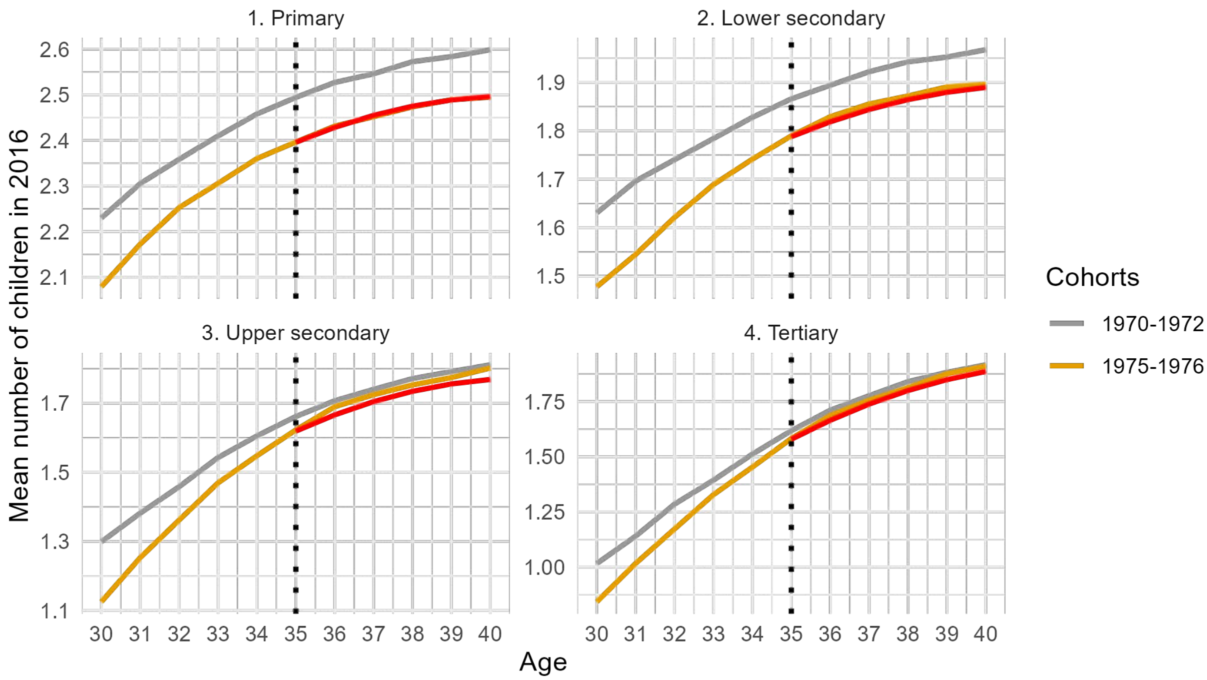


Note: based on the matched dataset of the 2016 Microcensus and the 2011 Census of Hungary.

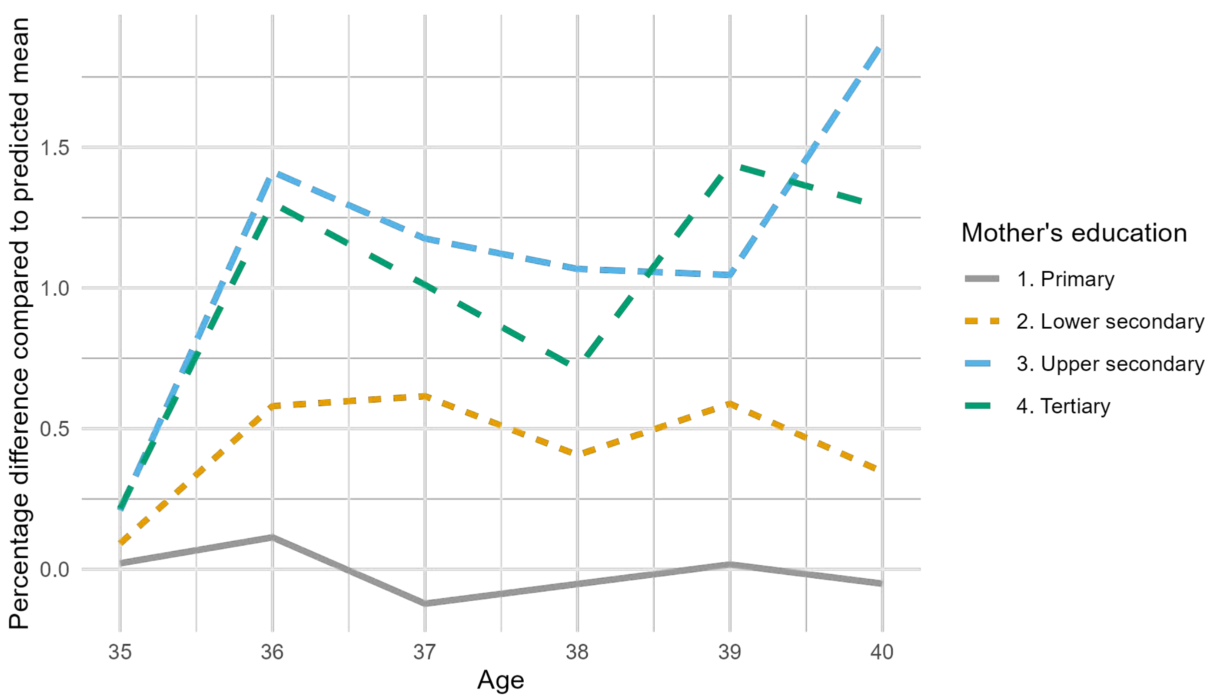
Finally, let us take a look at the evolution of the number of children between ages 30 and 40 by the education of the mother in Figure 1.7a, for the 1975-1976 cohorts vs. the 1970-72 cohorts. I also display with red how many children the younger cohorts would have if the fertility trends of the older cohorts persisted. While for women with primary and lower secondary education, the number of children shifts downwards uniformly at all ages, the initial gap at 30 for women with at least upper secondary education seems to almost close by age forty. Regarding the predicted number of children for the treated cohorts, we can see that for primary-educated mothers, the prediction is almost identical to the realized number, while for the other groups, realized fertility is slightly higher than the forecast. The extent of this difference is displayed in Figure 1.7b, showing that higher-educated women had more children later in their lives in the treated cohorts compared to the prediction by around 1.5%.

Figure 1.7: Mean number of children in the analysis sample by mother's education

(a) Evolution of the mean number of children



(b) Percentage difference compared to prediction



Note: based on the matched dataset of the 2016 Microcensus and the 2011 Census of Hungary. The family tax break policy's extension is marked at age 35. Predictions displayed with red for the 1975-1976 cohorts are based on the education and age of the 1970-1972 cohorts.

1.4 Empirical strategy

1.4.1 Identification

This study relies on the 2011 family tax break extension in Hungary, which introduced large tax base deductions as a function of the number of children and gross household income. The policy introduced non-linearities as shown earlier (Figure 1.2), allowing us to examine the treatment effects at the interaction of household income and the number of children. Due to the kinks in the policy structure, we can control linearly for both income and the number of children, as additional variation remains at the interaction as the tax break is non-refundable and eligibility is complete.

Still, an important obstacle is not yet eliminated in identifying the treatment effects. Even after including these controls, families that receive more treatment might have ended up with different completed fertility regardless. For instance, a poorer family with two children might have ended up with two children, while a richer family receiving more incentives might have ended up with three children even without the additional tax breaks (for instance, due to systematic differential preferences about the ideal number of children). Therefore a single cohort does not plausibly provide a valid counterfactual for different treatments, even if a rich set of pre-policy covariates is available from the 2011 Census. For this reason, to identify the treatment effect of the policy, it is necessary to construct an adequate control group. This group should not react to the policy, while on average, it should not differ in observed or unobserved characteristics and would have reacted to the policy on average the same way as the treated group did.

Therefore, I use both within-cohort and between-cohort variation to identify the treatment effects and argue that using the remaining variation can be considered as good as random. To eliminate endogeneity concerns, I interact the treatments with cohort indicators to remove any additional unobserved heterogeneity as a source for additional children. I argue that a valid counterfactual can be constructed using cohorts just past fertile age, rendering them unable to react to the policy. To those cohorts, the received family tax break acts as a placebo regarding fertility: even though a similar treatment is received, the researcher knows ex-ante that no effect can be expected. These cohorts are close in age and experienced a similar historical and sociological context. It is reasonable to assume that their reaction to the treatment would have been identical on average to the younger cohorts (conditional on several observable characteristics). Evidence indicates that fertility preferences remained stable throughout the years before the policy (Kapitány and Spéder, 2015), providing further support for these assumptions. Note that it is not required that the level of observed and unobserved heterogeneity is the same for each cohort. However, their reaction to the policy would have to be equal on average.

Let us denote by $Y_{i,c}(T, X)$ the potential outcome of family i in cohort $c \in \mathcal{C} = \{1966, \ldots, 1978\}$, as a function of treatment $T_i \in \mathbb{R}^+$ and controls X_i. Then the required assumption can be stated as follows:

$$\mathbb{E}\left[\frac{\partial Y_{i,c}(T,X)}{\partial T}\middle| X = X_i, c = 1966, T = \tau\right] = \mathbb{E}\left[\frac{\partial Y_{i,c}(T,X)}{\partial T}\middle| X = X_i, c = \zeta, T = \tau\right], \forall(\zeta,\tau) \in \mathcal{C} \times \mathbb{R}^+$$

meaning that, on average, the reaction of the potential outcome to the treatment would be the same for each cohort at each possible level of treatment.

Suppose that the following linear equation governs the outcome, and let us also assume that unobserved heterogeneity denoted by $\eta_{i,c}$ is additively separable. Then for cohort c the expected outcome can be written as:

$$\mathbb{E}_i[Y_{i,c}|X_i, T_i] = \alpha_c + \theta_c T_i + \gamma X_i + \mathbb{E}_i[\eta_{i,c}|X_i, T_i]$$

, meaning that the conditional expectation of the outcome $Y_{i,c}$ in cohort c is determined by a cohort-level intercept, the cohort-level reaction to the treatment, potentially some other covariates to control for selection into the sample, and the conditional expectation of the unobserved heterogeneity.

It is easy to imagine that $\mathbb{E}_i[\eta_{i,c}|X_i, T_i] \neq 0$, where a plausible illustration might be the following. Suppose that families within a cohort can be treated or untreated by the tax break and unobserved heterogeneity η represents the extra births from families with preferences for many children. If the extra births from such families are different in the treated group than they would have been in the control group, the condition fails. That could occur in either direction: if the share of families with preferences for many children is higher in the treated group than in the control group, then we would overestimate the effect of the policy by simply regressing the outcome on the treatment and other covariates.

Therefore we need additional help with identification. Due to biological reasons, we know that θ_{1966} (so the average treatment effect on the treated for cohort 1966) must be 0, as they were 45 years old at the point of implementation. However, as mentioned before, it is reasonable to assume that they would have reacted similarly to the treatment on average as those cohorts that were still fertile, meaning that regarding their unobserved heterogeneity, we assume that

$$\frac{\partial \mathbb{E}[\eta_{i,c}|X_i, T_i]}{\partial T_i} = \Delta, \forall c \in \mathcal{C},$$

, the parallel trends assumption in this context.

Given these two assumptions, we can identify the (Conditional) Average Treatment Effect on the Treated in the following way:

$$(C)ATET = \frac{\partial \mathbb{E}[Y_{i,c}|X_i, T_i]}{\partial T_i} - \frac{\partial \mathbb{E}[Y_{i,1966}|X_i, T_i]}{\partial T_i} = (\theta_c + \Delta) - (\theta_{1966} + \Delta) = \theta_c$$

which can be estimated using regression difference-in-differences by interacting cohorts and treatments. The estimation also includes a validation (placebo) exercise to test pre-trends. If our exclusion restriction is valid, we should not see higher or lower completed fertility corresponding to higher treatments in the older, infertile cohorts. This exercise (especially without including any control variables) validates that, on average, differential selection regarding the treatment sizes and cohorts does not strongly affect the estimates. As even the 1970-72 cohorts should not respond to the policy, instead of the 1966 cohort, I select those as the reference category for the regressions and use the other cohorts to test pre-trends while also grouping yearly cohorts to biannual ones to increase statistical power. The estimated regression equations then have the following form, including cohort-level treatments and intercepts, individual controls for income, the initial number of children, and some additional controls, so that identification relies on the non-linear structure of the policy:

$$Y_i = \sum_{c \in \mathcal{C}} \theta_c \mathbb{I}[i \in c] \times T_i + \sum_{c \in \mathcal{C}} \alpha_c \mathbb{I}[i \in c] + \Big(\beta \text{Inc}_i + \sum_{N \in \mathcal{N}} \gamma_N \mathbb{I}[N_i = N]\Big) + \delta X_i + \varepsilon_i \qquad (1.1)$$

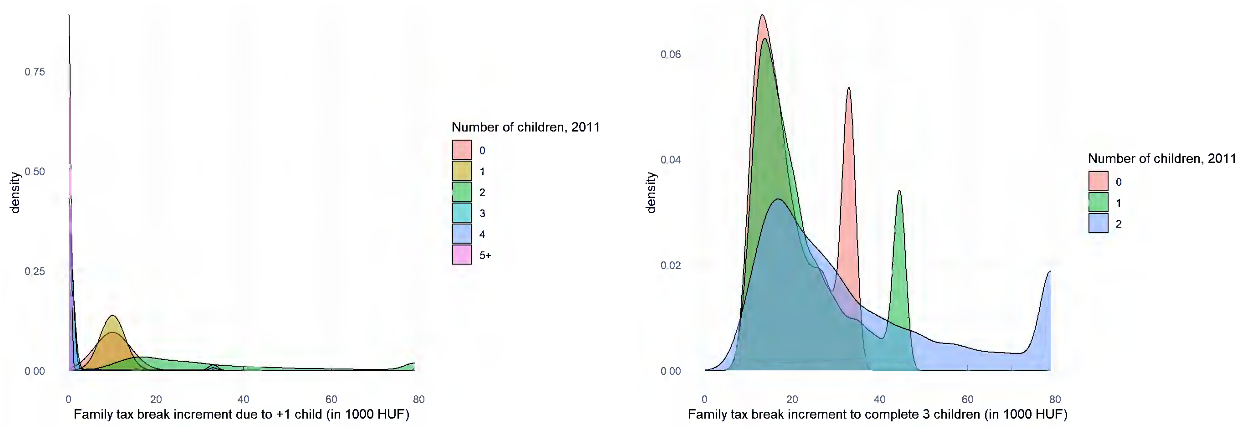
where for individual i in a biannual cohort c, $Y_{i,c}$ denotes the outcome, T_i denotes the amount of family tax break received as additional net household income, Inc_i the prospective gross household salary in 2011, N_i the number of children in 2011, and X_i the additional controls. In the main specification of this version of the paper, $Y_i = log(N_{i,2016} + 1)$ is used as the primary outcome variable.

1.4.2 Treatment definitions

The policy creates two types of changes: on the one hand, it induces differential financial incentives to have one additional child for families with at least two children, and on the other hand, it alters incentives for families with fewer than two children to end up with three children. The more straightforward treatment variable is the family tax break increment for one additional child, measuring directly the effect of lowering the relative cost of one additional child. In this setting, the identifying variation comes from the tax break kink at the third or higher birth parities. Therefore I attempt to capture the increment effect at lower birth parities (zero and one child) by calculating the additional family tax break per child received after the completion of three children. This approach incorporates the case when a family decides to increase the number of children from zero or one to three instead of two as a response to the policy. Figure 1.8 displays the estimated kernel densities for the income a household could get for one additional child (left panel) and the income a household could get per child if they reach three children (right panel). We can see in the left panel that there is minimal variation in the amount at lower parities, and similarly at the highest parities where only a few families (the richest ones) could claim more money for having additional children. The right panel displays that the family tax break increment follows the shape of household income densities by construction; however, due to the non-linearities in the structure of the policy, there is remaining residual variation even after controlling for the number of children and income.

Figure 1.8: Kernel densities of the family tax break variables

(a) Kernel density of the tax break due to one more child (b) Kernel density of the tax break after reaching three children



Note: the figures show estimated kernel densities of the two treatment variables: additional family tax break from one additional child and additional family tax break per child after reaching three children, based on the matched data of the 2016 Microcensus and the 2011 Census of Hungary.

Specifically, let us then define the treatment $T_i \in \mathbb{R}^+$ for a family i as the extra money per month in Hungarian Forints (HUF) corresponding to the tax break from additional children, based on the prospective gross salary of the household in 2011 ($S_{i,2011}$), the number of children in 2011 ($N_{i,2011}$), and the personal income tax rules of 2011. Then the tax base deduction rule per month is the following:

$$D_i = D(N_i) = \begin{cases} N_i \cdot 62,500 \text{ HUF} \sim N_i \cdot 230 \text{ EUR} & \text{if } N_i \leq 2 \\ N_i \cdot 206,250 \text{ HUF} \sim N_i \cdot 770 \text{ EUR} & \text{if } N_i \geq 3 \end{cases}$$

Based on the deduction, the net amount of tax break money received by the family as a function of the number of children N_i can be defined in the following way:

$$M(N_i, S_i) = \begin{cases} D(N_i) \cdot 0.16 & \text{if } S_i > D(N_i) \\ S_i \cdot 0.16 & \text{if } S_i \leq D(N_i) \end{cases}$$

A key component is that the tax break itself is a tax base deduction that is limited in size by the tax base of the family, the tax base being the sum of the gross salaries of the parents (who are not required to be married).[10] Based on this setup, I define the two types of continuous treatment variables targeting the increment effects, measuring how much additional money the family tax break policy would imply given the prospective household salaries. These are the following:

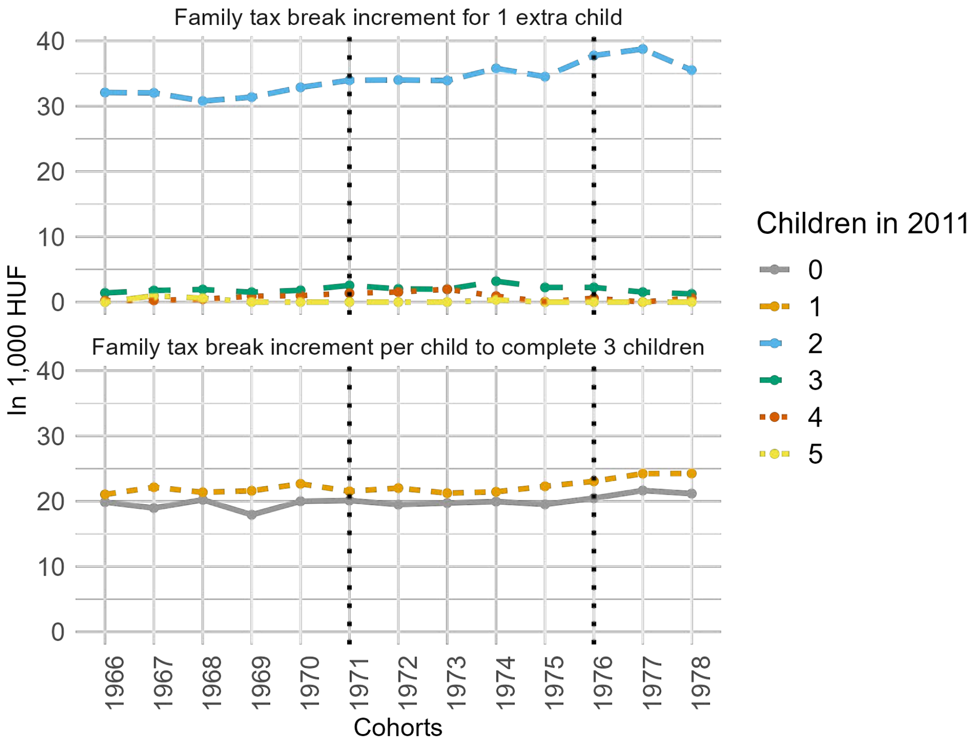
1. $T_{1,i} = M(N_{i,2011}+1, S_{i,2011}) - M(N_{i,2011}, S_{i,2011})$: Additional money in 10,000 Hungarian Forints (HUF) per month for one additional child (examine separately also for families with at least two children, as it varies very little due to the low amount of deduction for zero and one child)

2. $T_{2,i} = M(3, S_{i,2011}) - M(N_{i,2011}, S_{i,2011})$: Additional amount of money in 10,000 HUF per month completing three children compared to the amount received according to the initial number of children

Figure 1.9 shows the average values of the family tax break variables and the imputed prospective household income by maternal birth cohorts and the number of children. By definition, the number of children and income jointly determine the amount of the tax break an individual family would receive. This figure demonstrates that there is no considerable variation in the treatments families received across cohorts, and meaningful variation is expected to occur at the third and fourth birth parity.

[10] Rules of taxation over the period has changed slightly, after a significant overhaul in 2011 starting from 16% rate, which decreased to 15% in 2016.

Figure 1.9: Monthly average family tax break measures by cohorts and the number of children



Note: The figures show average values of family tax break measures and imputed household income by cohorts and the number of children in the matched sample, based on the 2016 Microcensus and the 2011 Census of Hungary.

Based on the treatment variable and an external threshold, we can introduce a discrete version of the treatment where we deem the treatment value 'large enough'. Poverty line estimates by the Hungarian Statistical Office indicate that the poverty line cost per child was around 20,000-25,000 HUF (HCSO, 2015). I use this value to define a discrete treatment such that if the continuous value per additional child exceeds this level, the family is considered treated. This setup requires less stringent identification assumptions in a difference-in-difference estimation compared to the continuous version (Callaway et al., 2021). I use the discrete treatment results as additional robustness checks to the main estimates, but we can conceptualize the treatment as the state financing a child's most basic needs.

1.5 Results

1.5.1 Completed fertility

I examine completed fertility (measured by the logarithm of the number of children in 2016 + 1)[11] and maternal employment as relevant outcomes of the policy.[12] Given the earlier assumptions, the coefficient on the interaction term, $\hat{\theta}_c$ estimates the conditional average treatment effect on the treated for cohort c. I employ a rich set of control variables from the 2011 Census about the family's education, employment, health, and living conditions, to eliminate any potential selection bias or composition change along these variables. [13] There are two treatment variables measured in units of 10,000 HUFs per month (~40EUR): the additional money receivable from one more child for families with at least two children, and the extra money receivable after completing three children for families with fewer than two children. Accordingly, the sample is split into subsamples by whether the mother had at least two children at age 35, the age of the 1976 cohort in 2011. The goal was to include those older families in the counterfactual group who are more similar to the fertile age cohorts, but the results are robust to this choice.

I display three specifications of the regressions with different sets of controls. In the first one, no additional variables besides the treatment and the base amount are included, so we regress the outcomes on the cohort group dummies, the treatment variables, and their interactions. This exercise aims to validate whether the estimates for the older cohorts are not different statistically from 0, providing evidence that differential sample selection along the treatment is not a problem. In the second specification, I include prospective household income and the number of children in 2011, as I argue that along these dimensions, we can expect quasi-randomness to identify the effects due to the non-linearities of the policy. Finally, in the third one, I add the rich set of observable characteristics in 2011 to control

[11]The results are robust to count-based regression specifications.

[12]Estimates for paternal employment, imputed salaries, and house size can be found in the Appendix, while regression tables are available at the author.

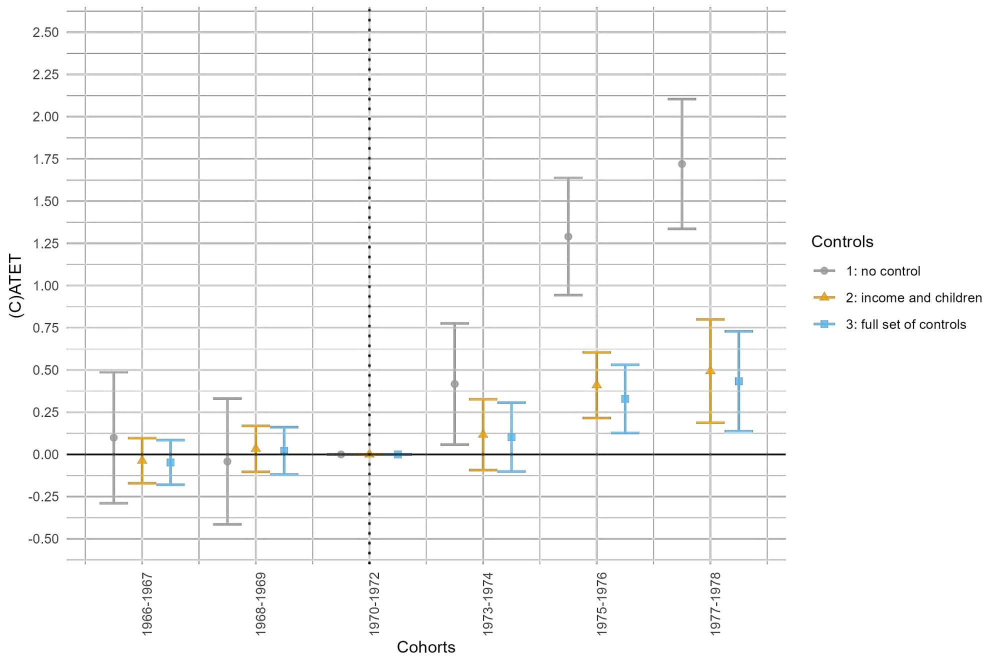
[13]The set of controls includes the following: number and gender composition of children; household income; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); location; ethnicity; illnesses and disabilities; foreign language knowledge; religiosity; wealth, pension and social care status; public employment status; and the type of income imputation that was used.

for any concerns regarding changes in composition. For each cohort, I display the point estimates of the interaction term with the 95% confidence interval bands calculated using cluster-robust standard errors on the sub-county administrative level.[14] For each cohort group, we can see three estimates, each corresponding to one of the specifications described previously. The estimates of the third specification should be interpreted causally under the assumptions we made earlier, giving us the policy's conditional average treatment effects on the treated. The 1975-1976 cohort is the primary subject as they are the ones that reach around completed fertility by the end of the time window while still having enough time to react to the incentives and adjust their completed fertility meaningfully. The following figures present the main results, I also report them in tables along with descriptive statistics in Appendix Table 1A2 (for count outcomes, see Table 1A3).

Figure 1.10 displays the findings for families with at least two children. We can see that without additional controls, the point estimate for the 1975-1976 cohorts is around 1.25 per 10,000 HUF, decreasing to 0.33 after including income and children. At the mean treatment size of 2.29, the policy is implied to have around 0.8% completed fertility effect at these birth parities. There are no statistically significant effects for the 1973-1974 cohorts, while for the younger 1977-1978, we identify a slightly higher effect (while they have not yet completed their fertility at age thirty-eight). Furthermore, the non-significant point estimates for the older cohorts without including additional control variables indicate that the assumption regarding unconfoundedness is supported, and we could interpret the estimates for the fertile cohorts causally.

[14] There are 198 of these units in Hungary, each containing around 50,000 people on average, and they are often used to approximate local labor markets.

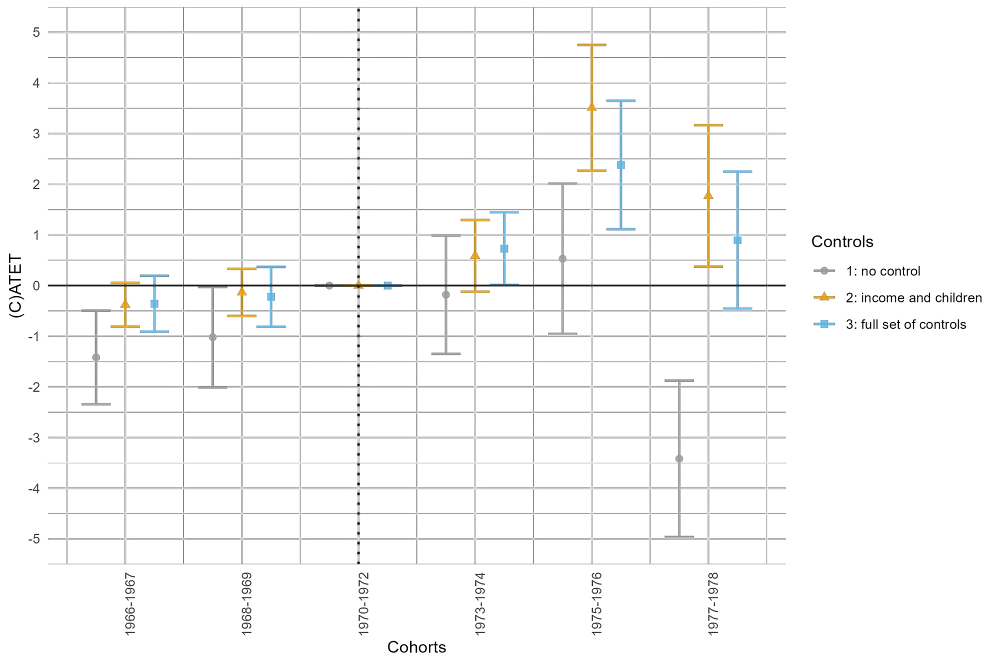
Figure 1.10: Effect of 10,000 HUF family tax break on completed fertility (in %), families with at least two children



Note: based on regressing the log(1+number of children) in 2016 on cohort groups interacted with treatments, with a baseline of 1970-1972. The full set of controls includes the following from 2011 besides household income and number of children: gender composition of children; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); location; ethnicity; illnesses and disabilities; foreign language knowledge; religiosity; wealth, pension and social care status; public employment status; and the type of income imputation. 95% confidence intervals are shown, standard errors clustered on the sub-county administrative unit ('járás') level. The sample includes families with at least two children by the mother's age of 35, with a sample size of 20,721. The mean treatment size is 2.29.

The results for the families with fewer than two children are reported by Figure 1.11. While we see non-significant effects without controls and some evidence for pre-trends for the older cohorts, after adding income and children as controls, pre-trends turn insignificant. In contrast, statistically significant effects are revealed for the 1975-1976 cohort. With a complete set of controls, an effect of 2.4% per 10,000 HUF is estimated, which at the mean treatment of 2.37 implies a 5.6% effect at the lower birth ranks. For the other female birth cohort groups, we cannot identify a treatment effect.

Figure 1.11: Effect of 10,000 HUF family tax break on completed fertility (in %), families with fewer than two children



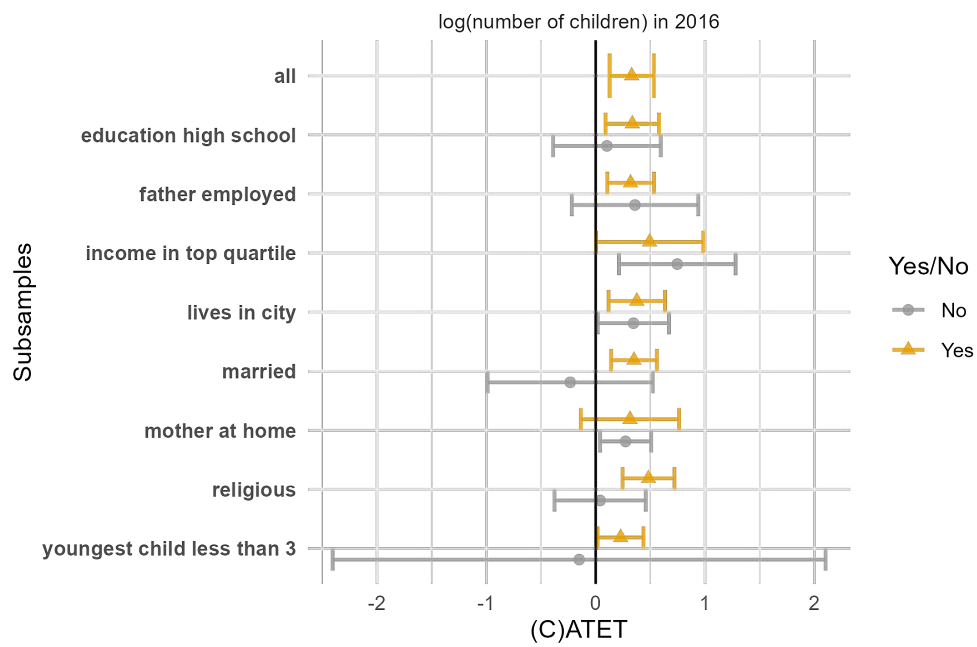
Note: based on regressing the log(1+number of children) in 2016 on cohort groups interacted with treatments, with a baseline of 1970-1972. The full set of controls includes the following from 2011 besides household income and number of children: gender composition of children; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); location; ethnicity; illnesses and disabilities; foreign language knowledge; religiosity; wealth, pension and social care status; public employment status; and the type of income imputation. 95% confidence intervals are shown, standard errors clustered on the sub-county administrative unit ('járás') level. The sample includes families with fewer than two children by the mother's age of 35, with a sample size of 10,243. The mean treatment size is 2.37.

The heterogeneity of the results reveals which segments of the population are more responsive to financial incentives. To address this question, I run the same regressions for several relevant subgroups[15], all estimates can be found in Appendix Tables 1A4-1A7. In Figures 1.12 and 1.13, I display the estimates for the splits of the 1975-1976 cohorts. Treatment effects are higher for families with at least two children, being married, religious, and having the youngest child older than three. Although the confidence intervals overlap in all cases,

[15] Relevant subgroups are selected based on the literature and created based on the 2011 Census information. 'Education high school' indicates completed high school education for the mother. Father employment indicates the employed status of the male parent the week previous to the Census survey with the ideal date of 01/10/2011. 'Income in top quartile' indicates households in the top 25% of the gross prospective income distribution. 'Lives in a city' indicates that the family lives in a city or a town. 'Mother at home' indicates that the mother is unemployed or at home with child allowances or benefits. Religiousness indicates that both parents report belonging to a religious group, in the context of Hungary, most likely to Christian churches. Married indicates a married couple, while 'youngest child less than 3' indicates that the youngest child in the family is younger than three years old.

the confidence bounds are above the point estimate of their counterpart. One-sided Z-tests[16] suggest that religiousness is significant at the 5% level, while marital status is at the 10% level.

Figure 1.12: Effect of 10,000 HUF family tax break on completed fertility for the 1975-1976 Cohorts (in %), families with at least two children



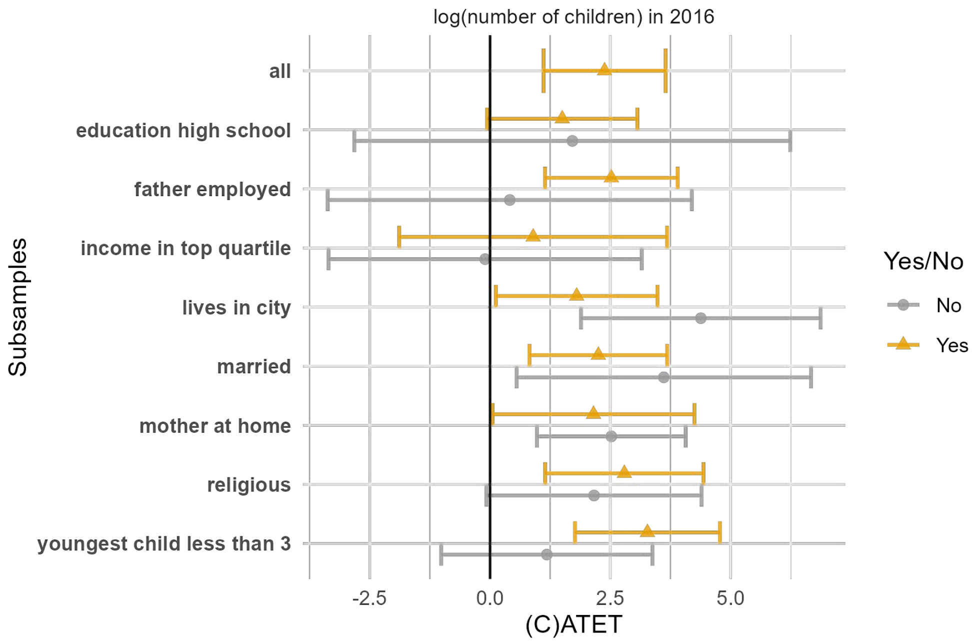
Note: based on regressing the log(1+number of children) in 2016 on cohort groups interacted with treatments for the 1975-1976 cohort, with a baseline of 1970-1972 for different subsamples. The full set of controls includes the following from 2011 besides household income and number of children: gender composition of children; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); location; ethnicity; illnesses and disabilities; foreign language knowledge; religiosity; wealth, pension and social care status; public employment status; and the type of income imputation. 95% confidence intervals are shown, standard errors clustered on the sub-county administrative unit ('járás') level. The sample includes families with at least two children by the mother's age of 35. Sample and treatment sizes vary between subsamples.

The results paint a different picture for families with fewer than two children (Figure 1.13). The youngest child being less than three years old along with living in a village seems to increase the effect size of the family tax break (one-sided Z-test indicating significance at the 5% and 10% levels), while no other variables alter the effect of the policy significantly. However, despite the non-significant difference along paternal employment statuses, families with an unemployed father did not experience an increase in fertility due to the policy, which

[16]Calculated as $Z = \frac{\beta_2 - \beta_1}{\sqrt{SE^2(\beta_1) + SE^2(\beta_2)}}$, similarly to seemingly unrelated regressions.

is in contrast with the results for families with at least two children.

Figure 1.13: Effect of 10,000 HUF family tax break on completed fertility for the 1975-1976 Cohorts (in %), families with at least two children



Note: based on regressing the log(1+number of children) in 2016 on cohort groups interacted with treatments for the 1975-1976 cohort, with a baseline of 1970-1972 for different subsamples. The full set of controls includes the following from 2011 besides household income and number of children: gender composition of children; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); location; ethnicity; illnesses and disabilities; foreign language knowledge; religiosity; wealth, pension and social care status; public employment status; and the type of income imputation. 95% confidence intervals are shown, standard errors clustered on the sub-county administrative unit ('járás') level. The sample includes families with fewer than two children by the mother's age of 35. Sample and treatment sizes vary between subsamples.

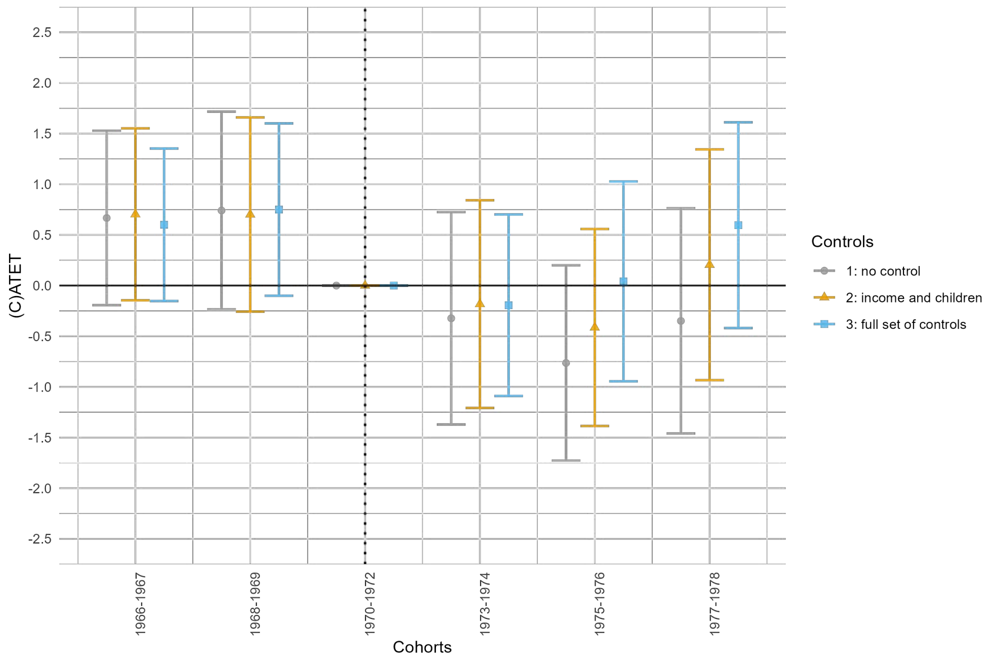
1.5.2 Mother's employment

The family tax break introduces large changes in the incentives for mothers to participate in the labor market, both for those who can and cannot have additional children. The potential effects are ambiguous: while there are higher returns from working more and increasing the tax base of the household for poorer families, the higher initial household income due to the initial or additional number of children provides counter-incentives via income effects such that a woman might leave or not reenter the labor market after giving birth, especially if her labor income is not high to start with. We can differentiate between the direct effect on labor supply which affects labor supply choice conditional on the number of children, and the indirect effect via having additional children. Additionally, there are long-run equilibrium effects after sufficient time that women would like to return to the labor market compared to short-run effects about the period for transitioning back to the labor market.

With the empirical strategy of the paper, we can only identify the short-run indirect effects of the policy. First, we can only observe the immediate future period after the policy is introduced. Second, we have to use older female birth cohorts as benchmarks that could also adjust their labor supply conditional on the number of children they have, so differencing out their adjustment would only let us identify the changes in employment due to the ability to have additional children. Therefore the following results reflect a short-run labor supply change. Labor supply decisions are intrinsically related to women's fertility choices due to the high career costs, which also affect their sorting into different occupations or education levels (Adda et al., 2017). I use a simple model simulation exercise later to shed light on the long-run effects in that regard. It is also important to note that the indirect change in the employment status of older cohorts should be unaffected by the policy, as any measured effect should come as a consequence of new children, so we can use the estimates for these older cohorts as validation exercises.

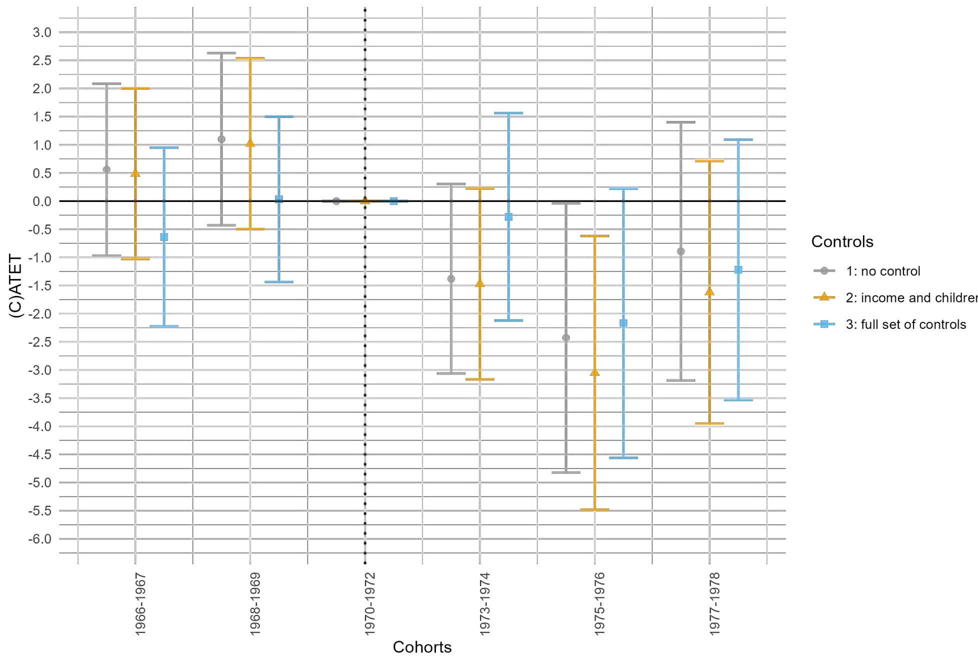
Figures 1.14 and 1.15 present the estimates for families with at least two, and fewer than two children, respectively. We can see that no robust effects can be detected in either case, with some significant negative point estimates for families with fewer than two children if we do not include all the control variables in the regression. These negative effects of around 2.5 percentage points (scaled up to around 5 percentage points at the mean treatment size) are likely to related to not yet being able to return to the labor market after giving birth. Reassuringly, the estimates for older cohorts are not statistically significant either, supporting the previous identifying assumptions.

Figure 1.14: Effect of 10,000 HUF family tax break on maternal employment (in percentage points), families with at least two children



Note: based on regressing the indicator for maternal employment in 2016 on cohort groups interacted with treatments, with a baseline of 1970-1972. The full set of controls includes the following from 2011 besides household income and number of children: gender composition of children; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); location; ethnicity; illnesses and disabilities; foreign language knowledge; religiosity; wealth, pension and social care status; public employment status; and the type of income imputation. 95% confidence intervals are shown, standard errors clustered on the sub-county administrative unit ('járás') level. The sample includes families with at least two children by the mother's age of 35, with a sample size of 20,721. The mean treatment size is 2.29.

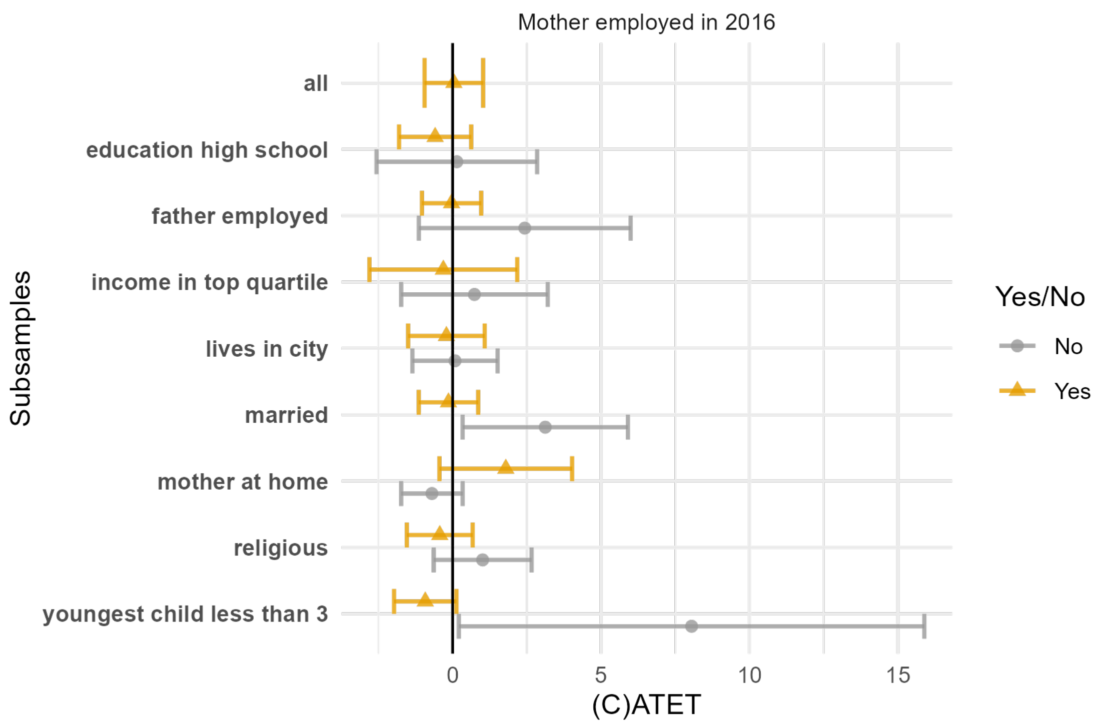
Figure 1.15: Effect of 10,000 HUF family tax break on maternal employment (in percentage points), families with fewer than two children



Note: based on regressing the indicator for maternal employment in 2016 on cohort groups interacted with treatments, with a baseline of 1970-1972. The full set of controls includes the following from 2011 besides household income and number of children: gender composition of children; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); location; ethnicity; illnesses and disabilities; foreign language knowledge; religiosity; wealth, pension and social care status; public employment status; and the type of income imputation. 95% confidence intervals are shown, standard errors clustered on the sub-county administrative unit ('járás') level. The sample includes families with fewer than two children by the mother's age of 35, with a sample size of 10,243. The mean treatment size is 2.37.

Despite not having overall average effects, there might be specific subgroups for whom the policy caused changes in labor market outcomes. The heterogeneity effects of the family tax break on employment are displayed in Figures 1.16 and 1.17, again using the 1970-1972 cohorts as a baseline. For families with at least two children (Figure 1.16), we can detect significantly lower employment for some of the groups where higher completed fertility was also detected earlier. The difference in employment using one-sided Z-tests is significantly negative at the 5% level for married couples, the youngest child is less than three years old. However, we detect some statistically significant positive employment effects at the 5% level for mothers at home in 2011 compared to their counterparts, despite neither being statistically significant one by one. So mothers who are at home receiving higher tax breaks return back to the labor market more compared to mothers who are not at home at the start of the policy.

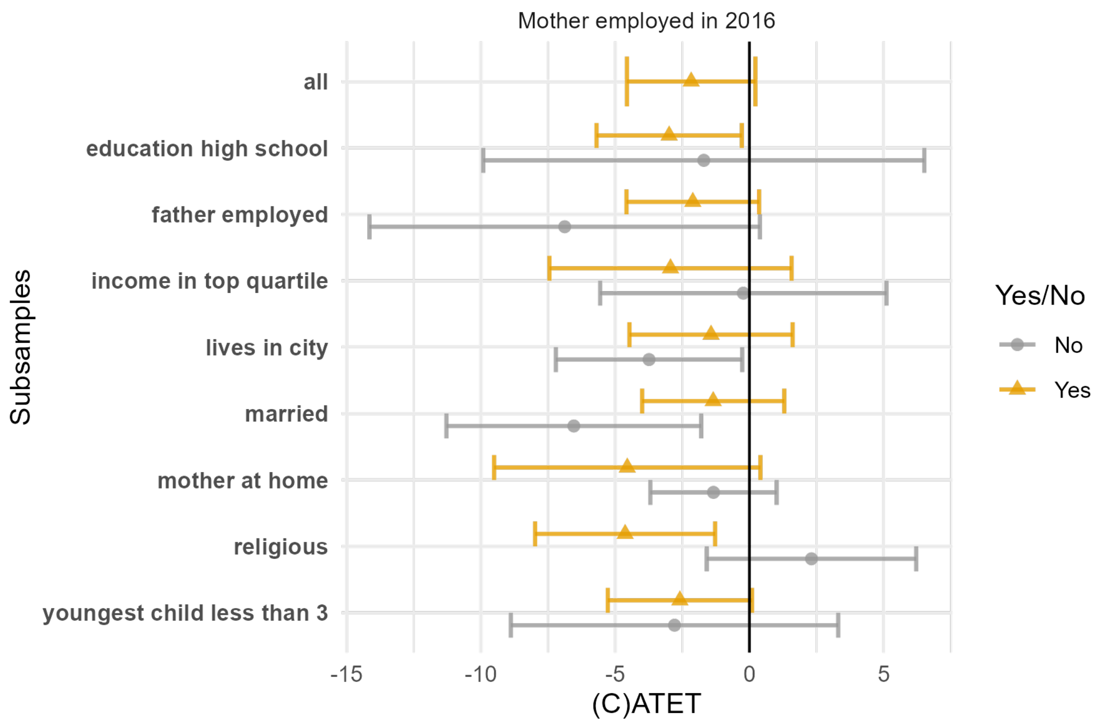
Figure 1.16: Effect of 10,000 HUF family tax break on maternal employment (in percentage points) for 1975-1976 Cohorts, families with at least two children



Note: based on regressing the indicator for maternal employment in 2016 on cohort groups interacted with treatments for the 1975-1976 cohort, with a baseline of 1970-1972 for different subsamples. The full set of controls includes the following from 2011 besides household income and number of children: gender composition of children; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); location; ethnicity; illnesses and disabilities; foreign language knowledge; religiosity; wealth, pension and social care status; public employment status; and the type of income imputation. 95% confidence intervals are shown, standard errors clustered on the sub-county administrative unit ('járás') level. The sample includes families with at least two children by the mother's age of 35. Sample and treatment sizes vary between subsamples.

Regarding families with fewer than two children, Figure 1.17 suggests significant differences along religiosity at the 5% level: for religious households, getting a higher tax break for additional children resulted in significantly lower employment (5 percentage points for 10,000 HUF) compared to their non-religious counterparts, whose point estimate is indeed positive. Additionally, at the 10% level, we can see that married couples are affected less negatively compared to non-married couples regarding employment status.

Figure 1.17: Effect of 10,000 HUF family tax break on maternal employment (in percentage points) for 1975-1976 Cohorts, families with fewer than two children



Note: based on regressing the indicator for maternal employment in 2016 on cohort groups interacted with treatments for the 1975-1976 cohort, with a baseline of 1970-1972 for different subsamples. The full set of controls includes the following from 2011 besides household income and number of children: gender composition of children; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); location; ethnicity; illnesses and disabilities; foreign language knowledge; religiosity; wealth, pension and social care status; public employment status; and the type of income imputation. 95% confidence intervals are shown, standard errors clustered on the sub-county administrative unit ('járás') level. The sample includes families with fewer than two children by the mother's age of 35. Sample and treatment sizes vary between subsamples.

1.6 Robustness of the results

1.6.1 Robustness tests based on the linked analysis sample

I explore several changes in the specification to test the sensitivity of the results. I also display the estimated coefficients for the 1977-1978 cohorts and further corroboration regarding the magnitudes of the main results. In Tables 1.3 and 1.4, the following specifications are presented. The first column shows the main estimates already displayed in the figures earlier, while the second column shows the same regression but without using the 2016 Microcensus population weights. In the third column, I use count outcomes which results should be interpreted in terms of the number of children. In the fourth column, I present results from an alternative specification where the family tax break's baseline amount is not included in the regressions, only the increments for additional children, and sample restrictions are only based on the number of children in 2011 with no regard for the mother's number of children at age 35. And finally, in the last column, I display the results using the discrete treatment definition mentioned earlier, where treatment equals one if its value exceeds the subsistence cost of a child.

First, let us consider the families with at least two children (Table 1.3). The second column containing the unweighted estimates suggests that the Microcensus weights hardly alter the estimates at all. The third column with the count outcome shows a treatment effect of around 0.01 child from an additional 10,000 HUF, at the mean treatment size of approximately 0.023. With the mean number of births at 2.0 of the 1971 cohort, a 1.2% effect size is implied, close to the main specification. The alternative, simpler specification in the fourth column suggests quantitatively similar point estimates, but the sample selection induces a larger mean treatment size of 2.41. Finally, the discrete treatment specification yields around 1.7% higher completed fertility if the treatment is above 25,000 HUF, which is the case for approximately 37% of the sample. That implies an around 0.6% effect of the policy, slightly smaller than the main specification.

Table 1.3: Robustness checks: completed fertility effect estimates for families with at least two children using different specifications

Mother's birth cohort	Main	Unweighted	Count outcome	Alternative spec.	Discrete treatment
1966-1967	-0.000475	-0.000619	-0.00117	-0.000360	0.000669
	(0.000673)	(0.000481)	(0.00265)	(0.000554)	(0.00228)
1968-1969	0.000211	-0.0000102	0.00167	-0.000146	0.00246
	(0.000716)	(0.000556)	(0.00277)	(0.000578)	(0.00273)
1970-1972	0	0	0	0	0
	()	()	()	()	()
1973-1974	0.00103	0.00122	0.00226	0.00145*	0.00273
	(0.00104)	(0.000855)	(0.00404)	(0.000843)	(0.00323)
1975-1976	0.00329***	0.00329***	0.00995**	0.00331***	0.0172***
	(0.00103)	(0.000932)	(0.00392)	(0.000893)	(0.00418)
1977-1978	0.00433***	0.00434***	0.0133**	0.00386***	0.0166***
	(0.00151)	(0.00114)	(0.00575)	(0.00130)	(0.00509)
Observations	20,721	20,721	20,721	21,445	20,377
Adjusted R-squared	0.896	0.906	0.902	0.896	0.901
Mean of outcome	1.2282	1.2342	2.4888	1.2245	1.2282
S.D. of outcome	.1966	.2011	.8013	.1946	.1966
Mean of family tax break increment	2.2892	2.0766	2.2892	2.4147	.3681
S.D. of family tax break increment	2.3148	2.1749	2.3148	2.3718	.4823

Note: The table reports the coefficients and cluster-robust standard errors of the interaction terms of the cohort group dummies and the treatment variable in different regression specifications. Cluster robust standard errors in parentheses. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Let us turn to Table 1.4, which reports the robustness checks for families with fewer than two children. While the estimates for the unweighted regression differ slightly in this case, the differences are within one standard error of the coefficients. With fertility measured in the number of children, the estimate suggests an around 0.07 child increase due to a 10,000 HUF tax break, which scaled up to the 2.37 mean treatment size leads to 0.17 children, a higher effect size of around 8.5%. Finally, using a discrete treatment definition of at least 25,000 HUF per child receivable tax break for the third child, I find about a 10.1% completed fertility effect, which is present at 29.6% of the population, leading to a lower around 3.0% suggested effect size. However, it is worth mentioning that in this case, we cannot reject the null of no pre-trends, as the estimate for the 1966-1967 cohorts is negative and statistically significant, suggesting that the 1969-1972 cohorts might have also reacted to the policy, resulting in an underestimation of the effect size. To conclude, the robustness checks for families with fewer than two children show that these results are more sensitive to specification, however, different magnitudes of statistically significant treatment effects can still be estimated.

Table 1.4: Robustness checks: completed fertility effect estimates for families with fewer than two children using different specifications

Mother's birth cohort	Main	Unweighted	Count outcome	Alternative spec.	Discrete treatment
1966-1967	-0.00358	-0.00156	-0.0106*	-0.00993*	-0.0251**
	(0.00281)	(0.00288)	(0.00642)	(0.00509)	(0.0122)
1968-1969	-0.00222	-0.00106	-0.00909	0.000197	0.00375
	(0.00301)	(0.00292)	(0.00656)	(0.00640)	(0.0144)
1970-1972	0	0	0	0	0
	()	()	()	()	()
1973-1974	0.00729**	0.00581*	0.0190**	0.0225**	0.0691***
	(0.00366)	(0.00338)	(0.00889)	(0.00878)	(0.0229)
1975-1976	0.0238***	0.0257***	0.0707***	0.0377***	0.101***
	(0.00647)	(0.00628)	(0.0159)	(0.00817)	(0.0205)
1977-1978	0.00898	0.00905	0.0260*	0.0353***	0.0619***
	(0.00690)	(0.00660)	(0.0137)	(0.00762)	(0.0189)
Observations	10,243	10,243	10,243	9,595	9,251
Adjusted R-squared	0.639	0.643	0.603	0.537	0.553
Mean of outcome	.6713	0.6708	1.0648	.6349	.6713
S.D. of outcome	.3397	.3332	.6501	.325	.3397
Mean of family tax break increment	2.3723	2.2497	2.3723	2.1829	.2958
S.D. of family tax break increment	1.3677	1.3116	1.3677	1.0229	.4564

Note: The table reports the coefficients and cluster-robust standard errors of the interaction terms of the cohort group dummies and the treatment variable in different regression specifications. Cluster robust standard errors in parentheses. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

1.6.2 Robustness tests based only on the 2016 Microcensus data

To alleviate concerns about the bias the data linking might create in the sample's composition, I develop an alternative identification strategy using only the 2016 Microcensus data for birth histories and demographic information. The primary advantage is that I can include all women from the 2016 Microcensus in the analysis regardless of being in a couple or not. However, there are some disadvantages to this approach. We cannot observe pre-policy characteristics, including the ones needed to impute prospective salaries. While male occupational choices are less expected to react to the tax incentives, it is less plausible for females. Therefore prospective salaries for them are imputed using only location and education information. Furthermore, we assume that women's relationships seen in 2016 can be projected back in time to represent the state in 2011. For these reasons, these findings can only augment the main results.

I altered the identification strategy regarding the older cohorts, i.e., testing for pre-trends. The within-cohort variation remained the same, relying on the non-linearities in the tax break along the dimensions of income and the number of children. However, for the older cohorts, instead of taking the number of children in 2011 as a basis for the potential tax break, I calculate the amount based on the number of children at the mother's age corresponding to the age of the younger, treated cohort. For instance, the 1975-1976 cohorts (which the analysis focuses on) were thirty-five or thirty-six years old at the policy's start and forty or forty-one at the end of the time window, so I compare all older cohorts at these same ages of the mothers while excluding the other younger cohorts from the regression. These regressions, therefore, capture the differences in the change in the number of children between ages thirty-five and forty (and, respectively, other age windows for the other female birth cohorts).

Table 1.5 presents the conditional average treatment effect on the treated estimates for families with at least two children, with a complete set of control variables.[17] The columns show the regression estimates with different treated cohorts going from the oldest to the younger ones. Let us consider first those cohorts for which the estimates reveal completed fertility effects. While there is no statistically significant result for the 1973-1974 cohort, we can see that for the pivotal 1975-1976 cohorts, an around 0.2% effect size is detected, or 0.4% at the average treatment size. Then the point estimate is similar for the 1977-1978 cohort, albeit not significant. For the youngest analyzed cohort, the estimated effect size has increased to around 0.5% (0.9% at the mean treatment), consistent with the fact that we expect higher short-run effects for cohorts not reaching completed fertility yet.

[17] Controls include prospective household income, mother's number of children in 2011, mother's education, father's education, mother's marital status in 2011, missing father dummy, log of paternal and maternal prospective salary, maternal place of living in terms of sub-county administrative unit and municipality type, age of the father, ethnicity of the parents.

Table 1.5: Robustness checks: completed fertility effect estimates for families with at least two children

Treated cohort:	1973-1974	1975-1976	1977-1978	1979-1980
Age 2011-2016:	37/38 - 42/43	35/36 - 40/41	33/34 - 38/39	31/32 - 36/37
Cohort 1961-1962	0	0	0	0
	(0)	(0)	(0)	(0)
Cohort 1963-1964	0.000271	-0.000224	-0.00163**	-0.00181*
	(0.000510)	(0.000660)	(0.000693)	(0.000924)
Cohort 1965-1966	0.000349	0.000570	-0.000236	2.17e-05
	(0.000431)	(0.000683)	(0.000820)	(0.00112)
Cohort 1967-1968	0.000506	0.000623	0.000876	0.00125
	(0.000457)	(0.000690)	(0.000714)	(0.000959)
Cohort 1969-1970	-0.000120	0.000498	0.000489	0.00229**
	(0.000429)	(0.000757)	(0.00105)	(0.000911)
Treated cohort	0.000969	0.00177**	0.00173	0.00458***
	(0.000628)	(0.000710)	(0.00111)	(0.00133)
Observations	44,342	43,210	40,176	36,427
Adjusted R-squared	0.907	0.854	0.798	0.739
Mean of outcome	1.2192	1.2184	1.2183	1.2167
S.D. of outcome	0.1887	0.1879	0.1869	0.1850
Mean of family tax break increment	2.0065	2.0584	2.0723	2.053
S.D. of family tax break increment	2.2527	2.2628	2.2441	2.1818

Note: The table reports the coefficients and cluster-robust standard errors of the interaction terms of the cohort group dummies and the treatment variable in different regression specifications, using a full set of control variables. Cluster robust standard errors in parentheses. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$

Table 1.6 reports the estimates for families with fewer than two children regarding fertility. Again first, let us inspect cohorts for which the estimates pertain to completed fertility. For the 1973-1974 cohorts, we cannot identify any significant effects of the policy, but for the cohorts of interest, the 1975-1976 cohorts, we can detect an around 1.2% effect (2.2% at the mean treatment) without any significant pre-trends suggesting validity. While the estimate for the 1977-1978 cohorts is non-significant (even negative), for the youngest 1979-1980 cohorts, the effect size elevates to around 2.1% (3.9% at the mean treatment).

Table 1.6: Robustness checks: completed fertility effect estimates for families with fewer than two children

Treated cohort:	1973-1974	1975-1976	1977-1978	1979-1980
Age 2011-2016:	37/38 - 42/43	35/36 - 40/41	33/34 - 38/39	31/32 - 36/37
Cohort 1961-1962	0	0	0	0
	(0)	(0)	(0)	(0)
Cohort 1963-1964	-0.00771*	-0.00293	-0.00245	-0.00640
	(0.00413)	(0.00495)	(0.00572)	(0.00587)
Cohort 1965-1966	-0.00584	-0.00551	-0.00365	-0.00430
	(0.00373)	(0.00591)	(0.00583)	(0.00589)
Cohort 1967-1968	0.00152	0.00337	-0.00575	-0.00948*
	(0.00409)	(0.00491)	(0.00558)	(0.00532)
Cohort 1969-1970	0.00203	0.00480	0.00102	0.00262
	(0.00359)	(0.00502)	(0.00561)	(0.00579)
Treated cohort	0.00661	0.0122**	-0.00190	0.0207***
	(0.00438)	(0.00553)	(0.00448)	(0.00625)
Observations	23,105	25,127	27,039	29,485
Adjusted R-squared	0.704	0.612	0.537	0.498
Mean of outcome	0.5424	0.5612	0.5789	0.5977
S.D. of outcome	0.3354	0.3554	0.3765	0.3945
Mean of family tax break increment	1.7826	1.8255	1.8616	1.8798
S.D. of family tax break increment	1.0686	1.0805	1.0803	1.077

Note: The table reports the coefficients and cluster-robust standard errors of the interaction terms of the cohort group dummies and the treatment variable in different regression specifications, using a full set of control variables. Cluster robust standard errors in parentheses. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$

To conclude these robustness exercises, we can note that we can again detect significant causal effects of the family tax break on completed fertility for the 1975-1976 cohorts. There is also some additional evidence for higher short-run effects for the younger cohorts. However, this 2016 sample and identification strategy indicates smaller effect sizes than the main results. This is not surprising, as women not in couples or in shorter relationships are also included in the analysis, who are expected to react less to tax breaks. Therefore these estimates can be considered a lower bound for the true effect of the reform.

1.7 Mechanism and alternative policy scenarios

In this section, I construct a simple model describing partial equilibrium household behavior, focusing on the key fertility and maternal employment variables. I estimate the model to fit empirical moments based on the 1966-1971 birth cohorts and compare the effect of two similar policies on fertility and female labor supply. The original family tax break policy (abbreviated as FTB in the figures) connects child support with employment status and the household's income level, providing a complex set of incentives regarding fertility and labor force participation. While it decreases the cost of additional children, especially for the third birth parity, it provides additional incentives to increase household income for lower-income households to the level where the entire tax break can be used. However, the additional income that families receive might decrease labor market participation of the mothers.

This model has two goals. First, I examine counterfactual policy scenarios to help to understand the effect of different policies on fertility and maternal employment. Second, I also validate the previous reduced-form empirical results while shedding light on the policy's potential female labor supply effects that could not be identified with the empirical strategy.

1.7.1 Framework

I introduce a simple model of household decisions over consumption, maternal labor supply, and the number of children, in parts building on the ideas of Becker and Tomes (1976) and Cigno (1986), but taking several shortcuts to keep the focus on the policies at hand. I augment the model with an ex-ante ideal number of children based on survey evidence of Kapitány and Spéder (2015), from which couples can deviate but will derive disutility in case they do. This is a novel approach where I abstract away from many potential drivers of a couple's fertility preferences and capture several non-economic reasons, such as cultural or sociological factors, with a single variable that has a solid basis in the demographic literature. But differently from most of the literature, additional children can appear as disutility to the household, not only as positive elements regarding utility.

The problem of the household can be formulated in the following way:

$$\max_{N\in\{0,1,2,3,4\},l^F\in\{0,1\}} \log(c) - \frac{\alpha}{2}(N-\nu)^2 + \beta q(l^F, N)\cdot \mathbb{I}[N>0]$$

$$s.t.$$

$$pc\left(2+0.4N\right) \leq (W^M + W^F l^F)\big(1-\tau(W^M, W^F, l^F, N)\big)$$

$$q(l^F, N) = \frac{2-l^F}{N}$$

where c denotes consumption, ν denotes the ex-ante ideal number of children, N denotes the total number of children, l^F denotes labor supply of the woman taking up either the value of 0 or 1 (being employed or not), and W^M and W^F denote the prospective maximum household gross labor income exogenous to the household, in case they work. The fraction of gross income paid in taxes and adjusted for the family tax break and allowances is given by $\tau(W^M, W^F, l^F, N)$, a function of the maximum prospective gross income of the household[18], labor supply, and the number of children. I abstract away from the male labor supply decision as men in Hungary predominantly work (around 90% employed in this sample).

The household derives utility from consumption and the quality of children while deriving disutility from not having the ideal number of children. Consumption is considered such that the price of a unitary consumption for adults is the poverty line level consumption of 50,000 HUF and for each child around 20,000 HUF reflecting the poverty line measures of the Hungarian Statistical Office (HCSO, 2015). This modeling choice abstracts away from having to split household consumption between the parents and the children, which would entail deriving separate utilities. The quality of children given by the function $q(l^F, N)$ depends negatively on maternal labor supply and the number of children the family has in total. This aspect captures that more children require more work assigned to the household rather than labor market activities. Note that I assume that employment is purely the choice of the household and that labor demand plays no part.

[18] I use the flat 16% personal income tax rate along with 17.5% rate of contributions, along with the family allowance amounts based on the government information website of the period Hungarian State Treasury.

1.7.2 Model fit

I study the model's behavior by solving it on a discrete grid of labor supply, which along with the discrete choice of the number of children, pins the consumption of the household via the budget constraint. I estimate the utility function parameters (α, β) to match the matched sample's completed fertility and maternal employment of the untreated 1966-1971 birth cohorts. I use a discrete distribution of paternal and maternal labor income based on the imputed prospective gross salaries, while I add the distributions on the ideal number of children based on survey evidence (Kapitány and Spéder, 2015). During the estimation, I maintain the assumption that the ideal number of children is independent of the household income.[19]

Table 1.7 reports the targeted and fitted simulated values of the moments. We can see that the model can reproduce aggregate empirical moments reasonably well, while Appendix Figure 3.11 shows that the found parameter values of $\alpha = 1.8838$ (SE: 0.0087), $\beta = 0.9087$ (SE: 0.0186) indeed minimize the loss function.

Table 1.7: Targeted empirical and simulated moments

Targeted moment	Empirical value	Simulated value
Completed fertility	2.01	2.00
Female employment rate (in %)	78.65	78.33

Note: The table reports the targeted empirical and simulated moments for the model estimation.

To test the model's predictive power, Table 1.8 compares non-targeted empirical and simulated moments. We can see that while the model is not directly matched on the fertility rates conditional on income, the model can accurately reproduce these moments for low and middle-income families. For high-income families, the model overpredicts completed fertility, which can be due to the violation of the independence assumption between income and the ideal number of children (indeed, these families might have fewer children on average ideally than the others). I also included in the table the implied policy effect of the model compared to the estimates I presented earlier in the paper: the model predicts a somewhat

[19] While this is a strong assumption, it ensures that the effects of the examined policies on households with different income levels will not be the result of differential preferences.

higher overall completed fertility effect of the policy than the main estimates; however, the magnitudes of the policy effect implied by the two methods are similar.

Table 1.8: Non-targeted empirical and simulated moments

Non-Targeted moment	Empirical value	Simulated value
Correlation between household income and completed fertility	-50.03%	-48.11%
Low income (40%) completed fertility	2.09	2.10
Middle income (40%) completed fertility	1.91	1.92
High income (20%) completed fertility	1.87	1.98
Change in completed fertility due to policy	2.36%	3.25%

Note: The table reports the non-targeted empirical and simulated moments based on the model. Household income is categorized into bins, and I use the middle value of the bins to calculate the correlation between household income and completed fertility. Low-income households (40% of households) are defined as earning less than 250,000 HUF (~1,000 EUR), middle-income households (40% of households) are defined as earning up to 400,000 HUF (~1,500 EUR), and the high-income group is the remaining 20%.

1.7.3 Policy scenarios

In this section, I use the estimated model to examine the equilibrium effects of the original family tax break policy while also introducing two alternatives:

- Alternative family tax break: 80,000 HUF per child per month tax base deduction, without a large jump at the third child[20]

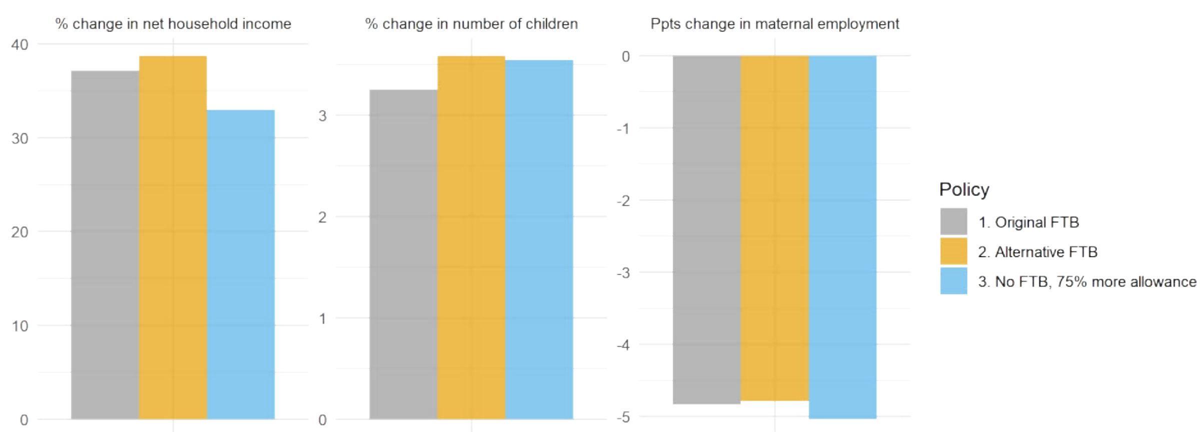
- No family tax break, but a 75% increase of the family allowance

The exercise aims to see whether the model corroborates the previously estimated long-run fertility effects and to provide predictions regarding long-run direct female employment effects. Additionally, it helps to identify which income groups react to the policy and which alternatives would work well. While the 'Alternative family tax break' simulates a simpler, kinkless version, the increase of the family allowance simply decreases the cost of children without any requirements to the income or employment status of the family. I also group households into three income categories with 40%-40%-20% shares based on reasonable household income thresholds (250,000 HUF and 400,000 HUF, $\sim 1,000$ EUR and ~1,500 EUR) to study how the different policies affect different income groups in the long run.

[20]The original version of the policy had 62,500 HUF deduction for families with fewer than two children, and 206,250 HUF for families with at least two children.

Figure 1.18 shows the aggregate changes due to different policies regarding the number of children (%), net household income (%), and maternal employment (in percentage points). We can see that the model suggests a 3.2% increase in the number of children due to the original family tax break policy, along with a 4.8 percentage point lower maternal employment. The change in completed fertility is greater but reasonably close to what the reduced form estimates indicate, yielding around 2.4% (taking the sample size weighted sum of the estimates for families with at least and fewer than two children). The alternative policies are comparable in government expenditure (households end up with around 30-35% more net income), so they can be considered appropriate comparisons. The alternative 'flat' family tax break would lead to slightly higher completed fertility effects (3.6%), with a similar effect on maternal employment rate (-4.8%) with approximately the same change in net income, while the family allowance scenario without tax breaks results in higher fertility effects (3.6%), slightly lower maternal employment (-5.0%), but around 5.0% less government expenditure.

Figure 1.18: Policy scenarios: aggregate changes in outcomes compared to the no-policy baseline

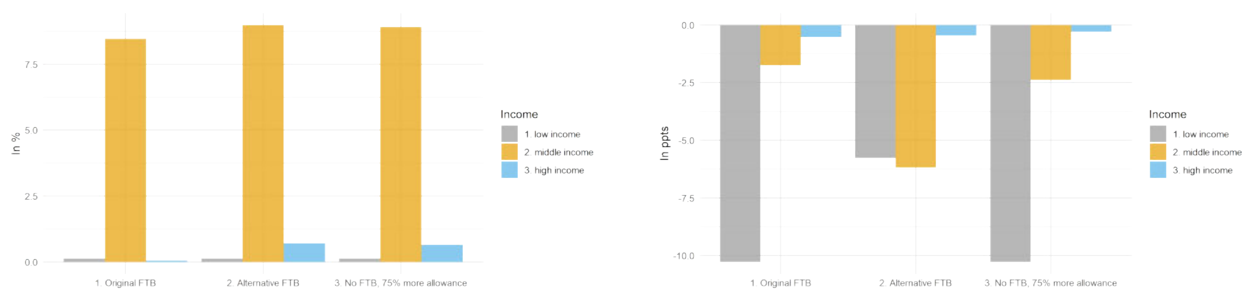


Note: the figure reports the model simulations' aggregate outcomes under different policies of the original family tax break (FTB), the alternative family tax break, and the no tax break scenario with elevated family allowance.

Figure 1.19 shows the policy effects conditional on income groups. We can see that the bulk of the fertility effect for both policies is generated by middle-income households, resulting in an 8% increase in the number of children for those families. In contrast, the impact on the poorer or richer groups of families is minimal. Both alternative policies result in higher fertility for middle and high-income families leaving low-income families unaffected. Regarding female labor supply the impacted groups are different. While the original family tax break and the family allowance policy both decrease maternal employment of lower-income families by around 10%, middle and higher-income families are only slightly affected or not at all. In comparison, the alternative tax break policy would lead to an around 5% decrease in female employment for both low and middle-income families in the long run while still keeping high-income families untouched. Overall, high-income families do not change their behavior due to these policies, while lower and middle-income families are more responsive concerning their fertility and employment decisions.

Figure 1.19: Policy effects compared to the baseline by income groups

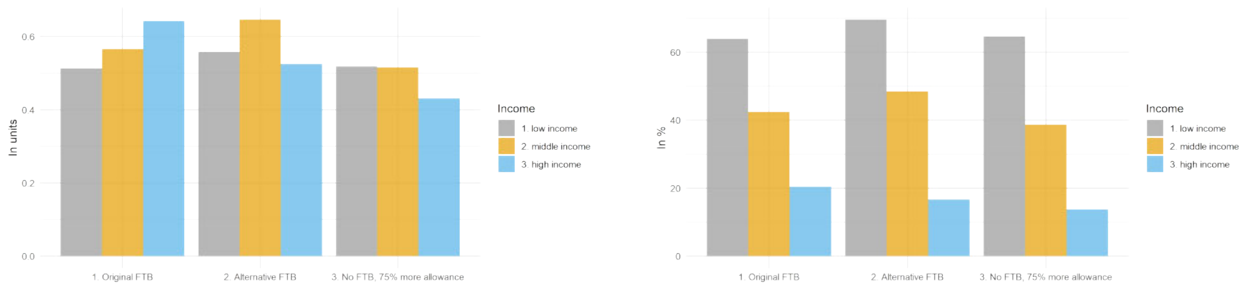
(a) Number of children (b) Maternal employment



Note: based on the policy simulations of the author, compared to the no-policy change baseline. The figure reports the model simulations' prediction for change in the number of children and maternal employment under different policies of the original family tax break (FTB), the alternative family tax break, and the no tax break scenario with elevated family allowance. Low-income households (40% of households) are defined as earning less than 250,000 HUF (∼1,000 EUR), middle-income households (40% of households) are defined as earning up to 400,000 HUF (∼1,500 EUR), and the high-income group is the remaining 20%.

Finally, it is also worth investigating the redistributive consequences of different policies. Figure 1.20 reports the change in net income: on the left-hand side in absolute terms (in units of poverty line consumption) and in percentage terms on the right-hand side. The net income effect of the original family tax break in absolute terms grows with the net income of households, which corroborates the finding of Tóth G. and Virovácz (2013). In contrast, the alternative tax break policy enriches middle-income families the most, while the family allowance favors low and middle-income families the most. In percentage terms, all the examined policies give more to the poorer households because they start with a lower income base, the alternative linear tax break providing the largest differential in net income.

Figure 1.20: Policy effects on net household income compared to the baseline by income groups



Note: based on the policy simulations of the author, compared to the no-policy change baseline. The figure reports the model simulations' prediction for change in net household income under different policies of the original family tax break (FTB), the alternative family tax break, and the no tax break scenario with elevated family allowance. Low-income households (40% of households) are defined as earning less than 250,000 HUF (∼1,000 EUR), middle-income households (40% of households) are defined as earning up to 400,000 HUF (∼1,500 EUR), and the high-income group is the remaining 20%.

1.8 Discussion

In this section, I put the findings into context: I discuss the internal and external validity of the findings and place them in the related literature. Key components of this discussion include the evaluation of the matching process regarding external validity, the pre-trend tests regarding internal validity, and the magnitude of the results in terms of short-run vs. long-run.

The analysis sample is a product of an exact matching process depending on the availability of the two birth dates of the adults in a household, which implies that the sample could only include couples whose relationships existed both in 2011 and 2016. That condition inherently constrains the external validity of the results, as the sample necessarily includes only more stable relationships. The family tax break policy affects families over an extended portion of the life cycle, so more stable relationships could expect more benefits from it in the long run. It might lead to overestimating the policy's effectiveness for the entire population. I combatted these difficulties in two ways. On the one hand, I included a rich set of control variables in the regressions to mitigate this bias. On the other hand, I examined the entire sample of women in the 2016 Microcensus without data linking and reestimated the effects using information only from that data source. In this exercise, I found smaller effect sizes of the reform; however, the new estimates remained statistically significant.

The identifying assumptions are tested in this paper by examining how pre-trends behave in different regression specifications. The results regarding families with at least two children have the most robust case in terms of internal validity: even before including additional controls in the regressions, I find no statistically significant estimates indicating 'placebo' effects. Further robustness checks corroborate these results, as different measurements and sample restrictions yield qualitatively similar estimates. For the results of families with fewer than two children, the picture is more mixed. While after controlling for household income and the number of children pre-trends disappear, the indicated treatment effects might deviate from the main estimates with different regression specifications, suggesting that we should be more careful with the interpretation.

With that in mind, we can summarise the findings of this paper in the following way. While the full family tax break benefit for three children would result in a maximum of around 4,400 EUR/year additional household income, the average treatment size is only around 1,000 EUR/year in this sample. The main estimates presented here suggest that for families with at least two children, the policy resulted in an around 0.8% increase in completed fertility. A less robust, around 5.6% effect is identified for families with fewer than two children. Taking the additional robustness exercises into account, the effect size of the policy for families with at least two children is bounded by the estimates of around 0.4% and 0.9%, while for families with fewer than two children, the bounding estimates fall between 2.2% and 8.5%. So using the sample sizes as weights, the overall main effect of the policy sizes up to be around 2.4%, with the other estimates ranging from 1.0% to 3.4%. The heterogeneity analysis revealed that the change in completed fertility overall was driven by religious and married families and by those that already had a child aged less than three at home. The structural model, estimated on the pre-policy sample's fertility and employment outcomes, predicted a similar, around 3.2% effect on completed fertility out-of-sample, providing additional support to the magnitude of the effect of the family tax break policy. In the model, the fertility effects are driven by middle-income households.

Regarding maternal employment, I could not detect any statistically significant short-run indirect (via having additional children) effects of the policy, although the estimates suggest more negative than positive short-run employment effects. However, the heterogeneity analysis suggested that in the short-run religious households with a higher amount of tax break ended up with significantly lower employment, while married couples ended up with higher employment than their non-married counterparts. The structural model indicated a negative, around 4.8% effect of the policy on the long-run partial equilibrium employment rate for women (as a baseline, the 1966 cohort had an 81% employment rate), driven by large decreases for lower-income households, leaving the middle and higher-income families almost unchanged.

Table 1.9 puts the magnitude of the findings of this paper into context by tabulating some relevant papers in the literature. I estimate the effects of the family tax break policy in Hungary to be lower than the short-run effects in the literature, consistent with the estimates not capturing the timing effects simultaneously. However, they are higher than the previous estimates of completed fertility effects. One explanation could be that I examine a country where the additional money provided by the government is large in size and could make a difference in the life of a family, and the policy is a recurring, dependable tax break compared to a one-time transfer per child. It is important to note that we cannot yet observe the effect on families that adjust their life cycle due to the changed tax incentives, using comparative statics, the structural model estimated in this paper would suggest only a slightly larger effect. As the last point, I mention that the magnitudes I estimate in this paper are in line with other studies on Hungary that use different empirical strategies as time series or quasi-panel regressions (Gábos et al., 2009; Bördős and Szabó-Morvai, 2021).

Table 1.9: Comparable results in the literature

Countries	Studies	Year	Policy	Size	Short-run	Long-run
Canada	Milligan (2005), Parent and Wang (2007)	1995	One-time transfer per child	+4,800 EUR	9%	close to 0%
Germany	Adda et al. (2017)	Sim.	One-time transfer per child	+6,000 EUR	4.5%	0.2%
Germany	Raute (2019)	2007	Maternity leave benefit	+5,000 EUR	16%	
Israel	Cohen et al. (2013)	2003	Child subsidy	-360 EUR/year	-1%	
Spain	Azmat and González (2010)	2003	Tax credit	+900 EUR/year	4.6%	
Spain	González (2013)	2007	One-time transfer per child	+2,500 EUR	6%	
Hungary	this paper	2011	Tax break	+1,000 EUR/year PV: 13,000 EUR (5%, 20 years)		2.4%

Note: based on the estimates in cited studies, converting to EUR according to the exchange rate of the policy year. For the present value calculation of the family tax break, I used a 5% discount rate for twenty years.

1.9 Conclusion

In this paper, I study the effects of tax incentives on completed fertility, and on maternal employment to a lesser degree, by examining the large-scale extension of the Hungarian family tax break in 2011. I argue that the policy introduced quasi-experimental variation along household income and the initial number of children due to its non-linear jump for the third child. I build a household-level linked dataset from the 2016 Microcensus and 2011 Census of Hungary, augmented with prospective parental salaries from the National Wage Surveys, and show how additional prospective income for having additional children affects the completed fertility of women and their labor market choices.

I find that the policy led to 0.8% higher completed fertility for families with at least two children and 5.6% for those with fewer than two children (weighted average of around 2.4%). The findings are driven by couples who claim to be religious, who are married, or has a child younger than three years old. Results regarding maternal employment effects due to having additional children are not statistically significant overall but some subsets of the population (religious, non-married) are detected to end up with lower maternal employment. I also fit a simple partial equilibrium model of household behavior to simulate counterfactual policies and to examine possible long-run effects regarding maternal employment, completed fertility, and income redistribution. The model suggests the policy causes around 3.2% higher completed fertility (somewhat larger than the main estimates but close in magnitude) driven by middle-income families, and around 4.8 percentage points lower employment rate for mothers driven by lower-income families.

This paper focuses on the 2011-2016 period, in which the total fertility rate of Hungary increased from 1.23 to 1.53, by around 24%. The change is in part mechanical due to the end of a demographic transition period in the country (Spéder, 2021), but according to the results presented in this paper, around one-tenth of the change can be attributed to the family tax break reform. However, present fertility measures still lag behind women's reported preferences in Hungary, indicating that there might be even more critical obstacles than financial constraints in realizing family size goals. These might include the availability of nurseries, the difficulty of returning to the labor market, fathers' participation, housing, or the state of the marriage market. Studying these factors present a natural continuation to this line of research, which could lay the foundations for effective future family policies in developed countries.

References

Abramitzky, R., Mill, R., and Pérez, S. (2020). Linking individuals across historical sources: A fully automated approach. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 53(2):94–111.

Adda, J., Dustmann, C., and Stevens, K. (2017). The Career Costs of Children. *Journal of Political Economy*, 125(2):293–337.

Azmat, G. and González, L. (2010). Targeting fertility and female participation through the income tax. *Labour Economics*, 17(3):487–502.

Balestrino, A. and Cigno, A. (2002). Endogenous Fertility and the Design of Family Taxation. *International Tax and Public Finance*, 9:175–193.

Baudin, T., de la Croix, D., and Gobbi, P. E. (2015). Fertility and Childlessness in the United States. *American Economic Review*, 105(6):1852–1882.

Baughman, R. and Dickert-Conlin, S. (2003). Did Expanding the EITC Promote Motherhood? *American Economic Review*, 93(2):247–251.

Baughman, R. and Dickert-Conlin, S. (2009). The earned income tax credit and fertility. *Journal of Population Economics*, 22(3):537–563.

Becker, G. S. (1960). An Economic Analysis of Fertility. In Bureau, U.-N., editor, *Demographic and Economic Change in Developed Countries*, pages 209–240. Columbia University Press, New York. OCLC: 38504679.

Becker, G. S. and Lewis, H. G. (1973). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy*, 81(2, Part 2):S279–S288.

Becker, G. S. and Tomes, N. (1976). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy*, 84(4):S143–S162.

Billari, F. and Kohler, H.-P. (2004). Patterns of low and lowest-low fertility in Europe. *Population Studies*, 58(2):161–176.

Bongaarts, J. and Sobotka, T. (2012). A Demographic Explanation for the Recent Rise in European Fertility. *Population and Development Review*, 38(1):83–120.

Bördős, K. and Szabó-Morvai, (2021). A 2000-2015 közötti családpolitikai intézkedések hatása a születésszámra Magyarországon [Impact of Family Policy Measures Between 2000 and 2015 on the Number of Births in Hungary]. *Demográfia*, 64(2-3).

Callaway, B., Goodman-Bacon, A., and Sant'Anna, P. H. (2021). Difference-in-differences with a continuous treatment. *arXiv preprint arXiv:2107.02637*.

Cigno, A. (1986). Fertility and the Tax-Benefit System: A Reconsideration of the Theory of Family Taxation. *The Economic Journal*, 96(384):1035.

Cigno, A. (2001). Comparative Advantage, Observability, and the Optimal Tax Treatment of Families with Children. *International Tax and Public Finance*, 8:455–470.

Cigno, A. and Pettini, A. (2002). Taxing family size and subsidizing child-specific commodities? *Journal of Public Economics*, 84(1):75–90.

Clark, A. E. and Lepinteur, A. (2020). A Natural Experiment on Job Insecurity and Fertility in France. *The Review of Economics and Statistics*, pages 1–47.

Cohen, A., Dehejia, R., and Romanov, D. (2013). Financial Incentives and Fertility. *Review of Economics and Statistics*, 95(1):1–20.

Currie, J. and Schwandt, H. (2014). Short- and long-term effects of unemployment on fertility. *Proceedings of the National Academy of Sciences*, 111(41):14734–14739.

Cygan-Rehm, K. (2016). Parental leave benefit and differential fertility responses: evidence from a German reform. *Journal of Population Economics*, 29(1):73–103.

Dahl, G. B., Løken, K. V., Mogstad, M., and Salvanes, K. V. (2016). What Is the Case for Paid Maternity Leave? *Review of Economics and Statistics*, 98(4):655–670.

Docquier, F. (2004). Income Distribution, Non-convexities and the Fertility-Income Relationship. *Economica*, 71(282):261–273.

Francesconi, M. (2002). A Joint Dynamic Model of Fertility and Work of Married Women. *Journal of Labor Economics*, 20(2):336–380.

Fraser, C. D. (2001). Income Risk, the Tax-benefit System and the Demand for Children. *Economica*, 68(269):105–126.

Garganta, S., Gasparini, L., Marchionni, M., and Tappatá, M. (2017). The Effect of Cash Transfers on Fertility: Evidence from Argentina. *Population Research and Policy Review*, 36(1):1–24.

Gauthier, A. H. (2007). The impact of family policies on fertility in industrialized countries: a review of the literature. *Population Research and Policy Review*, 26(3):323–346.

Goldin, J. and Michelmore, K. (2021). Who Benefits from the Child Tax Credit. Technical Report Working Paper 27940, National Bureau of Economic Research.

González, L. (2013). The Effect of a Universal Child Benefit on Conceptions, Abortions, and Early Maternal Labor Supply. *American Economic Journal: Economic Policy*, 5(3):160–188.

Gábos, A., Gál, R. I., and Kézdi, G. (2009). The effects of child-related benefits and pensions on fertility by birth order: A test on Hungarian data. *Population Studies*, 63(3):215–231.

HCSO (2015). Létminimum, 2014 [The Minimum Subsistence Figure, 2014]. Technical Report 462, Hungarian Central Statistical Office.

Hoynes, H. (2019). The Earned Income Tax Credit. *The ANNALS of the American Academy of Political and Social Science*, 686(1):180–203.

Kalwij, A. (2010). The Impact of Family Policy Expenditure on Fertility in Western Europe. *Demography*, 47(2):17.

Kapitány, B. and Spéder, Z. (2015). Fertility. In Monostori, J., Őri, P., and Spéder, Z., editors, *Demographic Portrait of Hungary 2015.*, pages 41–55. HDRI, Budapest.

Keane, M. P. and Wolpin, K. I. (2002). Estimating Welfare Effects Consistent with Forward-Looking Behavior. Part I: Lessonsfrom a Simulation Exercise. *The Journal of Human Resources*, 37(3):570–599.

Keane, M. P. and Wolpin, K. I. (2010). The Role of Labor and Marriage Markets, Preference Heterogeneity, and the Welfare System in the Life Cycle Decisions of Black, Hispanic, and White Women. *International Economic Review*, 51(3):851–892.

Kim, Y.-I. A. (2014). Lifetime impact of cash transfer on fertility. *Canadian Studies in Population*, (1):14.

Kohler, H.-P., Billari, F. C., and Ortega, J. A. (2006). Low Fertility in Europe: Causes, Implications and Policy Options. In Harris, F. R., editor, *The Baby Bust: Who will do the Work? Who will Pay the Taxes?*, pages 48–109. Rowman & Littlefield Publishers, Lanham, MD.

Lalive, R. and Zweimüller, J. (2009). How does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments. *Quarterly Journal of Economics*, 124(3):1363–1402.

Laroque, G. and Salanié, B. (2014). Identifying the Response of Fertility to Financial Incentives. *Journal of Applied Econometrics*, 29(2):314–332.

Lee, R. and Mason, A. (2014). Is low fertility really a problem? Population aging, dependency, and consumption. *Science*, 346(6206):229–234.

Makay, Z. (2015). Family support system – childraising – employment. In Monostori, J. Őri, P. S. Z., editor, *Demographic Portrait of Hungary 2015.*, pages 57–74. HDRI, Budapest.

Makay, Z. (2019). The Family Support System and Female Employment. In Monostori, J. Őri, P. S. Z., editor, *Demographic Portrait of Hungary 2018.*, pages 85–105. HDRI, Budapest.

Malthus, T. (1798). *An Essay on the Principle of Population*. London.

Meier, V. and Wrede, M. (2013). Reducing the excess burden of subsidizing the stork: Joint taxation, individual taxation, and family tax splitting. *Journal of Population Economics*, 26(3):1195–1207.

Milligan, K. (2005). Subsidizing the Stork: New Evidence on Tax Incentives and Fertility. *The Review of Economics and Statistics*, 87(3):539–555.

Moffitt, R. (1984). Profiles of Fertility, Labour Supply and Wages of Married Women. A Complete Life-Cycle Model. *Review of Economic Studies*, 51(2):263–278.

Moffitt, R. (1998). The Effect of Welfare on Marriage and Fertility: What Do We Know and What Do We Need To Know. In Moffitt, R., editor, *The family and reproductive behaviior: Research perspectives*, pages 50–96. National Academy Press, Washington DC.

National Assembly of Hungary (2010). Act CXXIII of 2010 on the Amendment of Laws on Taxes and Contributions. *Magyar Közlöny*, pages 25286–25355.

OECD (2021). PF1.1 Public spending on family benefits., Family Database.

Parent, D. and Wang, L. (2007). Tax incentives and fertility in Canada: quantum vs tempo effects. *Canadian Journal of Economics/Revue canadienne d'économique*, 40(2):371–400.

Raute, A. (2019). Can financial incentives reduce the baby gap? Evidence from a reform in maternity leave benefits. *Journal of Public Economics*, 169:203–222.

Riphahn, R. T. and Wiynck, F. (2017). Fertility effects of child benefits. *Journal of Population Economics*, 30(4):1135–1184.

Sobotka, T. (2011). Fertility in Central and Eastern Europe after 1989: Collapse and Gradual Recovery. *Historical Social Research / Historische Sozialforschung*, 36(2):246–296.

Sommer, K. (2016). Fertility choice in a life cycle model with idiosyncratic uninsurable earnings risk. *Journal of Monetary Economics*, 83:27–38.

Spéder, Z. (2021). Termékenységi mintaváltás – a családalapí- tás átalakulásának demográfiai nyomvona- lai Magyarországon [Fertility pattern change – demographic traces of the transformation of family formation in Hungary]. *Szociológiai szemle*, 31(2):4–29.

Spéder, Z. and Kapitány, B. (2009). How are Time-Dependent Childbearing Intentions Realized? Realization, Postponement, Abandonment, Bringing Forward: Les intentions de fécondité sont-elles réalisées dans le délai prévu ? Réalisation, report, abandon, avancement. *European Journal of Population / Revue européenne de Démographie*, 25(4):503–523.

Spéder, Z., Murinkó, L., and Oláh, L. S. (2020). Cash support vs. tax incentives: The differential impact of policy interventions on third births in contemporary Hungary. *Population Studies*, 74(1):39–54.

Thévenon, O. and Gauthier, A. H. (2011). Family policies in developed countries: a 'fertility-booster' with side-effects. *Community, Work & Family*, 14(2):197–216.

Tóth G., C. and Virovácz, P. (2013). Winners and Losers: An assessment of the Hungarian flat tax reform. *Public Finance Quarterly*, 58(4):369–385.

Varga, G., Nobilis, B., and Svraka, A. (2020). Adóköteles jövedelmek rugalmassága –egy identifikációs kísérlet a családi adókedvezmény 2011-es bevezetése alapján [elasticity of taxable income - an identification experiment based on the 2011 introduction of the family tax break]. *Statisztikai Szemle*, 98(1):5–33.

Őri, P. and Spéder, Z. (2020). Folytonos Átmenet: Magyarország népesedése 1920 és 2020 között [Permanent transition: demography of Hungary between 1920 and 2020]. *Statisztikai Szemle*, 98(6):481–521.

Appendix

Table 1A1: Family policies in Hungary around 2011-2016

Name	Age of child	Amount	Eligibility
Baby-care allowance (CSED)	0-6 months	70% of previous earnings	Employment prior to childbirth
Childcare benefit (GYED)	6-24 months	70% of previous earnings with upper limit of 1.4 times the minimum wage (in 2013: 1,646,400 HUF∼6,000 EUR a year)	Employment prior to childbirth
Childcare allowance (GYES)	24-36 months if employed 0-36 months if unemployed	statutory minimum of old age pension (in 2013:342,000 HUF∼1,270 EUR a year)	Universal
Childrearing support (GYET)	3-8 years old	statutory minimum of old age pension (in 2013: 342,000 HUF∼1,270 EUR a year)	Families of three or more children with a parent working 30 hours, or from home
Childrearing allowance	0-6 years old	144,000 HUF∼530 EUR a year	Child not yet enrolled in education
Schooling support	6-18/20 years old	144,000 HUF∼530 EUR a year	Child enrolled in education

Note: the table reports the most important family policies in Hungary apart from the family tax break, based on Makay (2015). CSED and GYED are subject to personal income tax, and GYES is subject to a 10% pension contribution.

Table 1A2: Fertility effects of the family tax break, log-specification

Mother's birth cohort	Families with at least two children			Families with fewer than two children		
Cohort 1966-1967	0.000991	-0.000370	-0.000475	-0.0142***	-0.00378*	-0.00358
	(0.00198)	(0.000682)	(0.000673)	(0.00472)	(0.00220)	(0.00281)
Cohort 1968-1969	-0.000418	0.000331	0.000211	-0.0102**	-0.00135	-0.00222
	(0.00190)	(0.000694)	(0.000716)	(0.00506)	(0.00237)	(0.00301)
Cohort 1970-1972	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)
Cohort 1973-1974	0.00417**	0.00117	0.00103	-0.00181	0.00586	0.00729**
	(0.00183)	(0.00107)	(0.00104)	(0.00595)	(0.00361)	(0.00366)
Cohort 1975-1976	**0.0129*** **	**0.00410*** **	**0.00329*** **	**0.00529**	**0.0351*** **	**0.0238*** **
	(0.00177)	**(0.000993)**	**(0.00103)**	**(0.00757)**	**(0.00633)**	**(0.00647)**
Cohort 1977-1978	0.0172***	0.00494***	0.00433***	-0.0342***	0.0177**	0.00898
	(0.00196)	(0.00156)	(0.00151)	(0.00788)	(0.00712)	(0.00690)
Observations	20,721	20,721	20,721	10,243	10,243	10,243
Adjusted R-squared	0.466	0.892	0.896	0.103	0.621	0.639
Controls	None	Income, children	All	None	Income, children	All
Mean of outcome	1.2282	1.2282	1.2282	.6713	.6713	.6713
S.D. of outcome	.1966	.1966	.1966	.3397	.3397	.3397
Mean of family tax break increment	2.2892	2.2892	2.2892	2.3723	2.3723	2.3723
S.D. of family tax break increment	2.3148	2.3148	2.3148	1.3677	1.3677	1.3677

Note: based on regressing the log(1+number of children) in 2016 on cohort groups interacted with treatments, with a baseline of 1970-1972. The full set of controls includes the following from 2011 besides household income and number of children: gender composition of children; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); location; ethnicity; illnesses and disabilities; foreign language knowledge; religiosity; wealth, pension and social care status; public employment status; and the type of income imputation. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 1A3: Fertility effects of the family tax break, count-specification

	Families with at least two children			Families with less than two children		
Cohort 1966-1967	0.00577	-0.000921	-0.00117	-0.0341***	-0.0133**	-0.0106*
	(0.00835)	(0.00263)	(0.00265)	(0.00924)	(0.00556)	(0.00642)
Cohort 1968-1969	-0.000520	0.00201	0.00167	-0.0257**	-0.00974*	-0.00909
	(0.00806)	(0.00265)	(0.00277)	(0.0107)	(0.00577)	(0.00656)
Cohort 1970-1972	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)
Cohort 1973-1974	0.0158**	0.00260	0.00226	-0.000241	0.0172**	0.0190**
	(0.00758)	(0.00418)	(0.00404)	(0.0128)	(0.00873)	(0.00889)
Cohort 1975-1976	**0.0472*** **	**0.0129*** **	**0.00995** **	**0.0125**	**0.0929*** **	**0.0707*** **
	(0.00733)	**(0.00380)**	**(0.00392)**	**(0.0168)**	**(0.0158)**	**(0.0159)**
Cohort 1977-1978	0.0643***	0.0153**	0.0133**	-0.0864***	0.0411***	0.0260*
	(0.00773)	(0.00599)	(0.00575)	(0.0151)	(0.0140)	(0.0137)
Observations	20,721	20,721	20,721	10,243	10,243	10,243
Adjusted R-squared	0.403	0.898	0.902	0.141	0.584	0.603
Controls	None	Income + Children	All	None	Income + Children	All
Mean of outcome	2.4888	2.4888	2.4888	1.0648	1.0648	1.0648
S.D. of outcome	.8013	.8013	.8013	.6501	.6501	.6501
Mean of family tax break increment	2.2892	2.2892	2.2892	2.3723	2.3723	2.3723
S.D. of family tax break increment	2.3148	2.3148	2.3148	1.3677	1.3677	1.3677

Note: based on regressing the number of children in 2016 on cohort groups interacted with treatments, with a baseline of 1970-1972. The full set of controls includes the following from 2011 besides household income and number of children: gender composition of children; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); location; ethnicity; illnesses and disabilities; foreign language knowledge; religiosity; wealth, pension and social care status; public employment status; and the type of income imputation. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Figure 1A1: Fertility: families with at least two children, unweighted

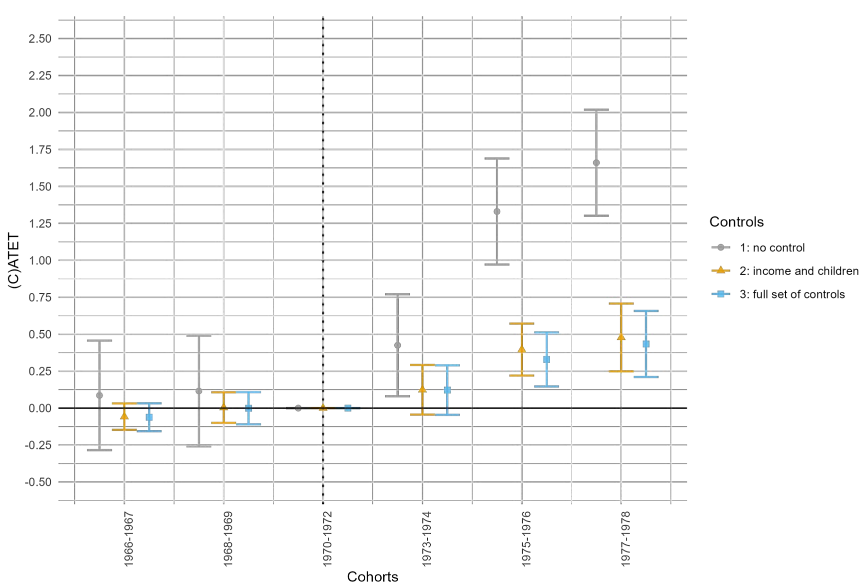


Figure 1A2: Fertility: families with fewer than two children, unweighted

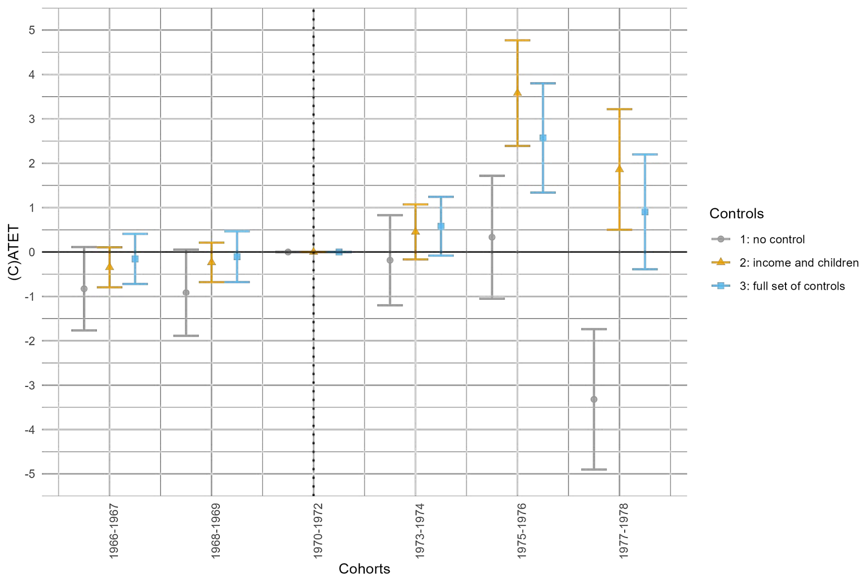


Figure 1A3: Fertility: families with at least two children, counts

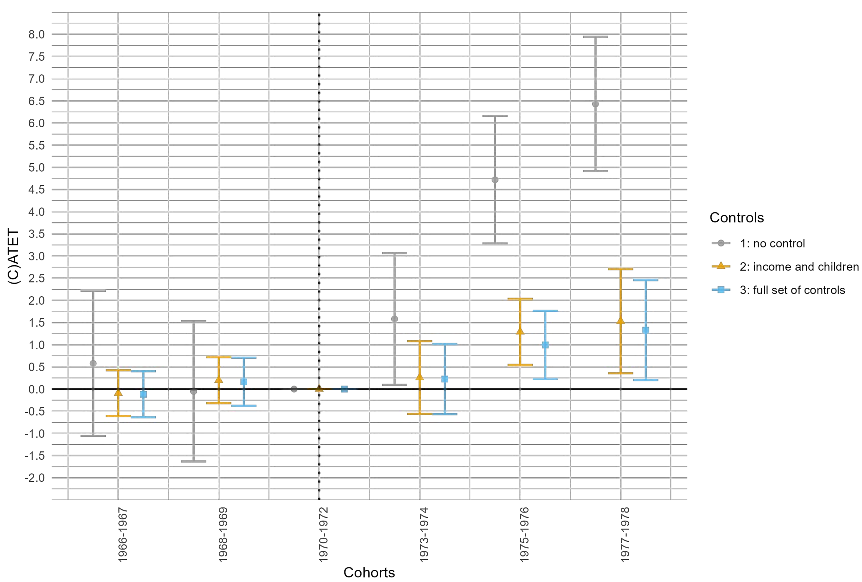


Figure 1A4: Fertility: families with fewer than two children, counts

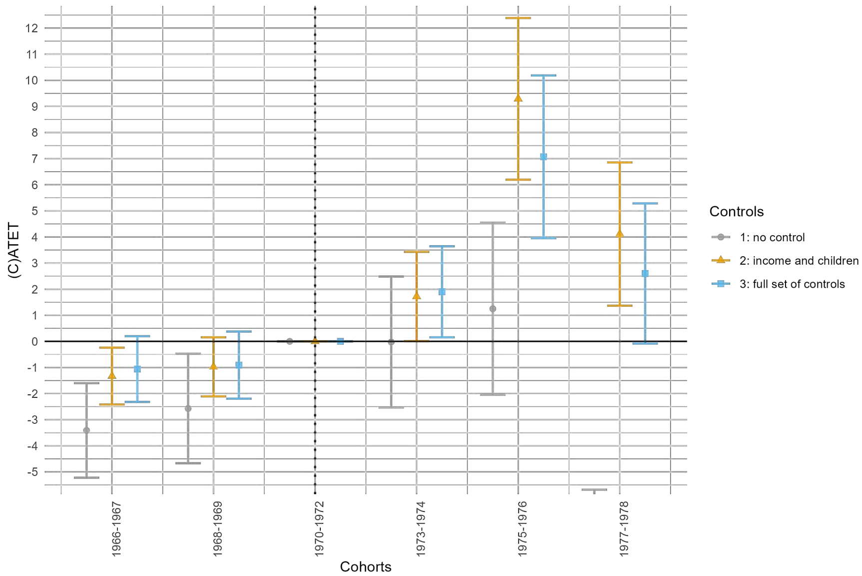


Figure 1A5: Fertility: families with at least two children, other spec.

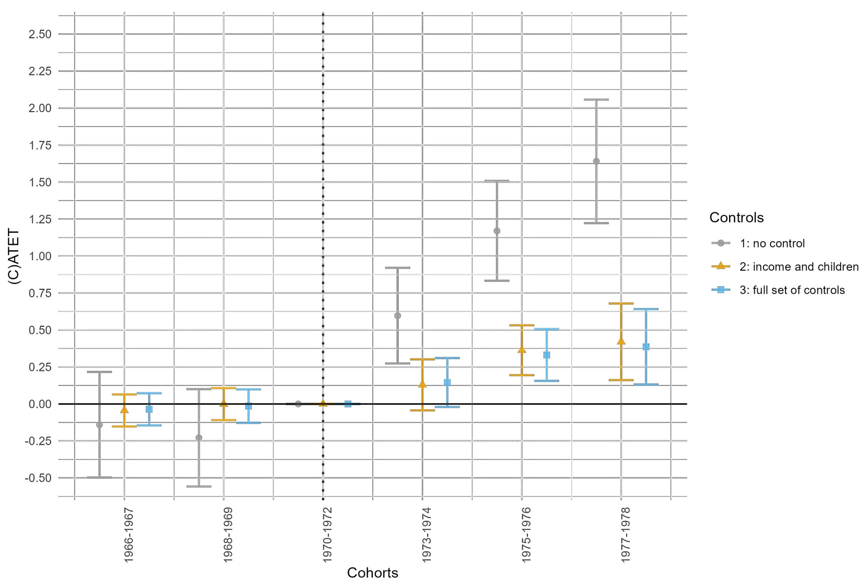


Figure 1A6: Fertility: families with fewer than two children, other spec.

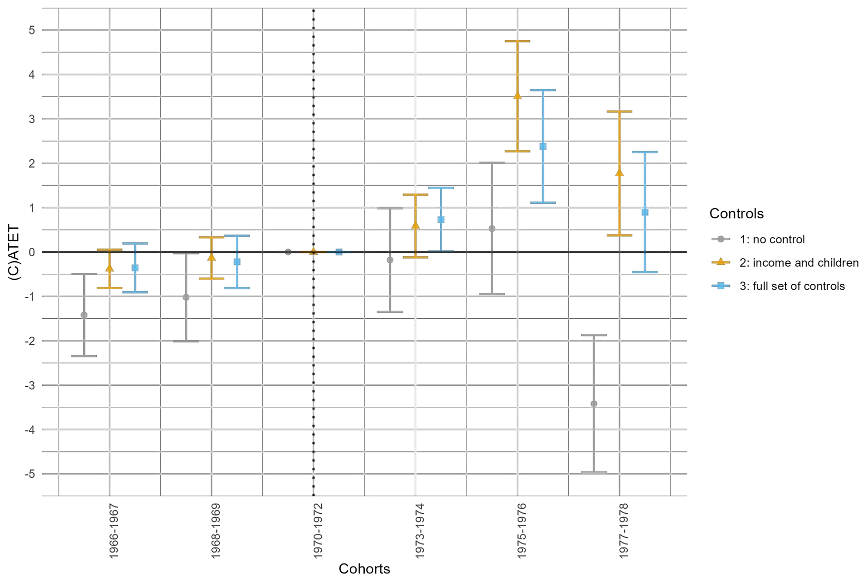


Figure 1A7: Fertility: families with at least two children, discrete

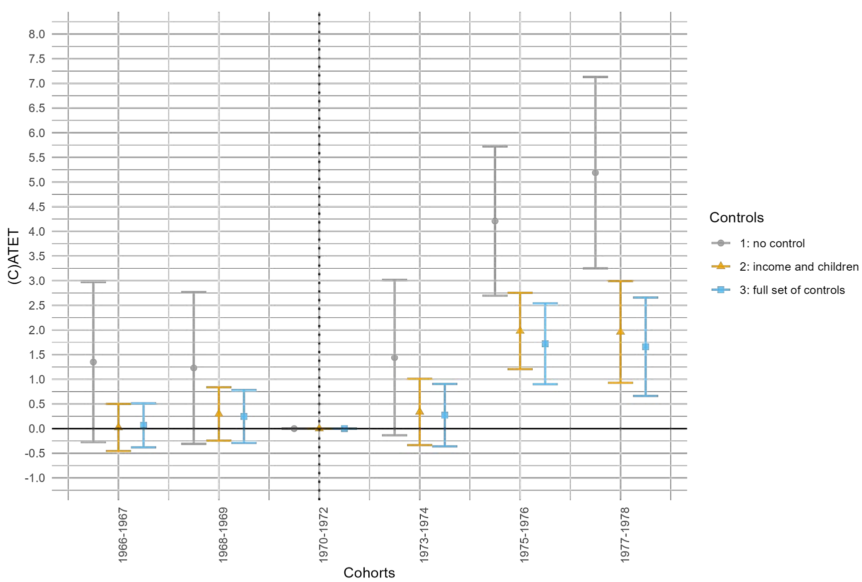


Figure 1A8: Fertility: families with fewer than two children, discrete

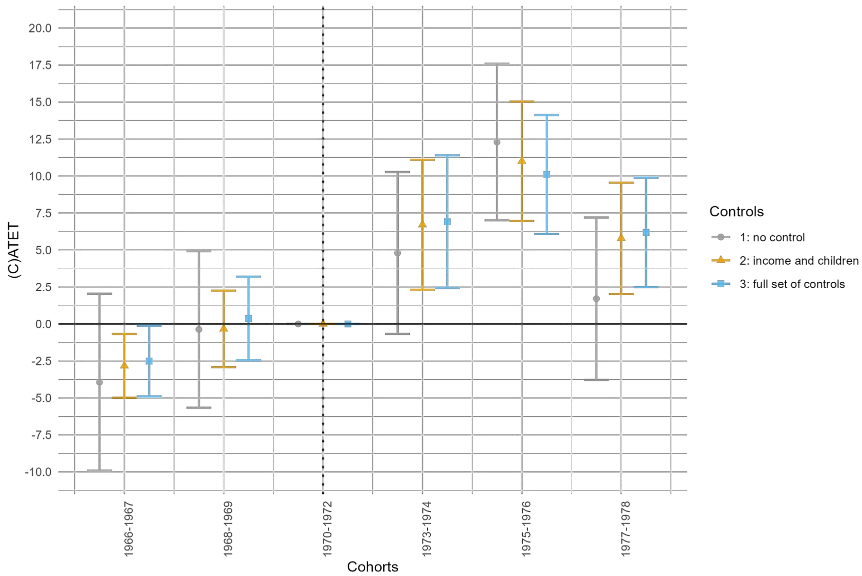


Figure 1A9: Male employment: families with at least two children

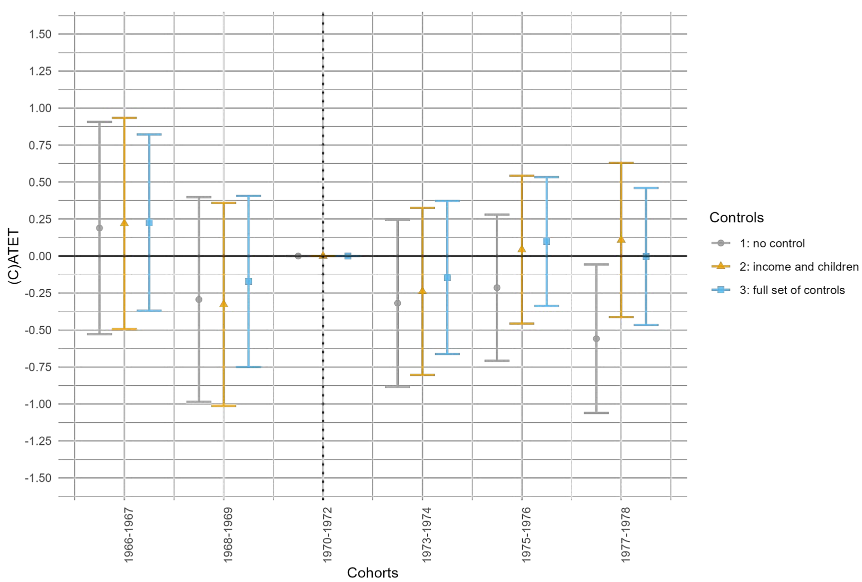


Figure 1A10: Male employment: families with fewer than two children

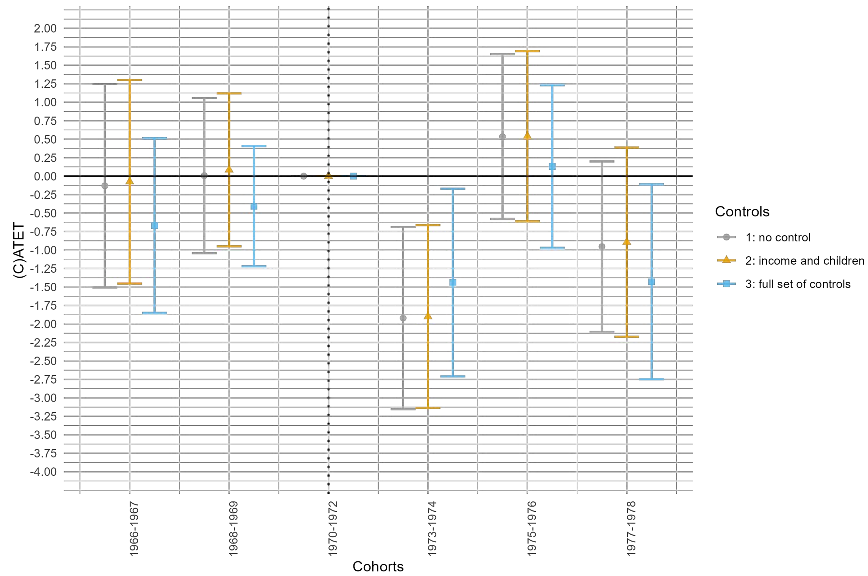


Figure 1A11: Male salary: families with at least two children

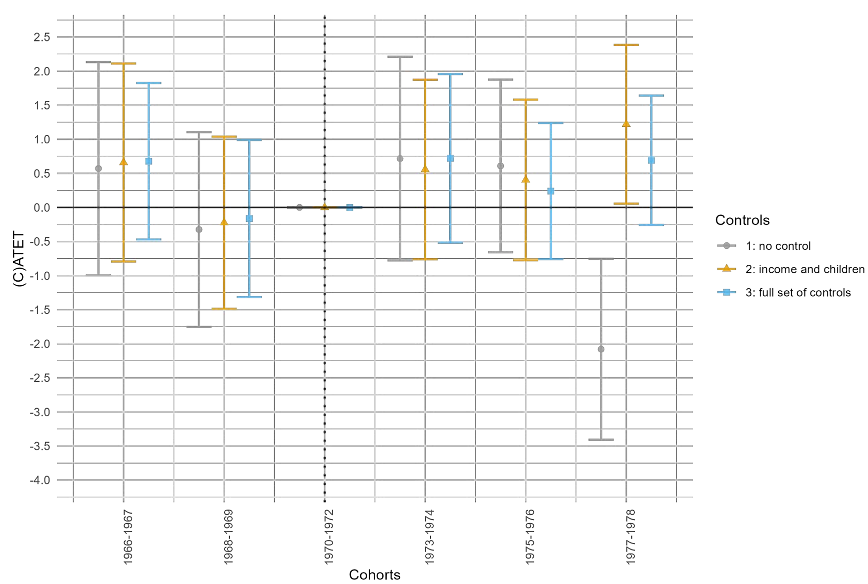


Figure 1A12: Male salary: families with fewer than two children

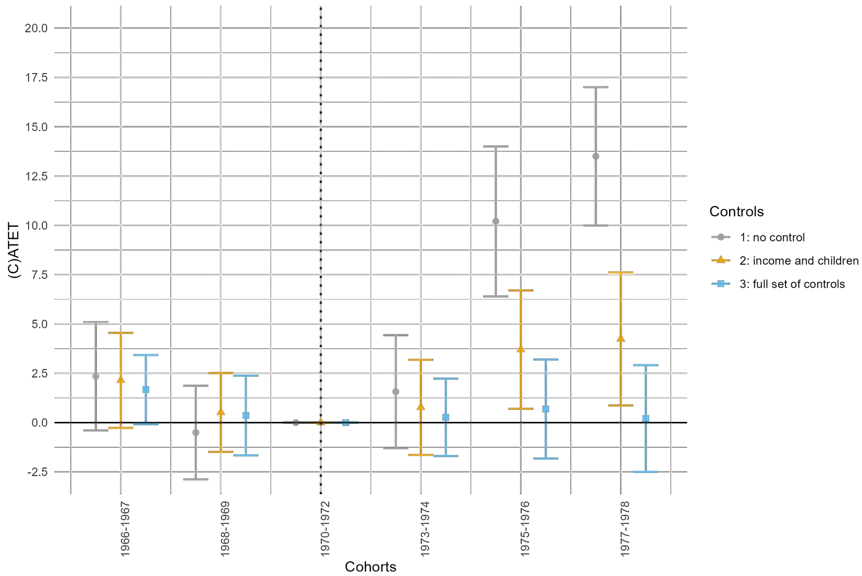


Figure 1A13: Female salary: families with at least two children

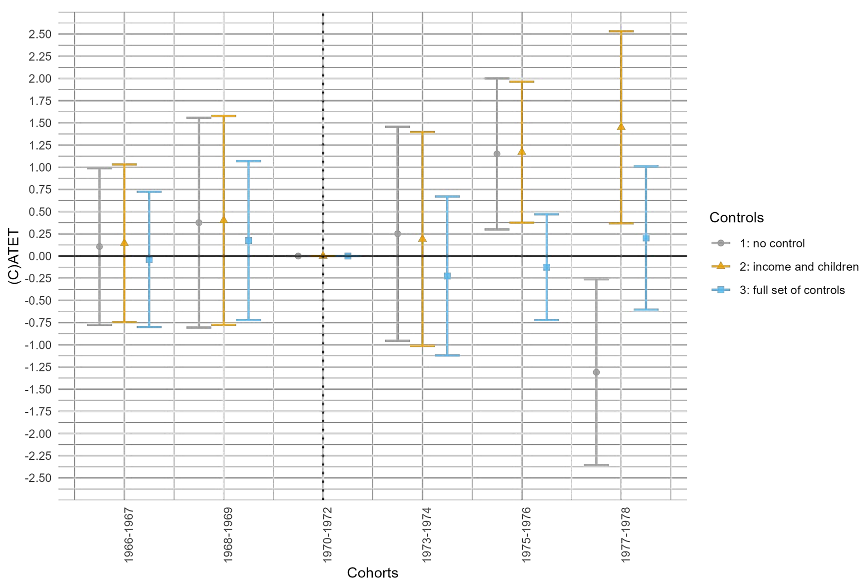


Figure 1A14: Female salary: families with fewer than two children

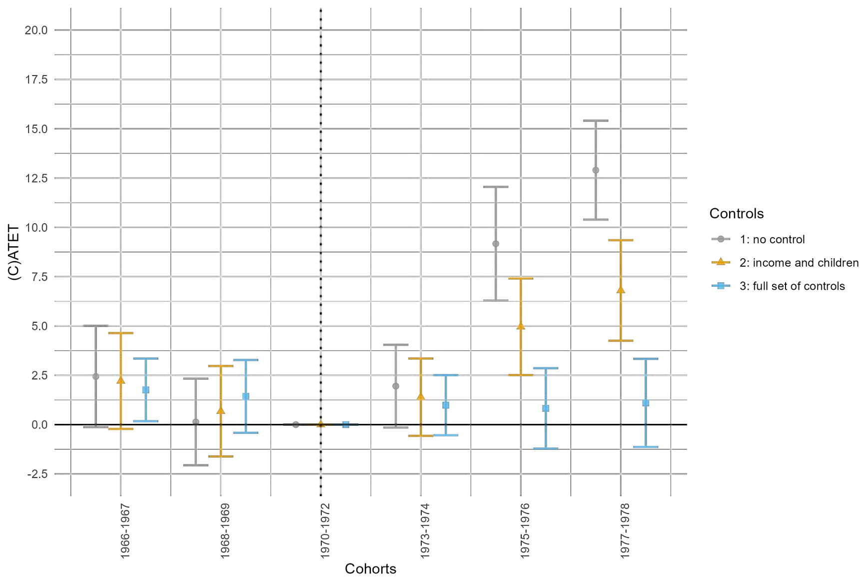


Figure 1A15: House size: families with at least two children

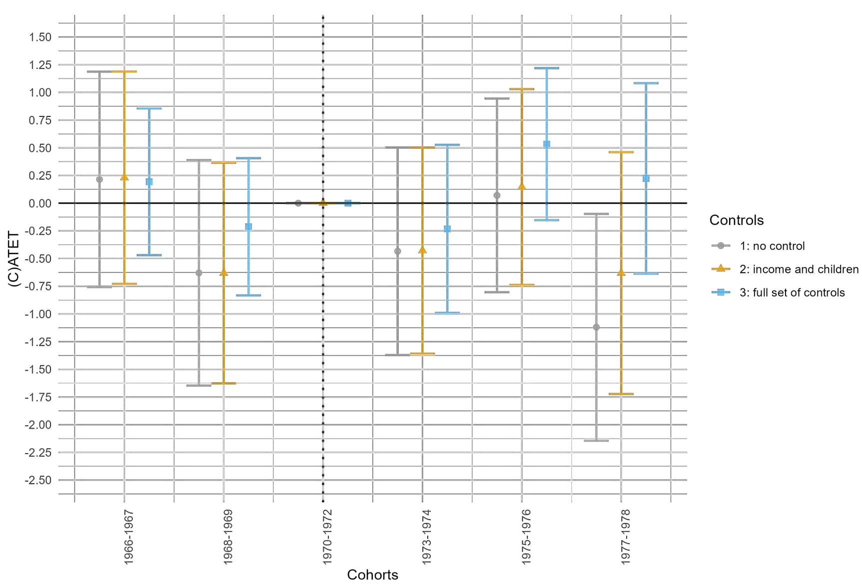


Figure 1A16: House size: families with fewer than two children

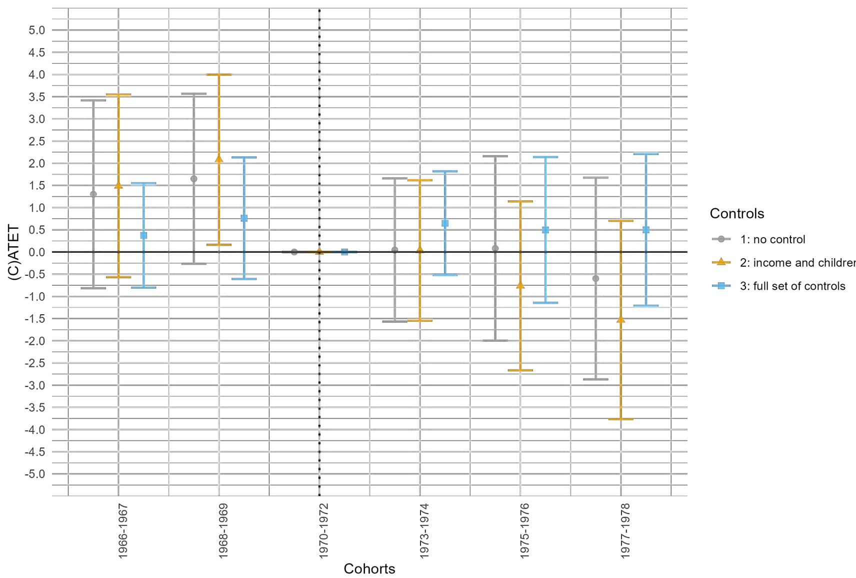
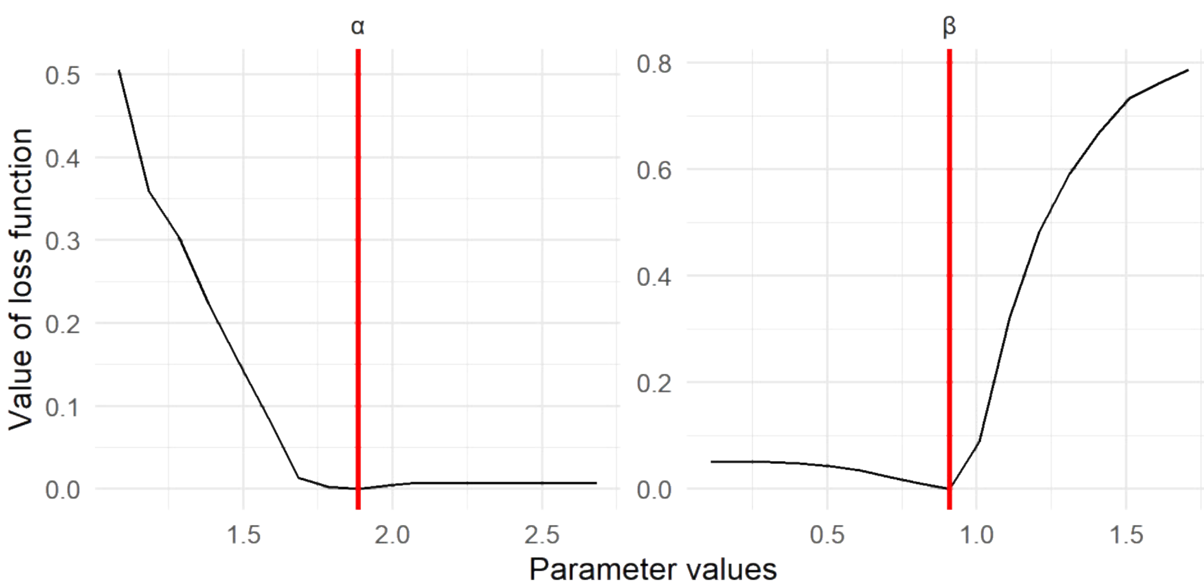


Figure 1A17: Loss function in the neighborhood of the estimated parameter values



Note: the figure shows the loss function around the estimated parameter values responding to changing one of the parameters.

Table 1A4: Heterogeneity of fertility effects, families with at least two children, log-specification

Mother's birth cohort	Religious household		Mother is at home		Married		Household income	
	No	Yes	No	Yes	No	Yes	Top25%	Lower75%
Cohort 1966-1967	0.000329	-0.000649	-0.000399	-0.00345	-0.00192	-0.000641	0.00166	0.000867
	(0.000914)	(0.000888)	(0.000669)	(0.00214)	(0.00352)	(0.000690)	(0.00150)	(0.00140)
Cohort 1968-1969	0.00184	0.000108	0.000150	0.00307	0.00582	-0.000367	0.00261	0.00161
	(0.00119)	(0.000883)	(0.000779)	(0.00229)	(0.00553)	(0.000685)	(0.00172)	(0.00164)
Cohort 1970-1972	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
Cohort 1973-1974	0.00338**	0.000354	0.00231*	-0.00217	-0.00388	0.00132	0.00335*	-0.00127
	(0.00169)	(0.00125)	(0.00124)	(0.00212)	(0.00371)	(0.00104)	(0.00200)	(0.00212)
Cohort 1975-1976	**0.000405**	**0.00483*** **	**0.00273** **	**0.00314**	**-0.00233**	**0.00350*** **	**0.00493* **	**0.00746*** **
	(0.00213)	**(0.00121)**	**(0.00119)**	**(0.00229)**	**(0.00385)**	**(0.00107)**	**(0.00250)**	**(0.00272)**
Cohort 1977-1978	0.00170	0.00598***	0.00421***	0.00433	-0.00146	0.00496***	0.00283	0.00493
	(0.00206)	(0.00204)	(0.00152)	(0.00277)	(0.00521)	(0.00151)	(0.00306)	(0.00318)
Observations	4,612	16,109	14,850	5,871	2,104	18,617	4,737	15,984
Adjusted R-squared	0.879	0.902	0.890	0.887	0.874	0.900	0.806	0.923
Controls	full	full	full	full	full	full	full	full
Mean of outcome	1.2201	1.231	1.1943	1.3137	1.2772	1.2228	1.2163	1.2328
S.D. of outcome	.1907	.1985	.1661	.2373	.2357	.191	.1818	.2018
Mean of family tax break increment	2.5641	2.198	2.4693	1.8352	1.7958	2.3437	4.573	1.418
S.D. of family tax break increment	2.4709	2.2533	2.2764	2.3485	2.0834	2.3326	2.8988	1.2035

Note: The table reports the coefficients and cluster-robust standard errors of the interaction terms of the cohort group dummies and the treatment variable in different regression specifications, using a full set of control variables. Cluster robust standard errors in parentheses. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 1A5: Heterogeneity of fertility effects, families with at least two children, log-specification

Mother's birth cohort	Father employed		Mother grad. high school		Youngest child is at most 3		Lives in a city	
	No	Yes	No	Yes	No	Yes	No	Yes
Cohort 1966-1967	0.00166	-0.000610	0.000812	-0.000347	-0.0252	-0.000679	-0.00130*	-2.48e-05
	(0.00162)	(0.000721)	(0.00120)	(0.000773)	(0.0227)	(0.000654)	(0.000763)	(0.000884)
Cohort 1968-1969	0.00336*	-1.49e-05	0.000591	0.000121	0.00324	-0.000193	-0.00125*	0.000651
	(0.00196)	(0.000790)	(0.00115)	(0.000837)	(0.0228)	(0.000713)	(0.000748)	(0.000898)
Cohort 1970-1972	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
Cohort 1973-1974	0.000807	0.000893	0.00138	0.00130	-0.00873	0.000934	0.000492	0.00142
	(0.00217)	(0.00110)	(0.00196)	(0.00124)	(0.0119)	(0.000935)	(0.00159)	(0.00116)
Cohort 1975-1976	**0.00359**	**0.00319*** **	**0.00103**	**0.00335*** **	**-0.00151**	**0.00228** **	**0.00345** **	**0.00376*** **
	(0.00295)	**(0.00109)**	**(0.00251)**	**(0.00125)**	**(0.0115)**	**(0.00107)**	**(0.00166)**	**(0.00132)**
Cohort 1977-1978	-0.00153	0.00449***	0.00145	0.00546***	0.000870	0.00280**	0.00302*	0.00516***
	(0.00335)	(0.00156)	(0.00333)	(0.00179)	(0.0117)	(0.00135)	(0.00181)	(0.00196)
Observations	2,869	17,852	9,044	11,677	2,667	18,005	9,315	11,406
Adjusted R-squared	0.958	0.878	0.944	0.837	0.817	0.918	0.925	0.871
Controls	full	full	full	full	full	full	full	full
Mean of outcome	1.2999	1.2178	1.2675	1.2031	1.3058	1.2156	1.2481	1.2159
S.D. of outcome	.2481	.1857	.226	.1706	.238	.1859	.212	.1853
Mean of family tax break increment	1.2995	2.433	1.1422	3.0222	2.5842	2.2443	1.7559	2.6198
S.D. of family tax break increment	1.6986	2.3568	1.2338	2.5361	2.7489	2.2334	1.9779	2.4433

Note: The table reports the coefficients and cluster-robust standard errors of the interaction terms of the cohort group dummies and the treatment variable in different regression specifications, using a full set of control variables. Cluster robust standard errors in parentheses. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$

Table 1A6: Heterogeneity of fertility effects, families with fewer than two children, log-specification

Mother's birth cohort	Religious household		Mother is at home		Married		Household income	
	No	Yes	No	Yes	No	Yes	Top25%	Lower75%
Cohort 1966-1967	0.00710	-0.00546	-0.00303	-0.00352	0.000838	-0.00545*	0.0135*	-0.00230
	(0.00680)	(0.00342)	(0.00368)	(0.00739)	(0.0121)	(0.00297)	(0.00696)	(0.00839)
Cohort 1968-1969	0.000693	-0.00273	0.000268	-0.00537	-0.0105	-0.00166	0.00474	-0.000599
	(0.00649)	(0.00392)	(0.00353)	(0.00711)	(0.0105)	(0.00312)	(0.00759)	(0.00887)
Cohort 1970-1972	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
Cohort 1973-1974	0.0135*	0.00351	0.0160**	-0.00585	0.0221**	0.00498	-0.00251	-0.00640
	(0.00741)	(0.00488)	(0.00639)	(0.00584)	(0.0105)	(0.00367)	(0.00764)	(0.0116)
Cohort 1975-1976	**0.0216***	**0.0279*****	**0.0252*****	**0.0215****	**0.0361****	**0.0225*****	**0.00896**	**-0.00102**
	(0.0114)	**(0.00839)**	**(0.00788)**	**(0.0107)**	**(0.0156)**	**(0.00730)**	**(0.0142)**	**(0.0166)**
Cohort 1977-1978	0.0248*	0.00539	0.00798	0.0128	0.00310	0.0113	-0.0354*	-0.0155
	(0.0137)	(0.00838)	(0.00865)	(0.0130)	(0.0165)	(0.00739)	(0.0196)	(0.0203)
Observations	2,909	7,334	8,118	2,125	2,586	7,657	3,059	7,184
Adjusted R-squared	0.617	0.650	0.604	0.706	0.606	0.649	0.574	0.682
Controls	full	full	full	full	full	full	full	full
Mean of outcome	0.6666	.6734	.63	0.821	.5986	.6964	.7027	0.6547
S.D. of outcome	.3421	.3387	.3309	.3289	.3755	.3227	.376	.3176
Mean of family tax break increment	2.4968	2.317	2.278	2.7218	2.1719	2.4418	3.7929	1.6255
S.D. of family tax break increment	1.4373	1.332	1.2432	1.7093	1.1865	1.4187	1.376	.4902

Note: The table reports the coefficients and cluster-robust standard errors of the interaction terms of the cohort group dummies and the treatment variable in different regression specifications, using a full set of control variables. Cluster robust standard errors in parentheses. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$

Table 1A7: Heterogeneity of fertility effects, families with fewer than two children, log-specification

Mother's birth cohort	Father employed		Mother grad. high school		Youngest child is at most 3		Lives in a city	
	No	Yes	No	Yes	No	Yes	No	Yes
Cohort 1966-1967	-0.0172	-0.000808	-0.00211	-0.000770	-0.0183	-0.00225	0.00727	-0.00400
	(0.0148)	(0.00283)	(0.00991)	(0.00334)	(0.0159)	(0.00196)	(0.00676)	(0.00340)
Cohort 1968-1969	-0.0270	-0.00103	-0.00395	-0.00176	0.0121	-0.00438**	-0.0117*	0.000150
	(0.0178)	(0.00339)	(0.0122)	(0.00338)	(0.0134)	(0.00202)	(0.00699)	(0.00354)
Cohort 1970-1972	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
Cohort 1973-1974	-0.00753	0.00755*	0.00891	0.00631*	-0.00377	0.0159**	0.000875	0.00724
	(0.0206)	(0.00384)	(0.0141)	(0.00379)	(0.00730)	(0.00725)	(0.00841)	(0.00459)
Cohort 1975-1976	**0.00410**	**0.0252*** **	**0.0171**	**0.0150***	**0.0118**	**0.0327*** **	**0.0438*** **	**0.0180** **
	(0.0193)	**(0.00703)**	**(0.0231)**	**(0.00797)**	**(0.0112)**	**(0.00768)**	**(0.0127)**	**(0.00857)**
Cohort 1977-1978	-0.000854	0.00954	-0.0157	0.00502	0.0216	-0.00784	0.0260**	0.00691
	(0.0370)	(0.00714)	(0.0194)	(0.00798)	(0.0159)	(0.00583)	(0.0121)	(0.00922)
Observations	1,211	9,032	3,209	7,034	1,398	6,582	3,401	6,842
Adjusted R-squared	0.787	0.622	0.746	0.607	0.401	0.541	0.122	0.627
Controls	full	full	full	full	full	full	none	full
Mean of outcome	0.5972	0.6807	.6266	0.688	.9563	.76	1.0346	.675
S.D. of outcome	.3621	.3356	.3136	.3475	.2389	.1573	.616	.3446
Mean of family tax break increment	1.9122	2.4298	1.5145	2.6935	3.3178	2.282	2.0145	2.499
S.D. of family tax break increment	1.1132	1.3857	.6025	1.4341	1.825	1.3132	1.2091	1.398

Note: The table reports the coefficients and cluster-robust standard errors of the interaction terms of the cohort group dummies and the treatment variable in different regression specifications, using a full set of control variables. Cluster robust standard errors in parentheses. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$

Chapter 2

Charitable Behavior and Public Intervention: a Survey Experiment (joint work with Francesca Leombroni)

Abstract

In this paper, we measure the extent of charitable behavior crowding out public intervention and how this phenomenon affects the welfare of the poor. To achieve this objective, we collect novel survey data on a representative sample of the U.S. adult population. In the survey, respondents are asked to go through several hypothetical scenarios, constructed on the basis of a simple model of public good contribution to learn about their preferences and expectations regarding donations and taxation. We find that when donations are available, government expenditure on the poor is lower in equilibrium. Yet, households in need are better off due to disproportionately higher donations. Therefore, in our setting, private charity crowds out public intervention only to a limited extent, affecting equilibrium-level taxes only slightly. We also estimate the structural parameters of preferences in our sample and find that individuals assign a sizable weight to both the utility of the poor and to the act of donating itself. The large contribution of the latter component rationalizes why taxation alone cannot fully compensate for the absence of donations.

Keywords: Altruism, Charity, Donation, Public Good, Anti-Poverty, Welfare

JEL Classification: D1, D8, H3, H4, I3

We thank our supervisors, Pamela Giustinelli, Gianmarco Ottaviano, Joseph-Simon Görlach, Alexia Delfino, Carlo Devillanova, and Francesco Billari for the excellent guidance, and professors Thomas Le Barbanchon, Giacomo Battiston, Stefano Fiorin, Salvatore Nunnari, Nicola Pavoni, Guido Tabellini, and Fernando Vega-Redondo, and for their insightful comments. We are also thankful to the participants of the Bocconi University reading groups and seminars for their valuable comments, and to the Ph.D. School of Bocconi University for providing financial support for this project.

2.1 Introduction

Charitable behavior plays a vital role in nearly all societies. For instance, donations account for more than 2% of the United States GDP (Andreoni and Payne, 2013), and more than 40% of U.S. households are involved in volunteering activities (Charities Aid Foundation, 2019). Similarly to taxation, private charity is a form of contribution to the public good. As such, the activity of charitable organizations is, to some extent, a substitute for public intervention. This is particularly true for areas such as poverty reduction, targeted by both charitable organizations and the public sector.[1] Some evidence of substitution between private charity and public intervention emerges from a cross-country comparison: among Western OECD countries, those that are characterized by a larger size of the government tend to show a lower prevalence of charity.[2]

The existence of some degree of substitution between charitable giving and public intervention has often been investigated in the literature, although most contributions have focused on one direction, that is whether government intervention could affect private donations through tax deductions (see for instance Schiff, 1985; Duncan, 1999; Brooks, 2000; Simmons and Emanuele, 2004; Garrett and Rhine, 2010; Bredtmann, 2016 and Peloza and Steel, 2005 for a meta-analysis of the estimates of the price elasticity of individual donations in the literature).

Whether the opposite direction is also relevant, that is whether the supply of private charity affects the extent of public intervention, has received considerably less attention. Becker and Lindsay (1994) and Sav (2012) find evidence for partial crowding out in the funding of US higher education; Heutel (2014) finds no evidence of private donations crowding out government grants to charities (while confirming that government grants crowd in private donations) while Werfel (2018) provides evidence that individuals are less likely to support higher taxation when informed of the size of charitable contributions in society.

While most contributions agree that the crowding out is not one-to-one in either direction,

[1] In the U.S., 35% of the donations as of 2017 were directed towards organizations in health, education, and human services, while more than 30% targeted religious organizations, most of which are also involved in poverty relief activities according to the nonprofit organization Charity Navigator https://www.charitynavigator.org/, accessed 06/01/2022.

[2] See for instance OECD (2021), Charities Aid Foundation (2019).

ruling out perfect substitution[3], answers concerning the size and even the direction of the relationship between charitable giving and public intervention are still discordant. Identifying a causal effect in either direction is difficult using observational data, as charitable giving does not happen in a vacuum; it is affected heavily by several unobserved confounders and equilibrium mechanisms.

In this paper we set to contribute to this debate with a survey experiment, by providing a causal estimate of the degree of crowding out in both directions. We identify and measure the impact of an increase in public intervention on private donations, and of its opposite, namely the effect of 'switching off' donations on tax preferences. To build the survey experiment, we rely on a simple framework that enriches the traditional models of public good contribution (Becker, 1974; Bergstrom et al., 1986) with elements that are typical of the more recent literature that investigates the determinants of private charity, such as impure altruism and reputational concerns (Andreoni, 1988, 1990; Bénabou and Tirole, 2006; Katz and Rosenberg, 2005).

We present a sample of 380 U.S. respondents[4] with hypothetical but realistic scenarios in which we vary the availability of donations and the respondents' gross income to measure the change in their taxation preferences. We also elicit respondents' donation choices at different taxation levels and their expectations about the average level of donations in society. Based on their answers, we simulate equilibrium outcomes in our public good model setting. The results of this exercise allow us to compare equilibrium tax rates and the welfare of the poor with and without donations available. Additionally, we conduct heterogeneity analyses based on the respondents' characteristics and elicited preferences and link them to their in-survey preferred levels of taxes and donations.

We find that government expenditure on the poor is lower when donations are available, but households in need are still better off due to disproportionately higher donations. In our setting, private charity crowds out public intervention only to a limited extent; equilibrium tax rates in the no-donations scenario are not high enough to compensate for the lack of

[3] Among the possible explanations for the lack of a complete crowding out, Eckel et al. (2005) highlight how individuals do not fully internalize their contribution to the government finances and, therefore, indirectly to the public good through taxation (fiscal illusion).

[4] The sample was selected by the survey company Prolific to be representative of the population of the United States according to gender, age bracket, and ethnicity.

private charity, suggesting that people are also driven by the direct utility of the act of donating (*warm glow*). We confirm this finding by retrieving the structure of preferences which generates the behaviors we elicit in the hypothetical scenarios with individual-level estimates of the main utility parameters of our model (generosity, warm glow and weight of reputational concerns). While the estimated average generosity in our sample is higher than the weight of the direct utility of donations, the latter component is positive and relatively large in magnitude: direct utility from donations (the *warm glow* component) is assigned an average weight of nearly 3% of the utility of one's own consumption, compared with a value of 6.6% for the weight of the utility of the poorest members of society (the generosity component). Reputational concerns are instead assigned a lower weight, at 0.2%. Overall, our results suggest that the widespread availability of private charity in the United States plays a pivotal role in alleviating poverty, which government intervention cannot substitute for due to the structure of voters' preferences.

The paper proceeds as follows: Section 2.2 describes the model and derives some predictions, Section 2.3 describes the survey and the characteristics of the sample, Section 2.4 presents our results, while Section 2.5 concludes.

2.2 A simple framework

We now provide an overview of a simple theoretical framework to guide the construction of the hypothetical scenarios in our survey. We are interested in the redistribution effects of the availability of charity in general equilibrium, where households form expectations over the charitable behavior of others which in turn affect their own ideal taxation levels. The latter are then reflected in the social choice of taxation with and without donations: if the expectations of households were substantially higher than the actual donation behavior of others, each household might prefer a suboptimally low level of taxes, leaving the poorest households potentially worse off when donations are available. However, if donations affected individual utility not only through their contribution to the benefit of poor households, but also directly (through a sizable enough *warm glow* component), incentives for charity could be largely beneficial for the poor.

We build a simple public good model where households derive utility from their own consumption, the public good, and their contribution to the public good[5]. In our setting, the society is composed of N_p households earning positive income, and N_z households earning zero income, and the public good is defined as the financial support accruing to zero-income households. Positive-income households can contribute to the public good through two channels: taxation and private donations. Their own donations, the expected level of donations, and the welfare of the households in need enter the optimal consumption choice of households, so that the value of the problem will depend on the level of taxes, allowing to pin down the preferred level of taxation for each household. The utility function of the households reflects inequity, warm glow, and reputational concerns. Finally, a neutral government sets the tax rate in accordance with the preferences of the median voter.

2.2.1 Baseline case: no charity

We first describe a simpler version of our model, where households earning a positive income can contribute to the public good only through taxation. Positive-income households maximize their utility, given by:

$$u(c_i, b) = log(c_i) + \alpha_i log(b).$$

where c_i is consumption, α_i is the *generosity* or *pure altruism* parameter, representing the weight of the public good in the utility function, and b is the public good, i.e. the transfer accruing to each household-in-need:

$$b = \frac{1}{N_z}(\tau - \underline{\tau})W,$$

where $W = \sum_{i=1}^{N_p} w_i$ s the total wage mass, τ is the tax rate selected by the government to support the households in need, and $\underline{\tau}$ is the fraction of total taxes devoted to the upkeep of the government, which is fixed at 20% of the gross wage.

Positive-income households cannot consume more than their net income w_i, resulting in

[5] Adapting the frameworks of Andreoni (1988) and Duncan (1999).

the following budget constraint:

$$c_i \leq (1-\tau)w_i.$$

Finally, zero-income households are characterized by the following utility function:

$$u(b) = log(b)$$

Solving for the preferred tax rate

In the baseline case where no charity is allowed, consumption is always set at the maximum available level $c_i = w_i(1-\tau)$. We can therefore solve for the preferred tax rate of each positive-income household, τ_i^*, by maximizing the value of the problem,

$$V(w_i, W, \tau, N_z) = log\big((1-\tau)w_i\big) + \alpha_i log\left(\frac{1}{N_z}(\tau - \underline{\tau})W\right).$$

This implies the following first-order condition and optimal taxation:

$$\frac{\partial V(w_i, W, \tau, N_z)}{\partial \tau} = -\frac{1}{(1-\tau_i^*)} + \frac{\alpha_i}{(\tau_i^* - \underline{\tau})} = 0 \tag{2.1}$$

$$\tau_i^* = \frac{\alpha_i + \underline{\tau}}{1+\alpha_i}. \tag{2.2}$$

Deriving the preferred tax rate τ_i^* with respect to the degree of inequity aversion α_i, we obtain:

$$\frac{\partial \tau^*}{\partial \alpha_i} = \frac{1-\underline{\tau}}{(1+\alpha_i)^2} > 0, \tag{2.3}$$

meaning that the preferred tax rate is increasing in inequity aversion.

Zero-income households instead simply wish to maximize the amount of public good, and therefore prefer the highest possible tax rate (which we assume bounded above by some amount τ^H). When donations are not allowed, α_i is pinned down by preferred taxes:

$$\alpha_i = \frac{\tau_i^* - \underline{\tau}}{1-\tau_i^*}$$

2.2.2 The government's problem

We close the model by solving the government's problem. The government knows the preferences of each household and sets the tax rate τ to match as closely as possible the preferences of the median voter. It, therefore, minimizes the sum of absolute deviations from each citizen's preferred tax rate:

$$\tau = \arg\min_{\tau' \geq \underline{\tau}} \sum_{i=1}^{N_p+N_z} |\tau_i^* - \tau'|$$

This expression is indeed minimized by choosing the median of the population's preferences, which is equivalent to the preferred tax rate of the median voter[6].

2.2.3 Complete case: reintroducing private charity

We now reintroduce private charity in the picture and present the complete framework. Positive-income households can contribute to the welfare of the households in need, both paying taxes and engaging in private charity. Their objective function is now:

$$\begin{aligned} u(c_i, d_i; d_{-i}, b) = {} & log(c_i) + \gamma_i log(1+d_i) + \alpha_i log(\mathbb{E}[b|\tau]) \\ & + \eta_i \Big(log(1 + (d_i - \mathbb{E}[d_{-i}|\tau])^2) \cdot 1\big[d_i \geq \mathbb{E}[d_{-i}|\tau]\big] \\ & - log\big(1 + (\mathbb{E}[d_{-i}|\tau] - d_i)^2\big) \cdot 1\big[d_i < \mathbb{E}[d_{-i}|\tau]\big] \Big), \end{aligned}$$

where, in addition to own consumption and the public good, utility depends on the amount of own donations (d_i) and on the deviation of own donations from the prevailing level of donations in the society, respectively weighted by γ_i, the *warm glow* parameter, that regulates the importance of one's own contribution to the public good in the utility function, and η_i, that is the weight of reputational concerns, or equivalently the cost of deviating from the social norm[7].

[6] Using a quadratic loss function would result in selecting the average of the ideal tax rates instead of the median.

[7] The role of reputational concerns in this context has been emphasized, for instance, by Bénabou and Tirole (2006)

The budget constraint is also modified to include donations:

$$c_i \leq w_i(1-\tau) - d_i.$$

It is important to highlight that now, differently from the simplified case with no donations, agents have to form expectations over the private charitable contributions of others. Indeed, they get utility (disutility) from both positive (negative) deviations between their own donations and the average societal level of donations, and from the total amount of public good, which is composed of taxes, own donations, and the not-yet-determined donations of other positive-income households in the society:

$$\mathbb{E}[b|\tau] = \frac{1}{N_z}\Bigg((\tau - \underline{\tau})W + d_i + \mathbb{E}\big[\sum_{j\neq i} d_j|\tau\big]\Bigg).$$

We can simplify this expression by allowing individuals to only form beliefs on the average level of donations in the society[8] conditional on the level of taxes:

$$\mathbb{E}[b|\tau] \approx \frac{1}{N_z}\Bigg((\tau - \underline{\tau})W + d_i + (N_p - 1)\mathbb{E}[d_{-i}|\tau]\Bigg),$$

The usual assumption of perfect rationality would require that agents' guesses matched the realized outcome. We choose not to make any assumption on the structure of beliefs and instead use our survey to test whether individuals hold accurate beliefs. For the sake of completeness, we will, however, present a brief analysis of the benchmark case, characterized by a representative household with rational expectations.

Benchmark case: representative agent with correct beliefs

In the benchmark case, the representative agent maximizes her utility while knowing that everybody else solves an identical problem. Since, in the benchmark case, all individuals have the same preferences and budget, we can treat expected donations such that the agent has correct beliefs about expected donations as if they were the solution to the individual

[8]Excluding themselves, which however, is of little importance for a big enough number of households

optimization problem,

$$V(w_i, W, \tau, N_z, N_p) = \max_{d_i} log\big((1-\tau)w_i - d_i\big) + \gamma_i log(1+d_i)$$
$$+ \alpha_i log\left(\frac{1}{N_z}\left((\tau - \underline{\tau})W + N_p \mathbb{E}[d_{-i}|\tau]\right)\right).$$

Since everybody is the same, the individual level of donation (d_i) coincides with the average societal level ($\mathbb{E}[d_{-i}|\tau]$), implying that the reputational concern term does not play any role. However, the individual decision maker, not internalizing this, still solves for her own level of donations as if her contribution was only infinitesimal for the overall benefit accruing to the unemployed so that the benefit term ($\mathbb{E}[b|\tau]$) is taken as given and does not appear in the first order condition.

The optimal level of donations thus results from maximizing the following first-order condition

$$\text{w.r.t. } d_i : \frac{1}{w_i(1-\tau) - d_i} = \frac{\gamma_i}{1+d_i} \tag{2.4}$$
$$d_i^* = max\left\{\frac{\gamma_i w_i(1-\tau) - 1}{1+\gamma_i}, 0\right\}. \tag{2.5}$$

The preferred level of donations is positive whenever:

$$\gamma_i \geq \frac{1}{w_i(1-\tau)},$$

meaning that there will be a positive level of donations in society whenever the level of warm glow is above a certain threshold (equal to at least the inverse of the net wage).

From 2.5, the optimal level of consumption can also be retrieved as:

$$c_i^* = min\left\{\frac{w_i(1-\tau) - 1}{1+\gamma_i}, w_i(1-\tau)\right\}.$$

Assuming a high enough level of warm glow, we can plug back the values of the interior solution to obtain the value of the problem in the benchmark case:

$$V(w_i, W, \tau, N_z, N_p) = log\left(\frac{w_i(1-\tau)-1}{1+\gamma_i}\right) + \gamma_i log\left(\frac{\gamma_i w_i(1-\tau)-1}{1+\gamma_i}\right)$$
$$+\alpha_i log\left(\frac{1}{N_z}\left((\tau-\underline{\tau})W + N_p\frac{\gamma_i w_i(1-\tau)-1}{1+\gamma_i}\right)\right),$$

from which we can compute the preferred tax rate of household i, τ_i^*, by finding the tax rate maximizing the value of her problem. Considering that $W = N_p w_i$, the first order condition with respect to τ,

$$\frac{\partial V(w_i, W, \tau, N_z, N_p)}{\partial \tau} = -\frac{w_i}{w_i(1-\tau)-1} - \frac{\gamma_i^2 w_i}{\gamma_i w_i(1-\tau)-1}$$
$$+\frac{\alpha_i w_i}{(1+\gamma_i)(\tau-\underline{\tau})w_i + \gamma_i w_i(1-\tau)-1} = 0$$

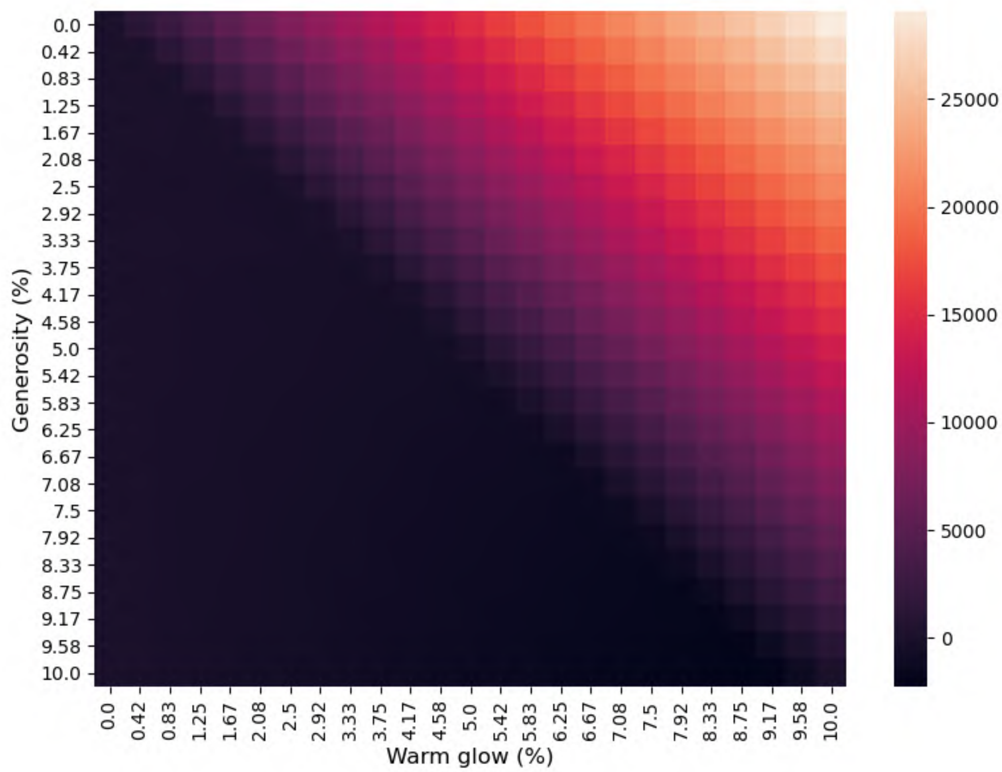
implies that:

$$\frac{1}{w_i(1-\tau)-1} + \frac{\gamma_i^2}{\gamma_i w_i(1-\tau)-1} = \frac{\alpha_i}{(1+\gamma_i)(\tau-\underline{\tau})w_i + \gamma_i w_i(1-\tau)-1},$$

which can be solved numerically for the preferred tax rate, τ^*.

Solving the problem for different values of the preference parameters governing inequity aversion (α) and warm glow (γ), we can infer their effect on preferred taxes, donations, and the level of benefits households-in-need receive. Preferred tax rates increase in inequity aversion but decrease in warm glow. Donation rates increase in warm glow, but as an individual's donations do not contribute to reducing inequity, inequity aversion does not affect optimal donation rates—consequently, total benefits for the poor increase in both dimensions. However, let's compare it with the scenario where donations are not allowed. We can see that for regions with somewhat high inequity aversion and warm glow, total benefits would decrease by allowing donations in society. So depending on the preferences representing social values, allowing donations might or might not benefit those that they are designed to target, even if donation expectations are correct, as a consequence of the equilibrium brought by the taxes set by the politician.

Figure 2.1: Benefit accruing to each zero-income household with versus without donations in the RERA benchmark for different levels of generosity and warm glow



Notes: authors' calculations based on solving the problem of the representative agent with rational expectations and correct beliefs about the level of average donations in society. Tax rates are constrained from below at $\underline{\tau} = 0.2$ representing a mandatory minimum level of taxation covering other government expenditures, and household income is set at \$60,000.

General case

We now move away from the representative agent, rational expectations benchmark; that is, we allow for heterogeneous household-level utility parameters and income and household-specific expectations. We treat these as model parameters without imposing any assumption and derive the optimal level of donations again for an employed household i in this general case. Analogously to the RERA case, households maximize their utility with respect to the donation level d_i:

$$
\begin{aligned}
V(w_i, W, d_{-i}, \tau, N_z, N_p) = & \max_{d_i} log(c_i) + \gamma_i log(1 + d_i) \\
& + \alpha_i log\Big(\frac{1}{N_z}\big((\tau - \underline{\tau})W + N_p \mathbb{E}[d_{-i}|\tau]\big)\Big) \\
& + \eta_i \Bigg(1\big[d_i \geq \mathbb{E}[d_{-i}|\tau]\big] \cdot log(1 + (d_i - \mathbb{E}[d_{-i}|\tau])^2) \\
& - 1\big[d_i < \mathbb{E}[d_{-i}|\tau]\big] \cdot log\big(1 + (\mathbb{E}[d_{-i}|\tau] - d_i)^2\big) \Bigg),
\end{aligned}
$$

resulting in the following first-order condition for the optimal level of donations:

$$
\text{w.r.t. } d_i : \frac{1}{w_i(1-\tau) - d_i} = \frac{\gamma_i}{1 + d_i} + \frac{2\eta_i(d_i - \mathbb{E}[d_{-i}|\tau])}{1 + (d_i - \mathbb{E}[d_{-i}|\tau])^2} \tag{2.6}
$$

from which the optimal level of donations can be retrieved as the solution to the third-degree equation:

$$
\begin{aligned}
0 = & \Big(1 + \gamma_i + 2\eta_i\Big) d_i^3 \\
& + \Big(1 - 2\mathbb{E}[d_{-i}|\tau](1 + \gamma_i + \eta_i) - w_i(1-\tau)(\gamma_i + 2\eta_i)\Big) d_i^2 \\
& + \Big(1 - \gamma_i - 2\eta_i w_i(1-\tau) + \mathbb{E}^2[d_{-i}|\tau] + 2\mathbb{E}[d_{-i}|\tau]\big(1 - \eta_i + (\eta_i + \gamma_i) w_i(1-\tau)\big)\Big) d_i \\
& + \Big(1 - \gamma_i w_i(1-\tau) + 2\eta_i w_i(1-\tau)\mathbb{E}[d_{-i}|\tau] + \mathbb{E}^2[d_{-i}|\tau](1 - \gamma_i w_i(1-\tau))\Big)
\end{aligned}
$$

This model setup constitutes the baseline for the hypothetical scenarios we use in our survey. In the survey, we ask respondents to choose their amount of donations conditional on different levels of income (w_i) and taxes (τ), along with the expected value of donations in society given the level of taxes $\mathbb{E}[d_{-i}|\tau]$. Furthermore, we also elicit respondents' preferred level of taxes for two types of society: one where donations are allowed and one where taxation is the only source of support for the households in need. This approach enables us to predict the effect of donations on the equilibrium level of taxes and the welfare of the households in need.

2.3 Survey experiment

To investigate the causal relationship between donations, taxation, and poverty, we implement a survey experiment in the spirit of the model detailed earlier. The data are provided by a sample of 380 U.S. adult residents selected through the professional survey company Prolific[9] to represent the population at large in terms of age, gender, and ethnicity. The survey requires approximately 40 minutes to complete and asks respondents to go through three main sections. The full text of the questionnaire is available upon request.

To provide context for the survey, we first present some aggregate descriptive evidence for the United States regarding the interrelatedness of charity, local taxation, and poverty at the county level. Table 2.1 and Figure 2.2 show donation rates, property taxes, and poverty rates as relevant proxy measures of these concepts with meaningful variance at the county level.[10] On average, people donate nearly 1.8% of their adjusted gross income, pay approximately 9.7 dollars on a thousand dollars worth of real estate, while the county-level average poverty rate is slightly below 16%. We can see a substantial geographic variation in the country. A look at pairwise correlations reveals that donations are negatively associated with local property taxes (-0.19) and positively with the poverty rate (0.07), while poverty correlates negatively to tax rates (-0.33).

[9] www.prolific.co.

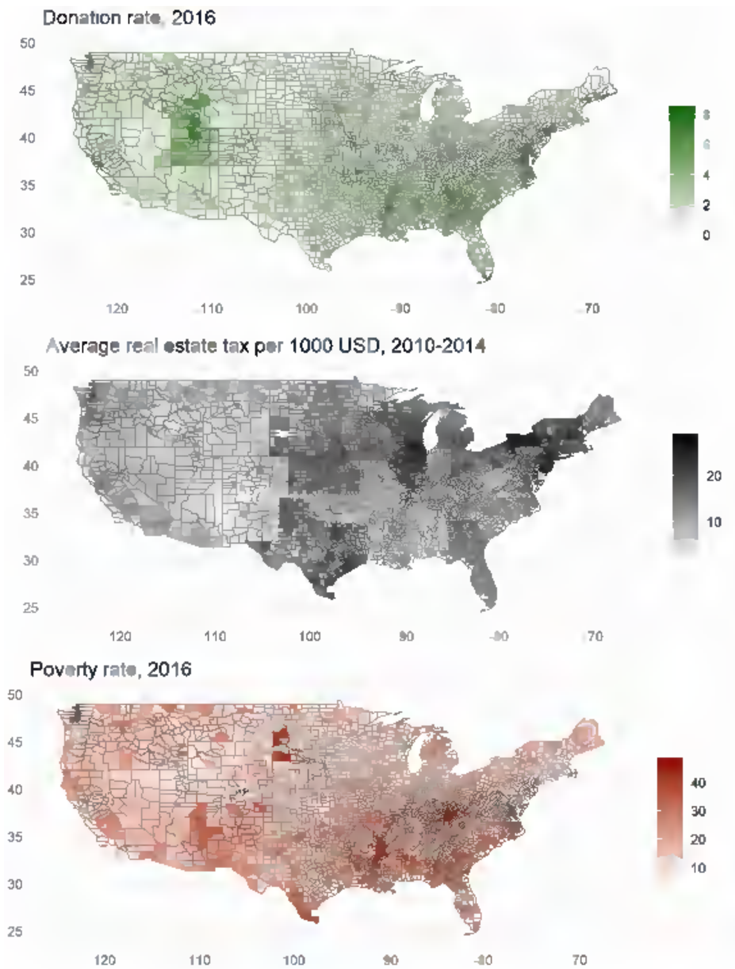
[10] We measure donation rates at the county level as the total charitable contributions reported in tax filings divided by the total adjusted gross income estimated by the Statistics of Income division, available at the Internal Revenue Service website: https://www.irs.gov/statistics/soi-tax-stats-county-data-2016, accessed 18/05/2021. For local taxation, we employ the five-year average (for 2010-2014) of the property taxes per $1000 worth of real estate collected by the National Association of Home Builders available at: https://www.nahbclassic.org/generic.aspx?genericContentID=250239&fromGSA=1, accessed 27/04/2021. Finally, poverty rates are based on the Annual Social and Economic Supplements of the 2016 Current Population Survey (CPS ASEC). Available at: https://www.census.gov/library/publications/2017/demo/p60-259.html, accessed 10/08/2021.

Table 2.1: Descriptive statistics of the key variables

	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Donation rate in 2016	3,129	1.816	0.802	0.000	1.275	2.193	8.552
Property taxes 2010-2014	3,129	9.700	4.635	1.085	6.124	12.503	29.001
Poverty rate in 2016	3,129	15.864	6.263	3.400	11.400	19.100	48.600

Note: The table reports the descriptive statistics for the main variables of the analysis dataset, which is collected by the authors from the following sources. Charitable tax deductions of 2016 are accessed via the website of the Internal Revenue Service, maintained by the Statistics of Income division. Data on property taxes from 2010-2014 are collected by the National Association of Home Builders. Poverty rates are calculated based on the CPS ASEC data by the U.S. Census Bureau.

Figure 2.2: The geographic variation in the key variables



Note: Figure shows the authors' calculations based on the following publicly available datasets. The donation rate for 2016 is calculated as donations over adjusted gross income. The data are accessed via the website of the Internal Revenue Service, maintained by the Statistics of Income division. Data on property taxes from 2010-2014 are collected by the National Association of Home Builders. Poverty rates are calculated based on the CPS ASEC data by the U.S. Census Bureau. Alaska and Hawaii are omitted from the map but are part of the dataset.

In order to shed light on the causal mechanisms behind these relationships, we ask respondents to imagine their preferences on taxation and their expectations and behavior concerning donations in six hypothetical scenarios. This first section of the survey replicates the game's structure presented in Section 2.2. In each scenario, respondents are asked to take up the roles of employed, income-earning households and to indicate their preferred contribution to the welfare of zero-income households (described as 'households in need'), which account for 15% of the overall population. In three out of six scenarios, respondents can contribute to the welfare of the households in need through additional taxation collected for that purpose or through private donations (complete scenarios). In addition to their own behavior and preferences, they are also asked to state how much they expect other employed households to donate. In the remaining three scenarios, individuals can only contribute through additional taxation, while donations are not allowed (no-charity scenarios). Within each category (with and without charity), scenarios differ according to the level of income accruing to the respondent's household: low, middle, or high.

The last two sections of the survey respectively ask for demographic information such as gender, age, ethnicity, state of birth, education level, occupation, income category and religion, and respondents' real-life charitable behavior (volunteering experiences and private donations) and elicit political preferences, personal attitudes towards economic redistribution and charitability, time- and risk-related preferences.

2.3.1 Descriptive statistics of the survey sample

Consistently with the 2019 American Community Survey estimates[11], our sample contains slightly more female than male respondents (51%) and is predominantly white (69%). Concerning age, the most represented category is the 58+ constituting 30% of the respondents, while the remaining categories all contain between 16 and 19% of the sample. Moving on to variables not targeted by the representative sample requirements, high-income households (that we defined as reporting a gross income of more than $90,000, consistently with the hypothetical scenarios of the first section of the survey) are over-represented in the sample, 40% versus 31% in the U.S. population.[12] The fraction of middle-income households (reporting a gross income of between $50,000 and $90,000) is slightly under-estimated, representing 27% of our sample but 30% of the overall population. Finally, low-income individuals (reporting a gross income below $50,000) represent 33% of our sample versus 38% of the U.S. population.

Hence, unsurprisingly, the sample has a low fraction of individuals with less than a high school diploma and high school graduates (0.3% and 7.9% versus 10% and 28% in the overall

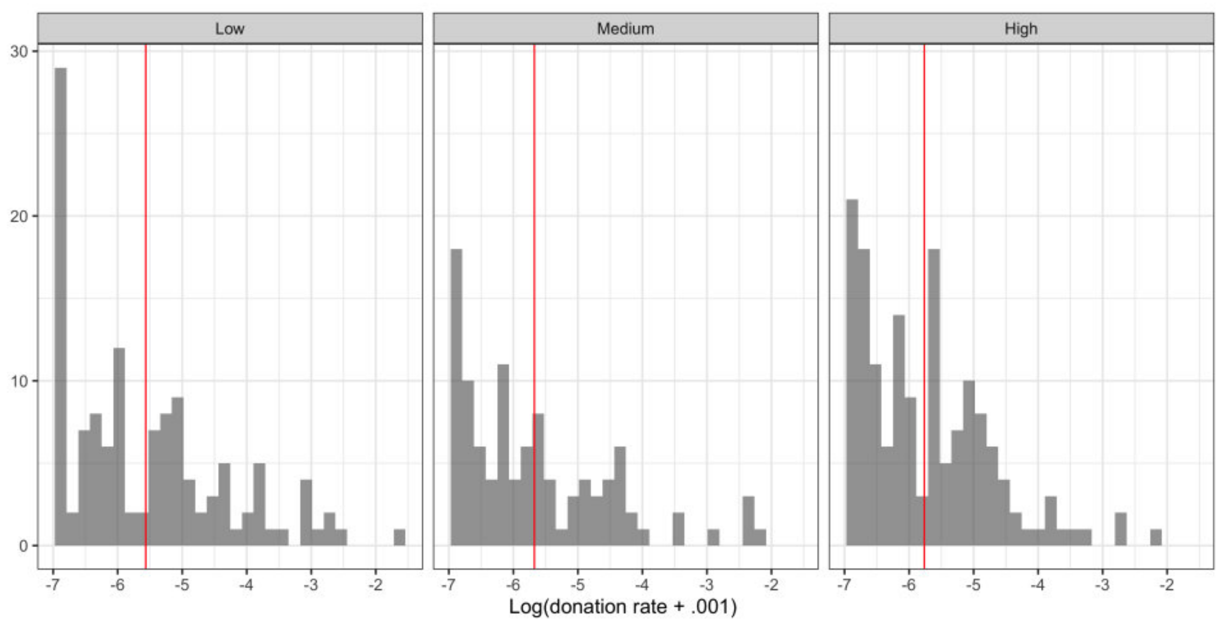
[11] Aggregate demographic information is available at: https://data.census.gov.

[12] Data on income brackets available at: https://censusreporter.org/topics/income/.

population). At the same time, more than 30% of respondents hold a master's or professional degree, versus 10.2% in the population at large. Concerning religion, more than 30% of the sample declared having no religious identity. Among those indicating some religious affiliation, the most prevalent creed is Catholicism (26% of the sample), followed by other Christian denominations (18%) and mainline Protestantism (11%).

Finally, concerning the reported patterns of charitable behavior, slightly less than 30% of the respondents report no experience with volunteering, and less than 15% have never engaged in monetary donations. More than half of the respondents volunteer occasionally and report having donated a few times. A sizable fraction (21% of the respondents) reports engaging in regular donations. The following graphs show the distribution of estimated yearly donations by income category. Although respondents from the lowest income category are more likely not to donate, the average donation rate slightly decreases with income.

Figure 2.3: Estimated in-life donation rate by income category



Note: We estimate donation rates in real life by combining survey responses on the frequency of donations and the average donation size. To compute total donations, we impute the middle value of the donation brackets available in the survey and multiply it by the reported number of yearly donations. To compute yearly income, we impute the middle value of the selected income bracket.

Table 2.2: Demographic characteristics of the sample

Category		Count	Fraction
Gender	Female	194	51%
	Male	186	49%
Age category	18-27	72	19%
	28-37	70	18%
	38-47	61	16%
	48-57	64	17%
	58+	113	30%
Ethnicity	White	264	69%
	Black	55	14%
	Other	61	16%
Education level	Less than high school degree	1	0%
	High school graduate	30	8%
	Some college but no degree	74	19%
	Bachelor or associate degree in college	130	37%
	Master's or professional degree	121	32%
	Doctoral degree	14	4%
Income category	Low income ($< \$50k$)	125	33%
	Middle income ($50k-$90k)	102	27%
	High income ($> \$90k$)	153	40%
Religion	No religious identity	120	32%
	Roman Catholic	96	26%
	Protestant (mainline)	41	11%
	Evangelical Protestant	21	6%
	Other Christian religion	68	18%
	Other non-Christian religion	34	9%
Frequency of volunteering	Never	105	28%
	Occasionally	202	53%
	At least once per month	46	12%
	At least once per week	27	7%
Frequency of donations	Never	56	15%
	Once	44	12%
	A few times	199	52%
	Regular donations	81	21%

2.3.2 Preferences and predicted behaviors in hypothetical scenarios

To analyze the effect of charity on economic redistribution, we analyze survey responses to the six hypothetical scenarios in the first section, where respondents are asked to report as truthfully as possible how much they would donate, what their preferred tax rate would be, and how much they would expect others to donate. The main components of each scenario are summarized in the table below.

Table 2.3: Scenarios description

Common elements						
Fraction of households-in-need	15%					
Baseline tax rate	20%					
Additional tax rate to support households-in-need	0%, 2.5%, 5%, 7.5%, 10%					
Elements differing across scenarios						
Donations allowed	Yes	Yes	Yes	No	No	No
Gross income	\$40k	\$60k	\$120k	\$40k	\$60k	\$120k
Tasks for respondents						
Choosing preferred additional tax rate	Yes	Yes	Yes	Yes	Yes	Yes
Selecting own donations	Yes	Yes	Yes	No	No	No
Declaring expected donation of the typical household	Yes	Yes	Yes	No	No	No

As shown in Table 2.3, respondents perform between one and three tasks in each scenario. First, for scenarios where charity is allowed, respondents are asked to select the dollar amount that they expect the typical middle-income household (where the middle income is set at \$60,000) to donate for each level of additional tax rate.[13] Secondly, in these scenarios, respondents are asked to state how much they would be willing to donate to support households in need, given each of the five levels of additional tax rates.[14] finally, they are asked to assign preference points across these five levels of additional tax rates (the table reports the options between 0% and 10%). Their preferred tax rate is then computed as a weighted average of their preferences. Afterward, they similarly provide taxation preferences for scenarios without donations available.

When facing these questions, respondents are explicitly reminded of the amount of benefits households-in-need would receive and the net income their own household would end up with, conditional on each tax level and their previous answers about donation expectations in society. For instance, we use built-in survey tools to calculate the implication of a tax level choice on total unemployment benefits, given the respondent's own expectations elicited

[13] In all the described scenarios, respondents are reminded that tax rates are flat and that donations cannot be deducted from their taxable income (i.e., they are subtracted from their net income).

[14] Dollar amounts are selected on a slider between a minimum of \$0 and a maximum of \$6,000.

earlier. This ensures that respondents do not need to engage in complicated calculation exercises and can express their preferences in a self-consistent manner.

2.4 Results

This section presents the main results obtained from the survey analysis. We first measure the extent of the substitution between taxes and donations in both directions (namely, the effect of taxes on preferred donations and the availability of charity on preferred taxes), and then we rationalize these results by retrieving the respondents' structure of preferences. To do so, we estimate the three main utility parameters of the model (generosity, warm glow, and weight of reputational concerns) for each respondent and present aggregate statistics for the whole sample.

2.4.1 First direction of crowding out: taxes on donations

Our first result is that taxes do crowd out donations in our setting, but to a very limited extent. By regressing donations and donation rates on in-survey income and tax rate, and including individual fixed effects, we obtain that a 1% increase in tax rates results in a 0.058% decrease in donation rates (column 1 of Table 2.4), implying a crowding out the magnitude of less than 6%. This result is very far from the 100% rate implied by the full crowding out hypothesis, suggesting that individuals are not only interested in the total amount of public good (pure altruism) but also in the extent of their own contribution (*warm glow*).

Table 2.4: In-survey donations, donation rate, and expected donation rate on tax rates and income

	Donations (in $ 1000)	Donation rate (%)	Expected donation rate (%)
Income (in $1000)	0.011***	-0.021***	
	(0.001)	(0.002)	
Tax rate (%)	-0.048***	-0.058***	-0.076***
	(0.011)	(0.018)	(0.028)
Individual F.E.	Yes	Yes	Yes
Observations	5,700	5,700	1,900
R^2	0.536	0.505	0.498
Note:	*p<0.1; **p<0.05; ***p<0.01		

The last column of Table 2.4 reports the result of an analogous regression, but with expected donation rates as the outcome variable (which, differently from own donations, are independent of income). The coefficient of the explanatory variable (in-survey tax rates) is negative and significant, but its magnitude is larger (by almost 2% in absolute terms) than for one's own donation rates. This discrepancy suggests that respondents' beliefs might be inaccurate, which we will now test more formally.

2.4.2 Correct beliefs

We now test whether respondents hold correct beliefs about the average level of donations in the hypothetical society described in the survey. Since we do not provide information on the income distribution in the society, but only a measure of central tendency[15], we aggregate actual donations by levels of income by using several sets of weights. For the primary analysis, we use the prevalence of low, middle, and high-income households[16] in the actual U.S. population, based on the 2019 version of the American Community Survey[17], but results are robust to using equal income weights, as well as to considering middle-income households only or to excluding middle-income households and considering equal weights for the remaining two categories. Figure 2.4 shows the distribution of the average difference between expected and realized donations for each level of the additional tax rate (0% to 10%). Standard errors are bootstrapped. Table 2.5 also reports the p value for the paired t-test for the difference in means, which leads us to reject the null hypothesis of accurate beliefs for all levels of taxes. Despite the difference between expected and actual donations being consistently positive and significant (implying overestimation of others' donations), the magnitude is larger for more extreme tax rates (on average \$420 versus \$310), suggesting that individuals tend to form better predictions in more realistic or preferable situations.

[15] Respondents are told that the typical income in society is \$60,000

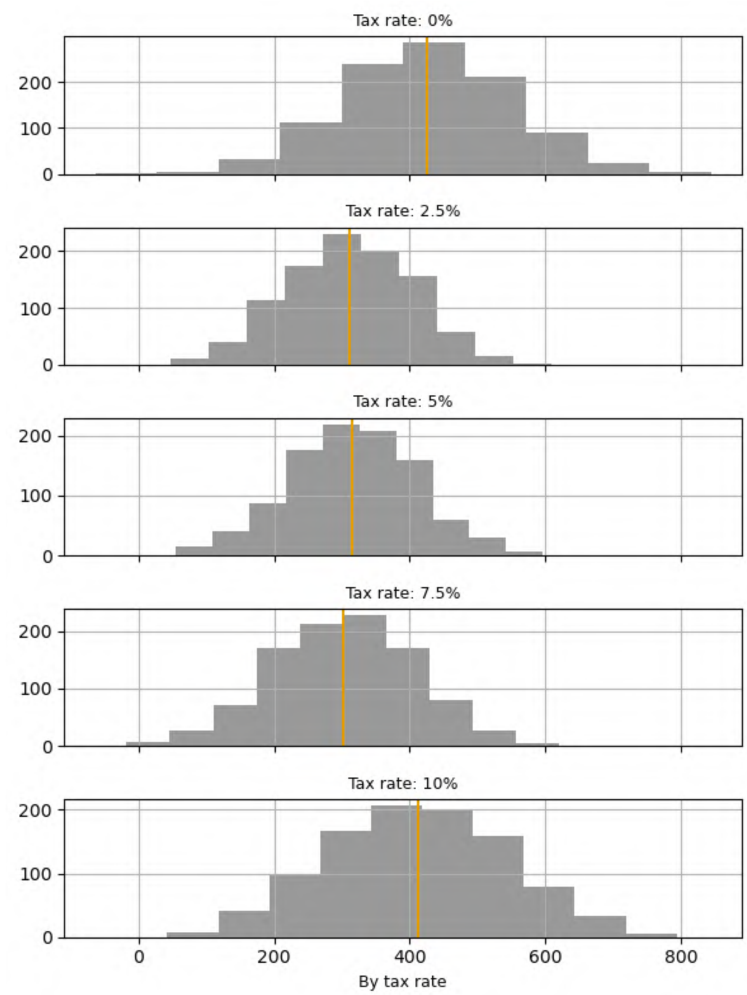
[16] Low income is defined as less than \$50,000, middle income as between \$50,000 and \$90,000 and high income as more than \$90,000

[17] Available at: https://censusreporter.org/topics/income/.

Table 2.5: Differences in expected and realized donations, and testing for accurate beliefs

Tax rate	Donations			
	Expected (in 1000 $)	Realized (in 1000 $)	Difference (in 1000 $)	Paired t-test p-value
0%	2.618	2.191	0.427	0.00
2.5%	2.299	1.987	0.313	0.00
5%	2.173	1.862	0.312	0.00
7.5%	2.070	1.759	0.311	0.00
10%	2.161	1.741	0.421	0.00

Figure 2.4: Distribution of the difference between expected and realized donations



Distribution of the bootstrapped difference between expected and realized donations. Income weights are 0.38, 0.30, 0.32

2.4.3 Second direction of crowding out

We next estimate the effect of the availability of donations on preferred tax rates, representing the second direction of crowding out. As all respondents are asked to state their preferences for all three imagined levels of household income, with and without donations available, we can interpret the estimates causally within the survey game's setup. Figure 2.5 shows the distribution of preferred tax rates for the two main scenarios (with versus without charity) and each level of in-survey income. Answers concentrate around 5%, especially for middle and higher income levels, while maximal levels appear more frequent with higher income and minimal levels with lower incomes. As expected, Table 2.6 reveals that respondents tend to prefer lower additional tax rates when donations are allowed: compared to the baseline of around 4.53% ideal tax rate for low-income households, donations decrease ideal taxes by 0.74% on average, while higher in-game income results in higher preferred tax rates. The availability of donations does not interact with the in-game income levels on average for the entire sample, so the effect of income on the ideal tax rate seems to be independent of donation availability.

Figure 2.5: Ideal tax rate with and without donations

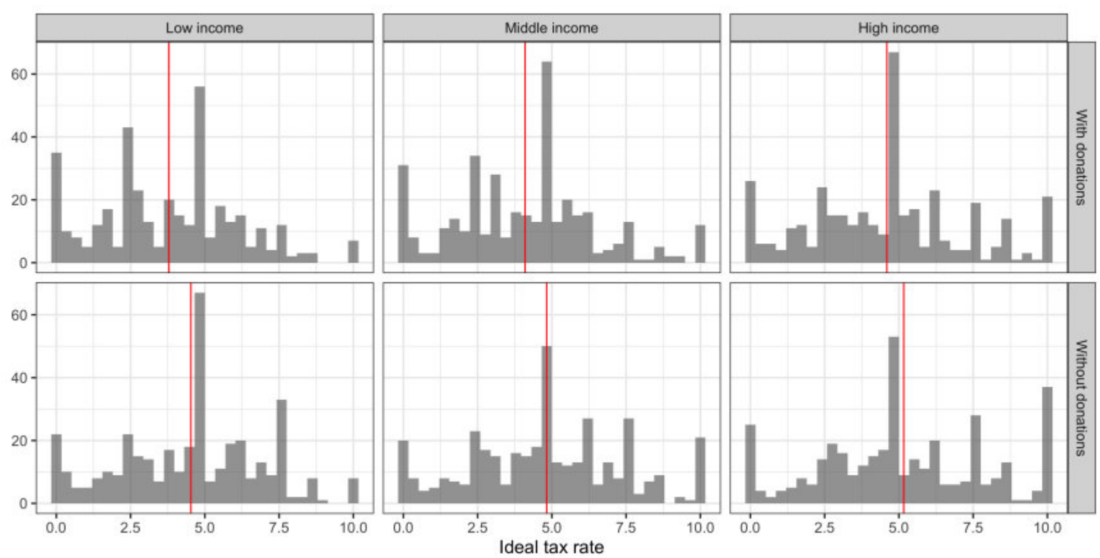


Table 2.6: Ideal tax rates regressed against in-game income and donation availability

	Ideal tax rate
Middle income (60k)	0.29378**
	(0.09288)
High income (120k)	0.63237***
	(0.10622)
Donations allowed	-0.73905***
	(0.12930)
Middle income (60k) X Donations allowed	0.01309
	(0.12726)
High income (120k) X Donations allowed	0.17031
	(0.13485)
Multiple R-squared (full model):	0.6719
Adjusted R-squared:	0.6054

Notes: The table displays the regression of ideal tax rates on in-game income level interacted with whether donations are allowed, using respondent-level fixed effects. Standard errors are clustered on the respondent level. *p<0.1; **p<0.05; ***p<0.01

Despite the usefulness of these individual-level results, which already point to some crowding out of donations on preferred public support for zero-income households, we are ultimately interested in the equilibrium tax rate at the societal level. Therefore, based on our model, we aggregate individual preferences by solving the neutral government's problem, which results in selecting the preferred tax rate of the median voter.

We rely on bootstrapping to simulate our hypothetical society repeatedly, where the bootstrapped preferences of survey respondents account for 85% of the votes (i.e., the proportion of positive-income households in society) while the remaining 15% of the votes are for the highest available additional tax rate(10%) since zero-income households optimize their utility by maximizing public support.

Table 2.7: Average realizations of the outcomes of interest

Variable	Private charity	Mean	SD
τ_a^{med} (%)	No	5.249	0.160
Benefit (in 1000 $)	No	21.167	0.646
τ_a^{med} (%)	Yes	4.848	0.110
Benefit (in 1000 $)	Yes	30.170	$0.608
Average donation rate (%)	Yes	2.966	0.103
Average donation (in 1000 $)	Yes	1.874	0.006
Average expected donation (in 1000 $)	Yes	2.183	0.007

Table 2.8: Average difference in the outcomes of interest

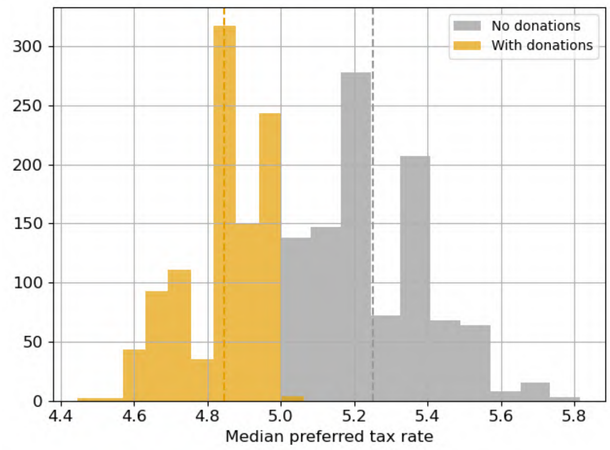
Difference	Mean	SD
τ_a^{med} (%)	-0.401	0.156
Benefit (in 1000 \$)	9.003	0.747

The resulting distribution of the equilibrium tax rate in the two main cases (with and without charity) and of the benefit accruing to each poor household are reported in figure 2.6, alongside the distribution of the difference in the two outcomes of interest. The average values and bootstrapped standard deviations are reported in table 2.7, while table 2.8 reports the average difference in equilibrium tax rate and benefit. The equilibrium tax rate is 5.25% in the taxation-only case compared to 4.85% when donations are allowed. Private donations (on average \$1,874 per positive-income household) more than compensate for the loss in public support, resulting in a much higher benefit per zero-income household in the case with charity (\$30,000 versus \$21,000)[18].

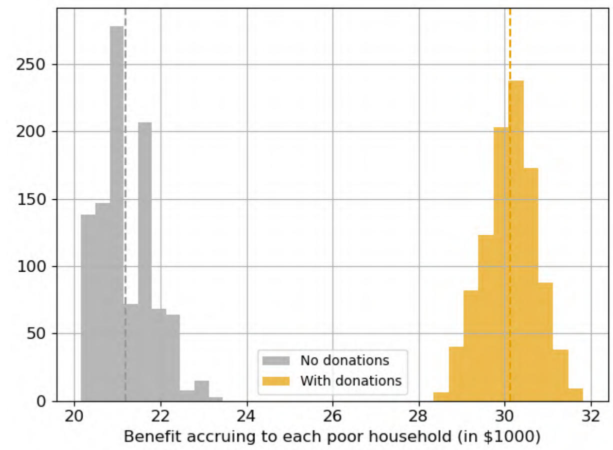
[18]To retrieve the average level of donations in the case with charity available we consider the preferred donation of each respondent for the two discrete levels of tax rate which are closest to the equilibrium level, weighting each by its distance to the equilibrium level.

Figure 2.6: Equilibrium tax rates and benefits of the simulations

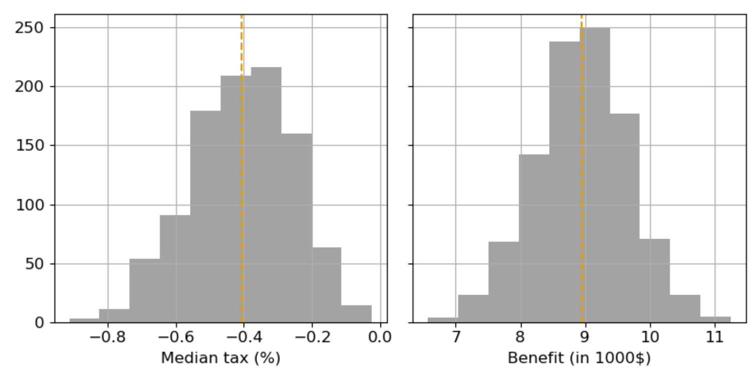
(a) Additional tax rate in equilibrium



(b) Poor household benefit in equilibrium



(c) Difference in equilibrium outcomes



2.4.4 Estimation of individual utility parameters

In order to better understand the structure of preferences leading to these results, we also estimate the individual utility parameters of our theoretical model, namely generosity (α), taste for donations (γ), and weight of reputational concerns (η). First, the value of generosity (α) is obtained from the no-donation case, by solving Equation 2.1 for each individual and each of the three possible levels of wage in the scenarios where donations are not allowed. We then average out the three individual-level observations to obtain the final estimate. Second, we retrieve the remaining parameters (γ and η) as the result of the minimization of the sum of squared deviations of observed in-game donations from the theoretical donations implied by solving the individual utility maximization problem with the given utility parameter values.[19]

Table 2.9 reports the main summary statistics for the estimated parameters, while figure 2.7 shows the distribution for the entire sample. On average, generosity has a value of 6.6%, meaning that individuals assign to the utility of the poorest members of society a weight of 6.6 percentage points compared to the weight of their own utility from consumption. Estimates at the individual level range from a minimum value of 0 to a maximum value of 14.3 percentage points. Direct utility from donations, or *warm glow*, is estimated to be 2.8% of the utility from one's own consumption for the average respondent, with a minimum of 0 and a maximum of 8.4 percentage points in the whole sample. Finally, the weight of reputation, proxied by the deviation of own donations from the expected societal level, is on average 0.2 percentage points, ranging from a minimum of 0 to a maximum of 3.8%.

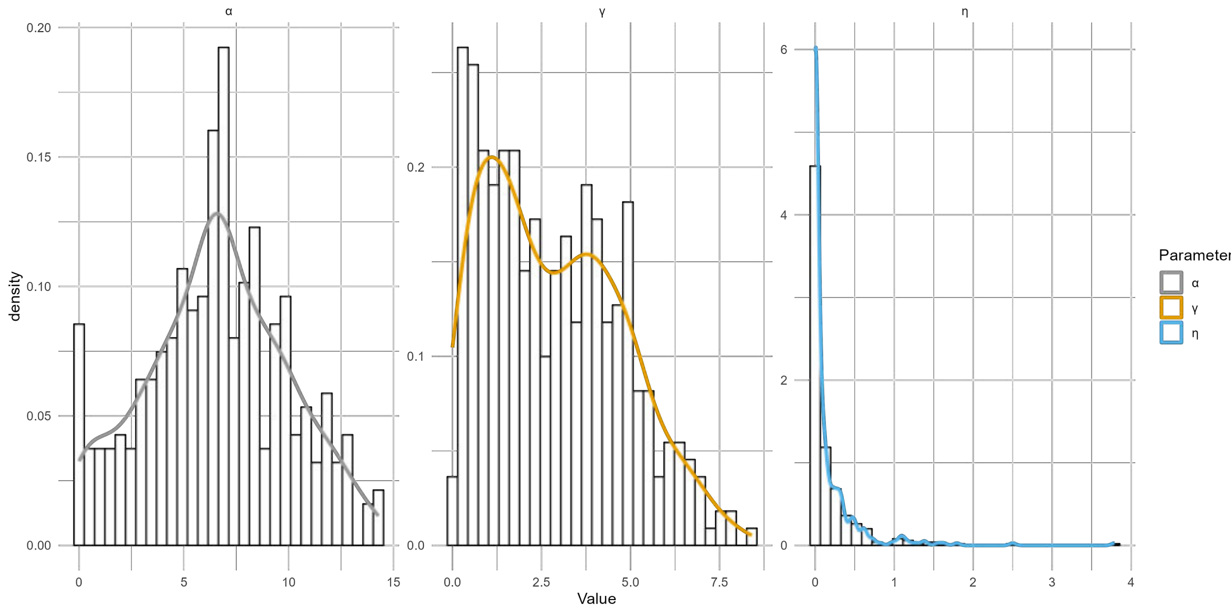
Generosity follows a unimodal shape with the majority of the sample being centered around the mean, with the exception of bunching around 0 for those people who do not derive utility from the welfare of the poor. In contrast, the shape of the warm glow parameter distribution suggests the presence of two different groups, one centered at approximately 2% and one assigning a larger weight of nearly 5% to the direct utility from donations. Finally, the weight of reputational concerns is very close to zero for the vast majority of the sample, with a longer tail, and some bunching around the value of 1%, suggesting the existence of a specific subset of respondents that care highly about social expectations.

Table 2.9: Summary statistics for the estimated utility parameters

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
α	380	0.066	0.034	0.000	0.044	0.089	0.143
γ	380	0.028	0.019	0.000	0.012	0.042	0.084
η	380	0.002	0.004	0.000	0.000	0.002	0.038

[19] The problem is solved for each level of wage and tax rate available in the hypothetical scenarios. Additional details on the estimation procedure are presented in the Appendix.

Figure 2.7: Societal distribution of utility parameters



2.4.5 Background characteristics and in-game behavior

Four key in-game behavioral variables (donation rates, expectations about donation rates, the difference between the two, and preferred tax levels) determine the simulation results, for which we can examine partial correlations with respect to other relevant background characteristics. We average through the values across scenarios for each individual, then regress them on demographic information, attitudes towards inequity and fairness, the preferred size of unemployment benefits, and psychological factors (risk-aversion and patience as in Falk et al., 2016, along with self-reported real-life charitable behavior. We construct principal components based on the variable groups of inequity attitudes and unemployment benefits due to their high cross-correlation, and we include those in the regressions.

Table 2.10 presents the OLS estimates of the regressions described above and displays the means and standard deviations for the outcome variables at the bottom. We can see that, on average, the in-game donation rates are close to the aggregates we observe in county-level data. As we already noted, the expectations of survey participants about the average donation rates do not match the realized average, as people overestimate how much others would donate by around 0.8 percentage points (23%). As a baseline, the positive relationship between real-life and in-game donation rates and the negative relationship between conservatism and preferred tax levels provide evidence of consistency between the survey respondents' in-game behavior and their in-life attributes.

Our results suggest that in-game donation rates are lower for women. Conservative opinions also seem to be associated negatively with the propensity to donate and the preferred level of taxes. In contrast, more inequity-averse individuals prefer lower donations and have higher ideal tax rates. Surprisingly, Protestant self-identification does not seem to correlate strongly with own in-game donations or ideal tax rates; however, it negatively relates to expected donation rates. In addition, living in a predominantly Protestant area negatively correlates with preferred tax rates, showing the importance of majority religion on regional social norms as described by Pugh (1980). Volunteering also negatively correlates with in-game own and expected donations, suggesting that donation and volunteering might be substitutes rather than complements. Finally, being more forward-looking and risk-averse are also negatively associated with donations, which is consistent with behavior that prioritizes higher savings against present expenditures, in this case, donations.

Table 2.10: In-survey behavior outcomes and participant background

	Dependent variable:			
	Donation rate	Expected donation rate	Preferred taxes	Donation difference
	(1)	(2)	(3)	(4)
Tertiary educated	−0.150	0.450*	−0.124	−0.519*
	(0.237)	(0.273)	(0.261)	(0.266)
Female	−0.466**	−0.507**	−0.366	0.007
	(0.209)	(0.241)	(0.231)	(0.235)
Age 28-37	−0.554	−0.887**	−0.723*	0.044
	(0.357)	(0.411)	(0.394)	(0.401)
Age 38-47	0.244	−0.238	−0.216	0.038
	(0.393)	(0.452)	(0.433)	(0.441)
Age 48-57	−0.179	0.142	−0.344	−0.675
	(0.390)	(0.449)	(0.430)	(0.438)
Age 58+	−0.584	−0.519	−0.494	−0.397
	(0.354)	(0.408)	(0.391)	(0.398)
Black	0.277	0.508	−0.492	−0.339
	(0.307)	(0.353)	(0.339)	(0.344)
Asian	−0.249	0.837**	−0.713*	−1.307***
	(0.364)	(0.419)	(0.402)	(0.408)
Hispanic	0.441	0.472	0.235	0.018
	(0.430)	(0.495)	(0.475)	(0.483)
Other	0.010	−0.671*	−0.148	0.693*
	(0.343)	(0.395)	(0.379)	(0.385)
Majority religion is Protestant	−0.246	0.041	−0.598**	−0.350
	(0.234)	(0.270)	(0.258)	(0.263)
Own religion is Protestant	−0.038	−0.653*	−0.022	0.649*
	(0.295)	(0.339)	(0.325)	(0.331)
Goes to church at least monthly	0.115	0.249	0.200	−0.128
	(0.228)	(0.262)	(0.251)	(0.256)
Conservative scale	−0.126**	−0.064	−0.173***	−0.047
	(0.054)	(0.062)	(0.060)	(0.061)
log(real life donation rate+0.001)	0.190***	0.089	0.130	0.109
	(0.072)	(0.083)	(0.079)	(0.081)
Real life donation rate is 0	0.378	0.547	0.201	−0.387
	(0.466)	(0.537)	(0.514)	(0.523)
Real life regular volunteering	−0.411*	−0.489**	−0.021	−0.074
	(0.210)	(0.242)	(0.232)	(0.236)
Inequity aversion princ. comp.	−0.348***	−0.188**	0.135	−0.097
	(0.079)	(0.091)	(0.087)	(0.089)
Unemployment benefit princ. comp.	0.190***	0.041	0.147**	0.196***
	(0.064)	(0.074)	(0.071)	(0.072)
Number of right answers	−0.441***	−0.499***	0.032	0.111
	(0.119)	(0.137)	(0.132)	(0.134)
Forward-looking preferences	−0.070	−0.038	0.083	−0.012
	(0.049)	(0.056)	(0.054)	(0.054)
Risk-loving preferences	−0.096**	−0.086*	−0.052	0.002
	(0.043)	(0.050)	(0.048)	(0.049)
Married or cohabiting	−0.033	−0.532*	−0.095	0.612**
	(0.250)	(0.288)	(0.276)	(0.281)
Number of children	0.128	0.267**	0.030	−0.118
	(0.098)	(0.113)	(0.108)	(0.110)
Constant	4.673***	4.185**	5.618***	0.995
	(1.547)	(1.782)	(1.706)	(1.735)
Observations	380	380	380	380
R^2	0.302	0.234	0.148	0.154
Adjusted R^2	0.246	0.173	0.080	0.086
Residual Std. Error (df = 351)	1.790	2.062	1.975	2.009

Notes: The population size of the respondent's area and the log of their estimated real-life income are also included in the regressions. We omitted them from the table to ease visibility as they are not statistically significant at the 10% level. Standard errors are in parentheses. Donation difference refers to the difference between the individual's own donation rate vs. their expected donation rate for the aggregate level.
*p<0.1; **p<0.05; ***p<0.01

2.5 Conclusion

Our results corroborate and extend several previous findings in the literature regarding the crowd-out between charity and the state, the drivers of charitable behavior, and individual behavior in public good games. By collecting and analyzing novel survey data, we provide evidence for the less-studied direction of charity crowding out the state in an abstract setting, connecting to the findings of Sav (2012), and Werfel (2018) amongst others. In our survey, we document that the other direction is also present: when taxes are higher, respondents choose to donate less. However, the relationship is not strictly monotonous for individual respondents or, on average. It suggests that even under the stylized and simplified conditions of our hypothetical scenarios, crowd-out might be only partial as people do not internalize the full effect of their choices on the public good provision, in line with the experimental findings of Eckel et al. (2005). We also find survey respondents to systematically overestimate the average donation rate in society compared to their realized average contribution, which might result in a sub-optimal public choice regarding poverty reduction. Our survey results are also in accord with aggregate evidence, suggesting a negative association of donations with poverty and taxes.

In our stylized setting, the higher equilibrium tax rates characterizing the no-donations scenario are not enough to compensate for the loss of private charity in terms of the benefit accruing to the poor. Retrieving the structure of preferences that generate the observed in-survey behavior, we confirm that individuals are not only interested in the welfare level of the poorest members of society but are also positively affected by the direct utility of contributing. This result corroborates the findings of Null (2011) that only a few donors are willing to pay to check whether their donations reach their declared target.

References

Andreoni, J. (1988). Privately Provided Public Goods in a Large Economy: The Limits of Altruism. *Journal of Public Economics*, 35:57–73.

Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *Economic Journal*, 100:464–77.

Andreoni, J. and Payne, A. A. (2013). Charitable Giving. In *Handbook of Public Economics*, volume 5, pages 1–50. Elsevier.

Becker, E. and Lindsay, C. M. (1994). Does the Government Free Ride? *The Journal of Law & Economics*, 37(1):277–296.

Becker, G. S. (1974). A Theory of Social Interactions. *Journal of Political Economy*, 82(6):1063–1093.

Bergstrom, T., Blume, L., and Varian, H. (1986). On the Private Provision of Public Goods. *Ted C Bergstrom*, 29.

Bredtmann, J. (2016). Does government spending crowd out voluntary labor and donations? *IZA World of Labor*.

Brooks, A. C. (2000). Is There a Dark Side to Government Support for Nonprofits? *Public Administration Review*, 60(3):211–218.

Bénabou, R. and Tirole, J. (2006). Incentives and Prosocial Behavior. *American Economic Review*, 96(5):1652–1678.

Charities Aid Foundation (2019). CAF World Giving Index 10th Edition: Ten years of giving trends. Technical report.

Duncan, B. (1999). Modeling charitable contributions of time and money. *Journal of Public Economics*, 72(2):213–242.

Eckel, C. C., Grossman, P. J., and Johnston, R. M. (2005). An experimental test of the crowding out hypothesis. *Journal of Public Economics*, 89(8):1543–1560.

Enke, B., Rodríguez-Padilla, R., and Zimmermann, F. (2020). Moral Universalism and the Structure of Ideology. Technical Report w27511, National Bureau of Economic Research.

Falk, A., Becker, A., Dohmen, T., Huffman, D., and Sunde, U. (2016). The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences. page 69.

Garrett, T. and Rhine, R. (2010). Government growth and private contributions to charity. *Public Choice*, 143(1/2):103–120.

Heutel, G. (2014). Crowding Out and Crowding In of Private Donations and Government Grants. *Public Finance Review*, 42(2):143–175.

Hoge, R. and Yang, F. (1994). Determinants of Religious Giving in American Denominations: Data from Two Nationwide Surveys. *Review of Religious Research*, 36(2):123.

Katz, E. and Rosenberg, J. (2005). An economic interpretation of institutional volunteering. *European Journal of Political Economy*, 21(2):429–443.

Null, C. (2011). Warm glow, information, and inefficient charitable giving. *Journal of Public Economics*, 95(5-6):455–465.

OECD (2021). Tax revenue (indicator). doi: 10.1787/d98b8cf5-en (Accessed on 28 May 2021).

Peloza, J. and Steel, P. (2005). The Price Elasticities of Charitable Contributions: A Meta-Analysis. *Journal of Public Policy & Marketing*, 24(2):260–272. Publisher: SAGE Publications Inc.

Pugh, W. J. (1980). Catholics, Protestants, and Testamentary Charity in Seventeenth-Century Lyon and Nimes. *French Historical Studies*, 11(4):479.

Pullan, B. (2005). Catholics, Protestants, and the Poor in Early Modern Europe. *The Journal of Interdisciplinary History*, 35(3):441–456.

Sav, G. T. (2012). Government free riding in the public provision of higher education: panel data estimates of possible crowding out. *Applied Economics*, 44(9):1133–1141.

Schiff, G. (1985). Does government spending crowd out charitable contributions? *National Tax Journal*, 38(4):535–546.

Simmons, W. and Emanuele, R. (2004). Does Government Spending Crowd Out Donations of Time and Money? *Public Finance Review*, 32:498–511.

van Elk, M., T. Rutjens, B., and van Harreveld, F. (2017). Why Are Protestants More Prosocial Than Catholics? A Comparative Study Among Orthodox Dutch Believers. *The International Journal for the Psychology of Religion*, 27(1):65–81.

Werfel, S. H. (2018). Does charitable giving crowd out support for government spending? *Economics Letters*, 171(C):83–86.

Appendix

2.5.1 Additional survey results

Figure 2A1: Donation rate by additional tax rate and level of income

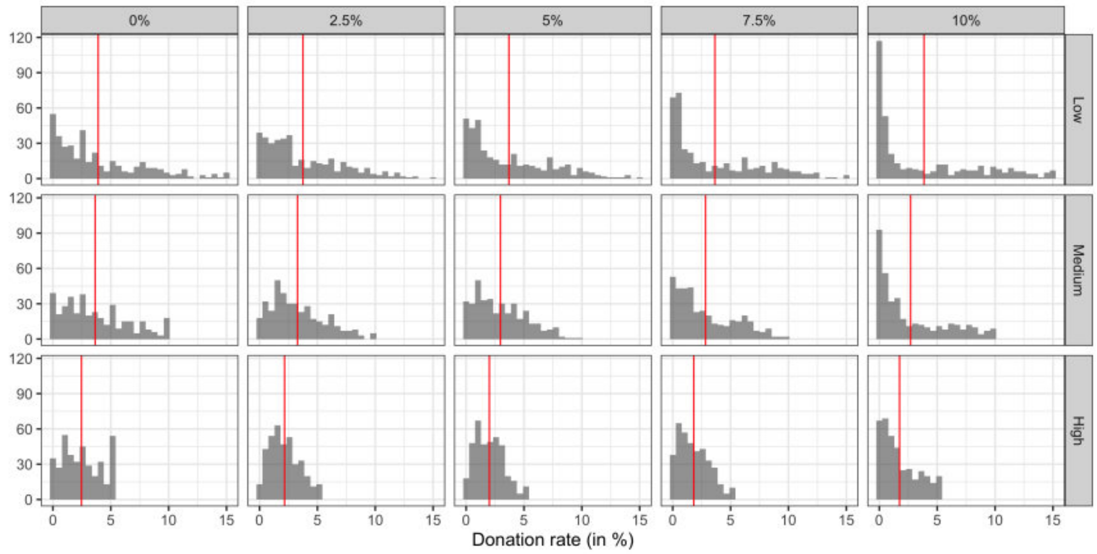
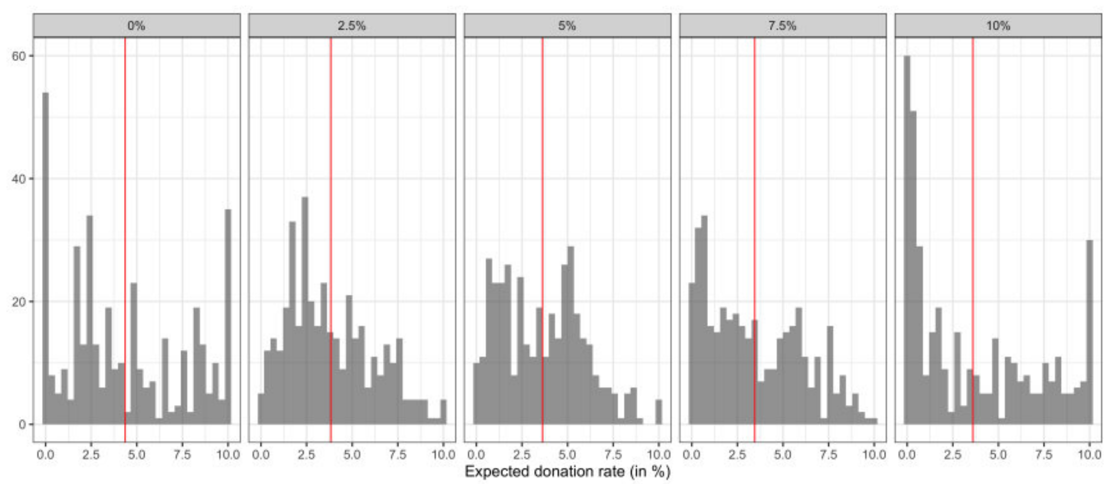


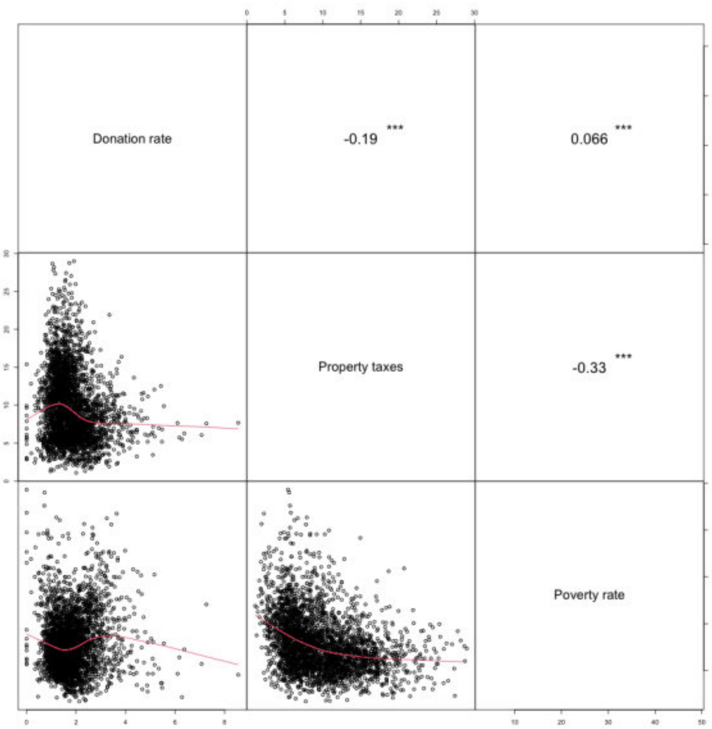
Figure 2A2: Expected donation rate by additional taxation level



2.5.2 Additional county-level descriptives

Figure 2A3 displays the pairwise correlations between the key variables relevant to the analysis on the county level.

Figure 2A3: Bivariate relations between the key variables



Note: Figure shows the bivariate relations of the county-level variables of the authors' calculations based on the following publicly available datasets. The donation rate for 2016 is calculated as donations over adjusted gross income. The data are accessed via the website of the Internal Revenue Service, maintained by the Statistics of Income division. Data on property taxes from 2010-2014 are collected by the National Association of Home Builders. Poverty rates are calculated based on the CPS ASEC data by the U.S. Census Bureau.

The causal links behind these correlations are unclear. Hence, as an additional exploratory step, we show regression results intended to provide a basic understanding before examining the question in more depth with our survey analysis. Indeed, regressions might not reveal causal links, as the actual mechanism could be driven by unobservable characteristics of the counties, such as different levels of inequity aversion or moral codes behind altruistic behavior (Enke et al., 2020), or by simultaneity as these variables are equilibrium outcomes.

In our first setup, we regress poverty rates on donation rates, property taxes, county-level characteristics, and state fixed effects. The first column of Table 2A1 reports the results of the simplest specification, where the poverty rate is regressed only on donation rate and property taxes. We can see that without additional covariates, there seems to be no statistically significant relationship between donations and poverty. At the same time, there is a strong negative partial association between taxes and the poverty rate. Including state fixed effects (column 2) results in a negative partial correlation between the poverty rate and donations and taxes. Columns 3 to 6 include an increasingly comprehensive set of controls, with column 6 controlling for real GDP per capita, population size, and demographic composition according to religion[20], age group, ethnicity and educational level, and presidential election results. The additional covariates turn the relationship between taxes and poverty insignificant while preserving the sign and significance of the negative correlation with donation rates. Other controls improve the precision of the estimated donation rate coefficient (-0.8). If we were to interpret this estimate as causal, increasing donations by one standard deviation (0.8) would result in a relatively small decrease in the poverty rate of 0.64 percentage points, which is approximately 4% of the mean poverty rate and 10% of its standard deviation.

Table 2A1: County-level regression associations of poverty rate with property taxes and donation rates

	Dependent variable:					
	Poverty rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Donation rate	0.034	-1.725^{***}	-0.916^{*}	-1.048^{**}	-0.931^{***}	-0.806^{***}
	(0.691)	(0.406)	(0.508)	(0.493)	(0.244)	(0.250)
Property taxes	-0.448^{***}	-0.232^{***}	-0.092	-0.050	-0.023	-0.039
	(0.090)	(0.084)	(0.108)	(0.096)	(0.039)	(0.037)
State FE	No	Yes	Yes	Yes	Yes	Yes
GDP per capita, population size	No	No	Yes	Yes	Yes	Yes
Religious composition	No	No	No	Yes	Yes	Yes
Age, ethnicity, education composition	No	No	No	No	Yes	Yes
Election results	No	No	No	No	No	Yes
Observations	3,129	3,129	3,129	3,129	3,128	3,128
R^2	0.110	0.367	0.400	0.422	0.781	0.786

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors clustered at the state level in parentheses.

In our second setup (whose results are reported in Table 2A2), we take the local property tax as the outcome variable, and we examine its partial correlations with the two remaining key variables (donation rate and poverty rate). Here, even the coefficients' sign reacts sub-

[20] In Christianity, the role of charity is central, however, due to the different historical institutional evolution Protestants are expected to donate more than Roman Catholics (Pugh, 1980; Hoge and Yang, 1994; Pullan, 2005; van Elk et al., 2017).

stantially to the set of controls we include in the regressions. In the simplest specification, which does not control for county characteristics nor state fixed effects, higher poverty rates and donation rates are associated with lower property taxes. Including state fixed effects reduces the coefficient on poverty rate by a magnitude and flips the sign on donation rates while accounting for most of the explained variance fraction of 0.78. The inclusion of demographic controls reverts the partial correlation between tax and donations to negative, while the impact of the poverty rate becomes non-significant. If we were to interpret the estimates causally, a standard deviation increase of donation rates would imply a slight decrease in local taxes of around 0.288 per 1000$ of property value, around 3% of the mean.

Table 2A2: County-level regression associations of property taxes with poverty and donation rates

	Dependent variable:					
	Property taxes					
	(1)	(2)	(3)	(4)	(5)	(6)
Donation rate	−0.960**	0.518***	−0.058	−0.101	−0.407***	−0.360***
	(0.374)	(0.149)	(0.116)	(0.120)	(0.102)	(0.099)
Poverty rate	−0.238***	−0.045***	−0.016	−0.009	−0.009	−0.016
	(0.051)	(0.015)	(0.019)	(0.017)	(0.015)	(0.014)
State FE	No	Yes	Yes	Yes	Yes	Yes
GDP per capita, population size	No	No	Yes	Yes	Yes	Yes
Religious composition	No	No	No	Yes	Yes	Yes
Age, ethnicity, education composition	No	No	No	No	Yes	Yes
Election results	No	No	No	No	No	Yes
Observations	3,129	3,129	3,129	3,129	3,128	3,128
R^2	0.138	0.778	0.814	0.816	0.837	0.839

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors clustered at the state level in parentheses.

The evidence we presented so far suggests that the inter-relatedness of poverty, charity, and taxation is challenging to clarify. According to these preliminary findings, higher donations seem to be associated with lower poverty and local property taxes, while the suggested magnitudes are relatively small. However, the sign and magnitude of these estimates are not robust to the inclusion of different controls, nor can we claim that they capture causal relationships.

2.5.3 Individual-level parameter estimation

Generosity parameter

The generosity parameter is obtained at the individual level by solving equation 2.1 for each individual and each level of wage in the scenarios when donations are not allowed:

$$\alpha_i(w_n) = \frac{\tau_i^*(w_n) - \underline{\tau}}{1 - \tau_i^*(w_n)}$$
$$\alpha_i^* = \frac{\sum_{n=1}^{3} \alpha_i(w_n)}{3}$$

Warm glow and reputation weight parameters

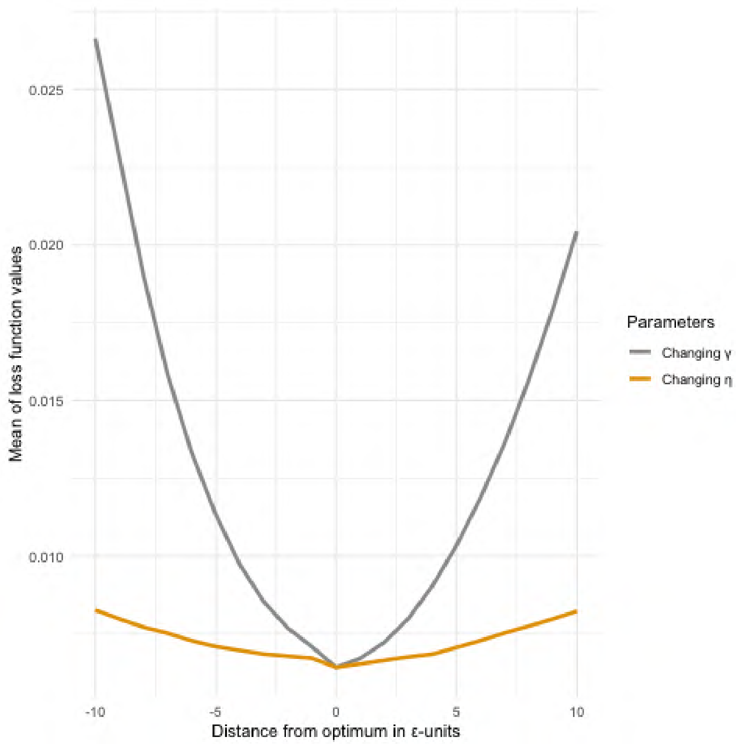
The remaining two parameters, measuring the weight of the direct utility from donations (γ_i) and of reputation (η_i), are estimated numerically via a non-linear optimization algorithm which minimizes the normalized sum of squared deviations of the observed in-survey donations from the theoretical solution of the utility maximization problem for each wage and tax rate level proposed in the hypothetical scenarios. For each individual i, (γ_i, η_i) is chosen to minimize the following expression given the levels of in-game taxation, income, and α_i, the previously estimated utility weight for generosity:

$$\min_{(\gamma_i, \eta_i)} \sum_{n=1}^{5} \sum_{m=1}^{3} \big(d_i(\tau_n, w_m) - d_i^*(\gamma_i, \eta_i; \alpha_i, \tau_n, w_m)\big)^2,$$

where $\tau_n \in \{0, 0.025, 0.05, 0.075, 0.1\}$ and $w_m \in \{40,000; 60,000; 120,000\}$.

While for each of the 380 respondents, we cannot show that the objective functions' minima were indeed reached, we can display in Figure 2A4 the average values of the objective function across respondents. We calculate them for a $\pm 10 \times 10\%$ neighborhood of the selected values for the two utility parameters. We can see that the γ and η values found indeed minimize the optimization problems on average.

Figure 2A4: Average value of the objective function in a neighborhood of the solution



Chapter 3

Housing and Fertility

Abstract

I build a dynamic equilibrium model of household behavior with unobserved heterogeneity in the desired number of children to examine how policies targeting the housing market affect choices of fertility, location, and house size of young households. I estimate the model's structural parameters using data from Hungary to evaluate the dynamic effects of the Family Housing Allowance policy, which provided a sizeable lump-sum subsidy for house purchases, with built-in commitment regarding the number of children the family would have. The model suggests that the combination of lower interest rates and the allowance increases house prices substantially compared to the baseline, which for poorer households counteract some of the positive welfare effects of the policy. While according to the model, completed fertility increases due to the policy by around 5-10% driven by poorer households, their housing conditions worsen in the long run due to the elevated house prices. Richer households experience no adverse effects, while their completed fertility is affected positively only to a smaller degree by the policy.

Keywords: Housing Demand, Fertility, Housing Policy, Residential Choice

JEL Classification: J11, J13, J61, H31, R21

I thank my thesis supervisors, Gianmarco Ottaviano, Joseph-Simon Görlach, and Francesco Billari, for the years of excellent guidance and support. I am grateful for the comments and discussions with Jérôme Adda, Kurt Mitman, Florian Oswald, Manuela Puente Beccar, and Balázs Reizer. I am also thankful for the valuable remarks of the Bocconi La Strada Seminar seminar participants. Finally, I thank the Centre for Economic and Regional Studies for supporting this research, the Databank of the Centre for Economic and Regional Studies, and the Hungarian Central Statistical Office for providing access to the databases and all the technical help their colleagues have given me.

3.1 Introduction

Housing plays a central role in the life of a typical household. In the United States and other developed countries, housing wealth has accounted for around 90% of young and poor households' assets in the last few decades and 35% of all household assets (Davis and Van Nieuwerburgh, 2015). Jordà et al. (2019) documents that housing has historically provided the largest part of capital stock, with returns comparable to equity. Furthermore, housing expenses have constituted around 20-25% of the total household expenses, a consistent finding for different periods and locations (Ortalo-Magné and Rady, 2006; Davis and Ortalo-Magné, 2011; Piazzesi and Schneider, 2016). Housing also contributed significantly to the Great Recession of 2008 (Mian et al., 2013; Famiglietti et al., 2020) and is understood to be a significant channel of monetary policy (Di Maggio et al., 2017). Real estate differs from other assets in several aspects. Houses are illiquid due to search frictions (Merlo et al., 2015); they are indivisible and relatively expensive, so most households must borrow to purchase one. Households usually end up owning only one house, where they reside and search for employment in its proximity (Davis and Van Nieuwerburgh, 2015). Compared to generic goods and services, housing is unique as its residents must consume its services. Furthermore, real estate functions as collateral to other types of borrowing, providing a channel through which household consumption is connected to the housing market (Attanasio et al., 2012; Di Maggio et al., 2017; Cloyne et al., 2019).

In the last few decades, the housing conditions of young households have been deteriorating. House ownership amongst young people has declined or been postponed notably in the developed world (Fisher and Gervais, 2011; Green and Lee, 2016; Flynn, 2020). Due to their high price, most households typically use mortgages to purchase houses. These mortgages are naturally mainly issued to young families who decide to suppress consumption in their early earning years to afford the down payment for the house (Chan et al., 2015). It is, however, necessary: most renters cannot afford a house. Therefore down payment, ranging between 20%-50% of the total price is the chief obstacle for young households' home ownership (Davis and Van Nieuwerburgh, 2015). Down payment depends on house prices, which means that escalating house prices could contribute mainly to young households' inability to get a mortgage (Barakova et al., 2014). Moreover, evidence shows that housing is deeply intertwined with fertility decisions and the choice of residential location. Difficult housing conditions in terms of prices or security are documented to decrease or delay fertility amongst young households who look for their first home ownership, a wide-spread finding in the literature (Ermisch, 1999; Ström, 2010; Öst, 2012; Vignoli et al., 2013; Kulu and Steele, 2013; Dettling and Kearney, 2014; Day and Guest, 2016; Lin et al., 2016; Öst and Wilhelmsson, 2019)[1]. Young couples are also found to adjust their housing status to budgetary circumstances, credit constraints, and anticipated fertility (Ortalo-Magné and Rady, 1999; Mulder and Lauster, 2010; Fisher and Gervais, 2011; Ermisch and Steele, 2016; Vidal et al., 2017; Mulder, 2018). This channel, alongside essential changes in the female role of the household (Esping-Andersen and Billari, 2015), could have contributed to low fertility rates in developed countries (Billari and Kohler, 2004), as high home ownership along with low access to mortgages is associated with the lowest fertility rates (Mulder and Billari, 2010). An unfriendly housing environment for young couples could then contribute to potential future difficulties in the social security systems of European welfare states (Flynn, 2017; Zeman et al., 2018).

In this paper, I study how fertility and the housing market interplay with the residential choices of young households and how policy interventions that target housing conditions could effectively improve fertility outcomes. To address this question, I build a life cycle model of forward-looking households choosing fertility, house size, ownership, and location, with unobserved heterogeneity in their desired number of children. I apply the model to analyze the effects of a government program in Hungary running since 2014/2015, called the Family Housing Allowance (further referred to as FHA)[2], which aimed to ease the borrowing constraints of young households at home purchases or mortgage down payments and to revitalize the housing market of the country. The policy provides a substantial non-refundable lump sum at house purchases to families if they commit to or already have at least three children.[3] The overall effects are not straightforward to assess either in the long run or short

[1]At the same time, increasing house prices provide one of the most important sources of asset value growth for households (Jordà et al., 2019).

[2]In Hungarian 'Családi Otthonteremtési Kedvezmény', abbreviated as CSOK

[3]The program since then has been extended to two children, with a substantially lower amount per child.

run (Banai et al., 2019; HNB, 2019). On the one hand, addressing the credit constraints of young households should lead to better housing conditions and fertility outcomes, especially considering direct incentives in this case. On the other, in an environment with an inelastic housing supply, the policy could result in a substantial increase in house prices which could force some families out of specific real estate markets, decreasing their welfare, and might result in the redistribution of resources between different types of households.

Due to the long-term nature of both mortgage and fertility decisions, dynamic models are the natural choice in the economic literature to analyze them, as fertility decisions and residential choices are jointly determined, possibly evolving with the housing markets in an endogenous manner. As mentioned, the housing market is documented to influence household welfare and decisions regarding fertility. At the same time, the other direction of causality is also present, as activity in the housing market is similarly affected by fertility decisions. Relatively higher mobility in the spatial dimension is well-documented for younger households, in which childbirth plays an instrumental role. Families look for areas with better amenities (Gambaro et al., 2017), which could indeed result in better life outcomes for small children (Chetty et al., 2016). Moving to agglomerations of larger cities can provide better amenities and labor market conditions, while it might increase the costs of housing, transportation, and consumption (Davis et al., 2014; Combes et al., 2019). It is also documented that families with different fertility choices can be observed to sort into different types of housing, more spacious dwellings giving a home to families with more children (Kulu and Vikat, 2007; Kulu and Steele, 2013; Chudnovskaya, 2019). Capturing endogenous relations of fertility and housing with dynamic structural models is critically important when we examine policies such as the Family Housing Allowance, as the reduced form approach cannot capture long-term unintended, potentially substantial consequences. An example of this issue, as studied by Parent and Wang (2007), is a benevolent expansion of the Canadian family tax exemption, which seemed to increase fertility in the short run. Still, the cohort-level long-run analysis shows that completed fertility remained unchanged, so the actual effect of the policy appeared only in the timing of births.

According to the model developed in this paper, the Family Housing Allowance (FHA) itself would not have increased house prices in the realized magnitude if the earlier somewhat higher interest rate conditions had stayed intact. However, I find that the combination of low interest rates, and the housing allowance would result in an around 70-90% increase in house prices in the medium-run (4-6 years) both in the urban and rural areas. So while the FHA by itself would have resulted in higher ownership rates and fertility compared to the baseline, homeownership declines slightly amongst the poorer households due to elevated house prices. Consequently, according to the model specification, they are often forced to rent smaller, central-location apartments despite their adverse welfare effects. Richer families' housing conditions are not affected by the policy. Regarding completed fertility, I find that the policy changes the timing and completed fertility as well: births occur earlier in the female life cycle, and around 5-10% increase in completed fertility is implied by the model, which seems to be high considering past results on the effect of welfare policies in Hungary (Spéder et al., 2020). The fertility effects of the policy are driven by families with lower education levels (of around 6-8%). However, the model also indicates an around 2.5% increase in births for families of higher education level. So overall, the policy seems to be effective in its goals regarding fertility, but it appears to damage the prospects of poorer households regarding housing welfare.

This paper primarily contributes to the literature by constructing a life cycle model of households' joint decisions over fertility, housing size, and residential location with unobserved heterogeneity regarding the desired number of children and the endogenous evolution of house prices. It is the first attempt, as far as I know. However, in demography and economics, many stylized facts and descriptive evidence have been gathered over the years, suggesting strong interconnectedness between fertility, housing type, and location choice (Ermisch, 1999; Kulu and Vikat, 2007; Öst, 2012; Vignoli et al., 2013; Mulder, 2013; Dettling and Kearney, 2014; Day and Guest, 2016; Chudnovskaya, 2019). In this effort, I attempt to synthesize the relevant findings of macro, urban, and labor economics literature and demography on fertility, housing, and location choice. In the macroeconomic literature, authors mainly focused on studying houses as unique assets within the frames of optimal portfolio decisions, treating fertility as an issue of primarily exogenous consumption commitment (Cocco, 2005; Yao and Zhang, 2005; Love, 2010; Li et al., 2016; Fischer and Khorunzhina, 2019), endogenous fertility appears rarely and not in the context of housing (Sommer, 2016).

The effect of housing on the life cycle of young households has been studied (Li and Yao, 2007; Attanasio et al., 2012), along with how migration and home ownership interact, impacting the welfare of households (Oswald, 2019). Still, these pieces did not consider the role of endogenous fertility in housing choices. In urban economics, work has been done on the joint modeling of housing services and location choice (Ortalo-Magné and Rady, 2006; Ortalo-Magné and Prat, 2016), along with how demographic changes interplay with urban costs (Combes et al., 2019). Similarly to the macroeconomic literature, fertility has been a less studied factor in urban economics. However, in labor economics, endogenous fertility plays a central role in the study of female labor supply (Becker, 1991), with substantial consequences for long-term labor market outcomes for women (Adda et al., 2017; Eckstein et al., 2019). I have not found examples of housing aspects as choice variables. Furthermore, the model enables the study of the effect of such policies in the medium and long run, which would not be possible in a reduced form setting so close in time to the start of the policy, while also connecting to the literature of combining structural models to study the effects of policy shocks, using the housing market as an ex-post model validation relating to the idea of Todd and Wolpin (2006).

The structure of the paper is as follows. First, in Section , I describe the policy context of the question in Hungary. Then Section introduces the model in detail. Afterward, I turn to the empirical strategy of the paper in Section , where I estimate and calibrate the parameters. And finally, I use the model to run counterfactual policy experiments to evaluate the effects on fertility and housing in . I conclude in .

3.2 Context: the Family Housing Allowance program of Hungary

Since 2010, the Hungarian government has introduced several new policies targeting directly or indirectly the fertility decisions of young households (Makay, 2020), one of them is the Family Housing Allowance announced in 2014 (further on: FHA)[4]. The allowance in its original form provided families that commit to having three children in total with a 10 million

[4]Legislated by Government Decrees No. 16/2016 and No. 17/2016 (Hungarian Government, 2016a,b)

HUF (~30,000 EUR) lump sum for purchasing a newly constructed house that satisfies specific minimum quality requirements. At the same time, another 10 million HUF interest-subsidized mortgage was also introduced to go along with the allowance. Later, the policy was extended to purchasing owner-occupied dwellings and to families that plan to have one or two children, but the latter only with a substantially lower sum.

There were several restrictions and penalties built into the policy. Only couples where one of the parties is below 40 years of age could apply. If a family does not fulfill the requirement of the number of children they committed to (except for medical reasons), they must pay a penalty. If they have two instead of three children, they have to pay 7,400,000 HUF (~20,000 EUR); if less than two, they have to pay back the entire 10,000,000 HUF (~30,000 EUR), with interest at a yearly rate of five times the size of the central bank's policy rate (set at 0.9% at the time).[5] The 'schedule' of promised children depends on the number of children the family has at the time of applying for the subsidy. If they have two children (so they only promise one extra child), the couple has four years to fulfill the commitment; for two additional children, it is eight years; for three, it is ten. Another restriction is that a dwelling purchased with the subsidy cannot be resold for ten years.

The Hungarian Government declared two objectives (Hungarian Government, 2016a; Sági et al., 2017). On the one hand, she aimed to support families in raising children and purchasing new homes. Survey evidence shows that in Hungary, house ownership is perceived as a necessary condition for raising children (Szalma and Takács, 2015). Other survey evidence by Kapitány (2016) suggests that when young adults were asked to rank several obstacles to having children, housing was named a significant but not a top issue. However, two demographic groups were found for whom housing was mentioned as a more critical aspect. One group includes those who do not plan to have a child in the short run but do in the long run, and the other group consists of those who do not want more children due to inadequate housing conditions. We can think of these groups as those the government might 'intend to treat'. The other declared aim of the legislation was to revitalize a struggling housing

[5] Since then, the roll-out of penalties has already started. However, due to the interest rate hikes in 2022 and also to moratoria introduced earlier due to the COVID crisis, the original implementation did not go into action as the government capped the penalty interest rates at 5%: https://bank360.hu/blog/mar-hitelkartyakamaton-ketyeg-a-csok-buntetes-es-szemelyi-kolcsonen-a-babavaro.

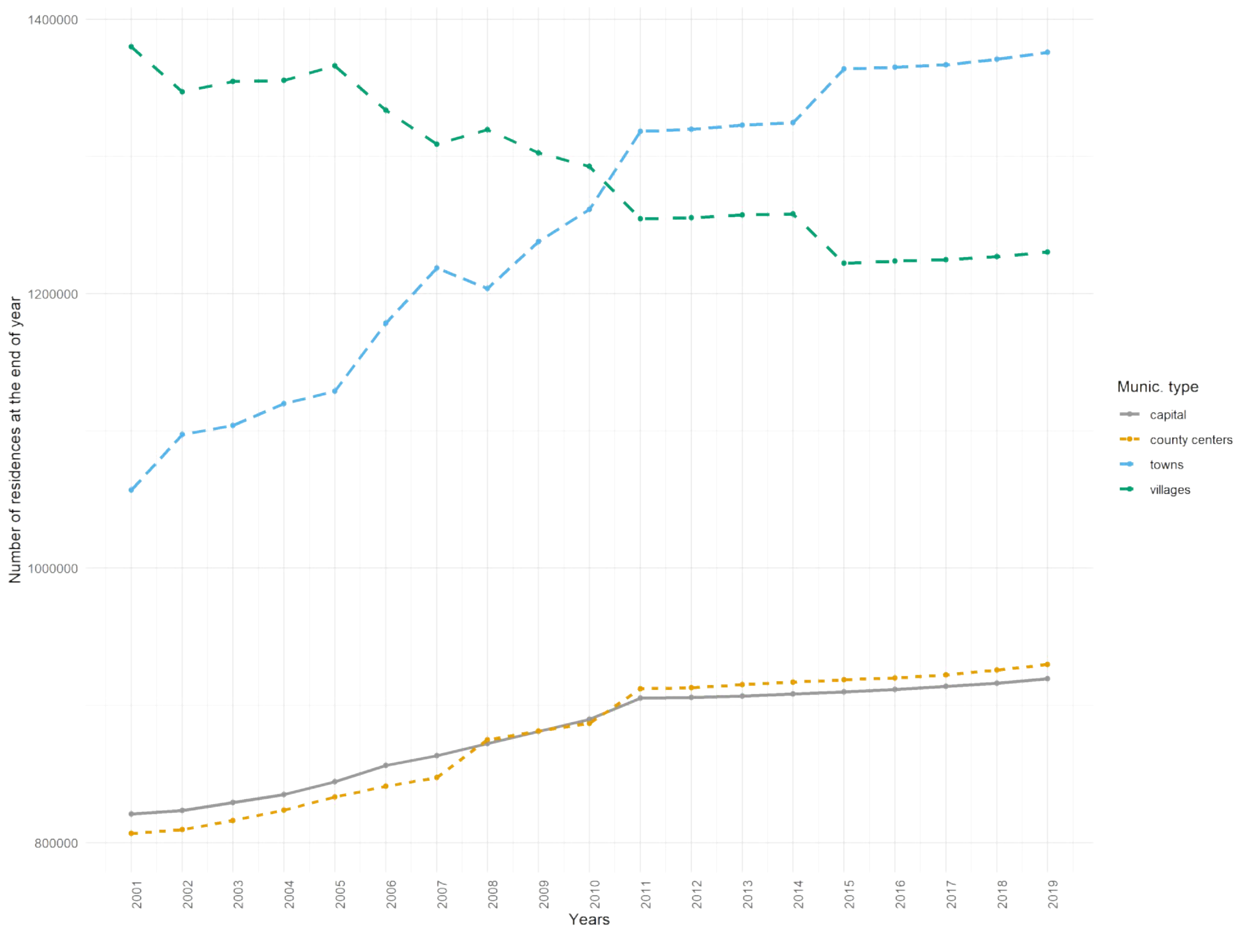
market with a demand push by young households, as it could not reach the pre-crisis output levels of the 2000s. The report of the Hungarian National Bank (HNB, 2019) shows that in 2007 the number of newly built dwellings stood at around 35,000, while in 2015, only at about 8,000, and did not show signs of recovering then.

As illustrated by Figure 3.1, after 2014 and 2015, several macroeconomic indicators potentially affected by the policy seem to have been perturbed. The average price and number of occupied houses sold increased dramatically, along with the number of marriages and internal migration. In comparison, the number of births seemed to experience more modest growth. Figure 3.2 displays the total number of houses at the end of calendar years by type of municipality. It shows that while the number of houses in rural towns seemed to increase after the policy, the rate of change for other municipality types does not seem to be substantially affected, at least in the short run.

Figure 3.1: Prices and number of sold houses, marriages, internal migration, and live births, 2008-2018

Note: The yearly time series are based on the publicly available settlement level data of the HCSO, accessed 16/02/2020. The vertical line indicates the announcement of the Family Housing Allowance program.

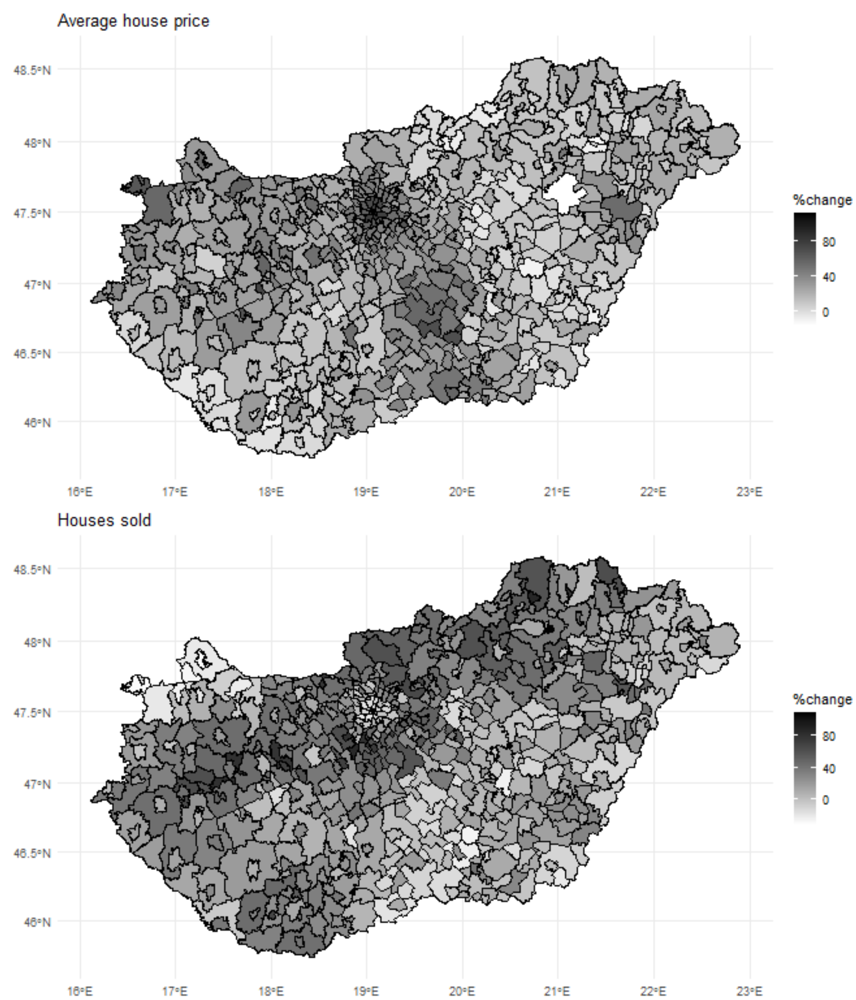
Figure 3.2: Total number of houses by municipality type



Note: The yearly time series are based on the publicly available settlement level data of the HCSO, accessed 16/02/2020.

There is also substantial heterogeneity within the country regarding prices and the number of houses sold, based on the comparison of the periods 2015-2019 and 2008-2014. Figure 3.3 shows the percentage change on the municipality level, with smaller municipalities aggregated to their larger administrative unit ('járás'). Prices increased the most in Budapest and closer towns (by more than 50%, reaching levels of 80%), while at the bottom of the distribution we find the country's eastern regions. However, the change in quantities sold correlates negatively with price changes: Budapest districts experienced a decline in the number of houses sold, while in rural areas, higher numbers were sold than before, also reflected in increasing immigration to rural areas, and decreasing immigration to the center, Budapest.

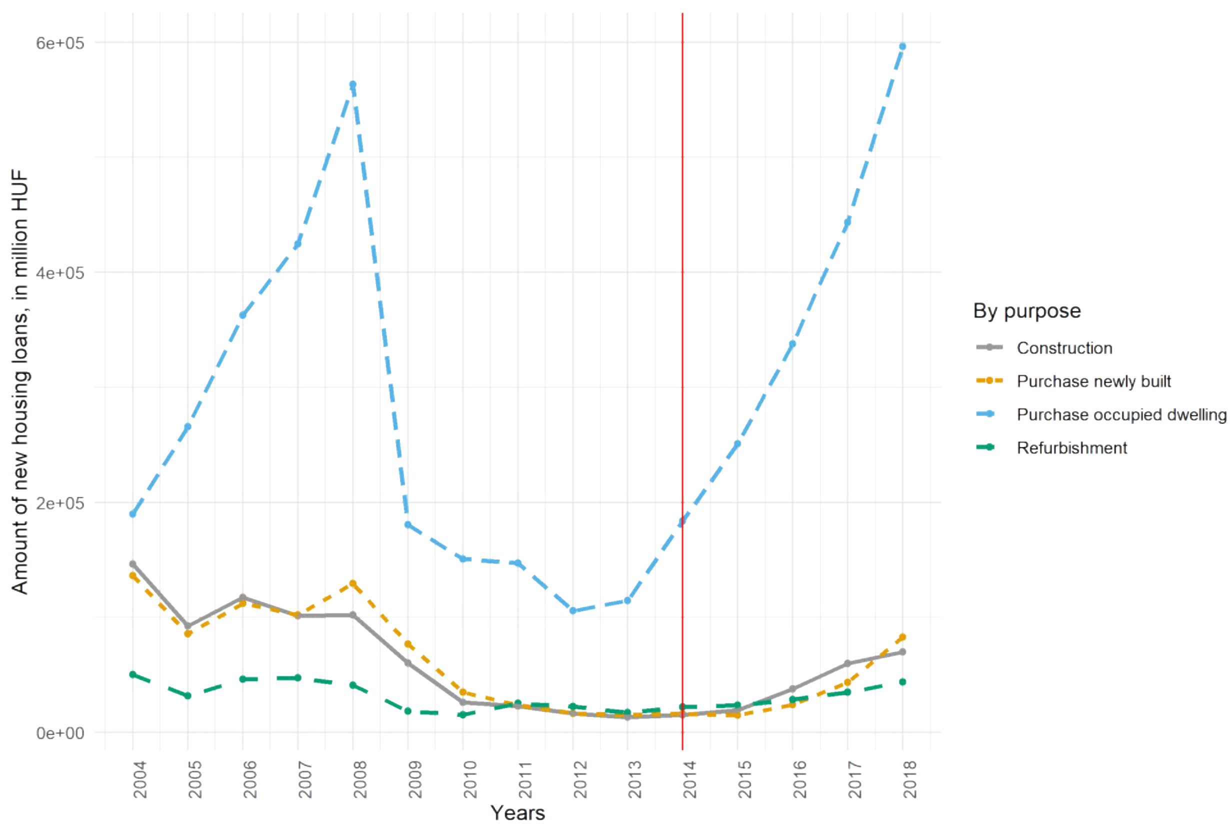
Figure 3.3: Change in average prices and number of houses sold in Hungary, 2015-2019 vs. 2008-2014



Note: based on the publicly available municipality-level data of the HCSO, accessed 16/02/2020. The time windows reflect the state of the housing market in Hungary before and after the introduction of the Family Housing Allowance.

The effects of the policy are far from trivial (HNB, 2021). The impact on housing demand seems substantial but not necessarily positive overall. Since 2016, close to 20% of all mortgage contracts are linked to Family Housing Allowance. At the same time, the surging house prices restricted the accessibility of purchasing a house (HNB, 2019). That is especially true for the capital Budapest, which experienced the largest growth of house prices within the last decade in the European Union, reaching by the second quarter of 2019 around 230% of the 2009 price level. The supply side effect also manifests in the growing number of permits issued starting from 2015; the number of newly built dwellings remains at a low yearly level of 12,000, less than half of the level in 2007. Figure 3.4 displays the amount of newly issued housing loans, which shows that the growth in the related outstanding debt can be mainly attributed to purchasing occupied dwellings. A debate about whether the FHA policy could have affected the number of births in Hungary is also ongoing. Kapitány and Spéder (2020) suggests that the increasing number of live births fit into the regional trends of recovering fertility levels and might have happened regardless of the family policies. At the same time, Sági et al. (2017) shows indicative evidence that some of the growth could be attributed to the new FHA policy.

Figure 3.4: The amount of new housing loan value issued in millions of HUF by purpose, 2004-2018



Note: based on the publicly available data of the Hungarian Central Statistical Office.

There are also other stylized facts to consider, which caution against overly simplistic interpretations of the data points, and show important limitations to the model this paper presents. Based on the survey published by the Hungarian National Bank HNB (2019), between 2015 and 2019, more than 40% of all house purchases in Budapest could be categorized as 'with investment purpose'. At the same time, the fraction of 'first home purchases' has fallen from around 35% to 20% of all transactions. While the FHA policy aims at buying newly constructed houses, their magnitude is small compared to the approximately 160,000 house market transactions a year, as shown by Figure 3A9, even though the entire market might be affected due to substitution effects.

3.3 The model

3.3.1 Demand side: the problem of the household

The demand side of the model examines how fertility, residence size, and location choice of young households evolve over the life cycle, with a finite number of periods. Household formation and divorce are not included in the model, the couple has already finished their education, and the female member is 25 years old. Each period, they make a forward-looking decision to have an additional child, and about the size of the house, ownership, location, and choosing between non-durable consumption and savings. They can also apply for a government subsidy representing the Hungarian government's Family Housing Allowance (FHA), providing a one-time, lump-sum subsidy in case of a house purchase with the additional commitment to having three children. Each household i is characterized by their desired number of children (ν_i), the education of the male (ed_i^M), and the education of the female (ed_i^F).[6] Education levels are observed, but the desired number of children remains hidden, providing the source of unobserved heterogeneity in the model. However, the population distribution of desired children can be estimated, which allows for assigning households a value drawn from this distribution. Kapitány and Spéder (2020) also reports that these preferences have stayed stable during the last decades. This approach allows the model to avoid assuming an error term distribution driven by the inference strategy itself, as we can treat this distri-

[6] I differentiate between different levels of education to be able to address the redistributive aspects of the policy, and education levels are the main factors in the data that allows for it.

bution known. Households also face idiosyncratic uncertainty due to unemployment shocks, infertility shocks, and global uncertainty due to changing house prices.

The household's utility function $u(.)$ depends on the non-durable period consumption and other relevant state variables: the number of children, the size of the residence, its location, and ownership status. Novelty in this paper's approach is that each household has its desired number of children, compared to the usual way of treating children in the utility as purely 'normal goods' such as consumption. So more than the desired number will decrease utility ceteris paribus, implying that households would not opt to have infinitely many children, even with unbounded resources. Hence, the housing allowance might induce households to end up with more than their desired number of children, as using the subsidy enables them to reach better housing conditions.

Another feature of the model is that 'housing services' receives an intuitive meaning as the household suffers disutility from the crowdedness of the dwelling, by having the number of children divided by the house size in the utility function. This component provides a natural mechanism for larger families to sort into larger houses. House sizes are discretized instead of continuous housing services, and households derive extra utility from ownership, both aspects also featured in Attanasio et al. (2012).

Households suffer disutility (or gain utility) due to distance from the central location, resembling the standard monocentric city model (Duranton and Puga, 2015). It represents amenity or disamenity effects of living in the central area of the economy, an idea appearing, for instance, in Brueckner and Zenou (1999) or Combes et al. (2019). In the model, I introduce two locations representing the center and rural areas, each with a separate housing market supply, connected by the demand for housing from households concerning both locations. Wages received by the agents also differ according to the location of the household, calibrated to the Hungarian case: wages are higher in the central location but not reacting endogenously to household decisions.

I introduce two house sizes. Size 0 represents a 50-m^2-apartment (approximately one or two rooms), while Size 1 represents the ownership of a 100-m^2 house (three or more rooms). The price of housing can vary by location but is given in m^2-prices; hence the price of 50 m^2-

residence is precisely half of a 100-m^2 residence[7]. Table 3.1 shows the empirical counterparts of these idealized concepts based on the 2011 Census of Hungary. Looking at all residences in Hungary, around 2 million have one or two rooms and around 2 million have three or more, with median sizes of 54 and 90 m^2: these types of housing are represented in the model. Not being exposed to costs of owned housing is another simplification that captures the idea of imputed rents, which is found to benefit house owners substantially (Kilgarriff et al., 2019). Households can choose between renting a house and purchasing one while selecting the house size. If the residence is not owned, the household must pay rent, set at 5% of the house price corresponding to a Hungarian context (and similarly to Attanasio et al. (2012)).

Table 3.1: Distribution of house sizes in Hungary, 2011

No. of rooms	No. of houses	Median size (m2)	Mean size (m2)	Std. Dev.
1-2	2,135,857	54	56.93	19.41
3+	2,233,623	90	93.99	29.14

Note: based on the 2011 Census of Hungary.

I incorporate the Family Housing Allowance into the model in the following way. As the policy is intricate and requires too many new state variables, I refrain from including all details. First, I only consider the original version of the allowance with the condition of three children, as the size of the subsidy is considerably larger than for two children (10 million HUF compared to 2.6 million HUF). Second, the fertility commitments of the households are checked only at the terminal period, representing the end of fecundity, even though the actual regulation gives couples four years to have one additional child, eight years for two, and ten years for three children. This way, I avoid tracking the children's birth years, reducing computational needs significantly. Third, the dwelling purchased using the FHA subsidy cannot be sold in the model, even though selling is forbidden only for ten years; this simplifies the model's state space. Fourth, a family can only apply for the subsidy when they purchase a large house, here in the model, a 100 m^2 house. This choice represents that

[7] The ratio of the two house prices could have been set as a parameter such as Attanasio et al. (2012), here I simplify to have fewer parameters by assuming that house prices scale linearly with size

the policy has a minimum 60 m^2 built-in threshold regarding size[8]. And finally, only those households are eligible for which the age of the household (age of the female) is lower than the maximum age for which birth is possible, which is set at 40 years in the model.

The problem setup builds on the standard macroeconomic mortgage models, presented in Davis and Van Nieuwerburgh (2015). Households are modeled to be alive for a finite number of periods, making decisions from 25 until 45 years of age of the female (10 periods, or 20 years), and living an additional 20 years [9]. When they remain both in the same type of house and the same location, they are considered not to move; otherwise, they are always assumed to be able to sell or buy for the house's market price. In case they move to a residence of a different size or a different location, they suffer a moving cost as in Attanasio et al. (2012), set at 10% of the price[10], representing transportation and legal costs of the choice. Houses are expensive (prices are calibrated to reflect the property prices of Hungary), and households can use savings, their property, and mortgages to purchase them. If they do not own a house, they must rent an apartment. Outside of mortgage loans, households are only allowed to have positive savings, which they could use for consumption smoothing. The repayment schedule of mortgages is not specified, providing flexibility and inclusion of several regimes, although different institutional contexts are found to affect household behavior (Chambers et al., 2009).

Choice variables

Choice variables include non-durable consumption ($c_{i,t} \geq 0$) set simultaneously with the next period's savings ($S_{i,t+1} \in \mathbb{R}$), the willingness to have an additional child ($n_{i,t} \in \{0,1\}$), house ownership ($o_{i,t} \in \{0,1\}$), size ($h_{i,t} \in \{0,1\}$) and location ($l_{i,t} \in \{0,1\}$), the choice of taking out a mortgage ($m_{i,t} \in \{0,1\}$), and the choice of applying for the government subsidy Family Housing Allowance ($f_{i,t} \in \{0,1\}$).

The rules for the choices are as follows. Each household can own up to one residence.

[8]Note, that a study of the Hungarian National Bank (HNB, 2019) finds that this hard threshold itself might have induced a distribution change of size among newly built apartments such that instead of the 50 m^2 apartments 60 m^2 ones have started to be constructed.

[9]About the impact of choosing finite vs. an infinite number of periods, see Hedlund (2018).

[10]Compared to Attanasio et al. (2012) which uses 5% for the United Kingdom, I implement a 10% cost to reflect the considerably lower levels of real estate prices in Hungary

Savings can be used for consumption smoothing or mortgage repayments; however, a household can only get indebted when they take out a mortgage. This restriction is in accord with evidence that historically, mortgage debt has constituted around 70% of all household debt (Piazzesi and Schneider, 2016). Following Attanasio et al. (2012), I allow households to take out a mortgage even if they choose not to move (borrow against their property), following the results of Cloyne et al. (2019), who find that house prices are the drivers of households' ability to borrow via collateral effects. Households are not allowed to have more than one mortgage contract at the same time, but they are allowed to apply for a new one after repaying the previous one. They are also forbidden to sell their house if they have a mortgage against it. They are also prohibited from selling using the FHA subsidy if they purchase it. If they are indebted in one period, they must weakly increase their position over time. Children can only be conceived until 15 years into the model, representing the end of the female fertility cycle at around 40 years. They are forbidden to apply for the subsidy past that age, which condition reflects the regulation's intentions.

State variables

Households are assigned at the beginning their desired number of children ($\nu_i \in \{0, 1, 2, 3\}$), and the education of male and female adults ($\mathrm{ed}_i^M \in \{0, 1\}, \mathrm{ed}_i^F \in \{0, 1\}$, 1 representing completed tertiary education). $\Omega_{i,t}$ collects the state variables for household i in period t, consisting of the following: house prices ($p_t^H(h_{i,t}, l_{i,t}), \forall h \in \mathcal{H}, l \in \mathcal{L}$), employment status ($\mathrm{e}_{i,t}^M \in \{0, 1\}, \mathrm{e}_{i,t}^F \in \{0, 1\}$), number of children ($N_{i,t} \in \{0, 1, 2, 3\}$), current house type and location ($H_{i,t} \in \mathcal{H} = \{0, 1\}, L_{i,t} \in \mathcal{L} = \{0, 1\}$), and the savings of the household ($S_{i,t} \in \mathbb{R}$), with additional states tracking the household's status status regarding mortgage ($M_{i,t} \in \{0, 1\}$), ownership status ($O_{i,t} \in \{0, 1\}$) and the government subsidy ($\mathrm{F}_{i,t} \in \{0, 1\}$ for Family Housing Allowance). Lastly, households face infertility shocks ($\iota_{i,t} \in \{0, 1\}$), which are realized only after deciding on having an additional child or not. The distribution of infertility shocks depends on the female's age and education and is considered to be known by the households.

The number of children, the house size and location, savings, mortgage, and government subsidy are determined in the previous period endogenously by the optimal choices of the households, taking the infertility shocks into account. I do not introduce uncertainty into the household's survival, examined, for instance, for the case of divorce by Fischer and Khorunzhina (2019), or death. It is plausible that the policy affects couple formation and separation as well. However, these aspects are beyond the scope of this paper.

Stochastic processes

There are five stochastic processes in the model: the employment status of the male and the female adults of the households separately, infertility shocks depending on the age and education of the female, and the house prices of the central and the rural location. Unemployment shocks are realized before the choices of each period, and house prices are determined jointly with household decisions. In contrast, infertility shocks are realized after choices are made, but the distribution of infertility shocks is known to households.

The employment status of the male and the female adults follow separate, two-state Markov processes, which depend on the level of education[11]. Households form correct beliefs over the transition probabilities of employment status. These probabilities are set such that the stationary state unemployment rate corresponds to the long-run unemployment rates in Hungary conditional on education for the relevant period, which is around 9% for lower than tertiary, and 3% for tertiary educated.[12].

House prices are calculated as feasible 'pseudo-equilibrium' outcomes, which concept I introduce later in detail. Households are assumed to be price takers, and they form their expectations in a naive way over future m^2-prices such that $\mathbb{E}_t[p^H_{l,t+1}|\Omega_{i,t}] = p^H_{l,t}$, for $l \in \{0,1\}$. This expectation formation can also be seen as households falsely believing that the time series of house prices follow a simple random walk process. Even if this belief is false within the model, there is evidence that real house price time series are indeed random walks (Holly et al., 2010). It is also worth noting that more sophisticated expectation formation could be introduced to the model, which is a possibility for a future extension.

Infertility shocks are realized after households have decided to have an additional child.

[11] Although there could be good reasons to consider location-dependent employment transitions, a possible future improvement of the model.

[12] Based on the data of the Databank of the Institute of Economics, available as of 03/02/2021 here, and here.

At the same time, they are also aware that having a child at different points in their life cycle implies different levels of infertility risk. I estimate these distributions conditionally on female age and education[13], as the fraction of miscarriages of births added to miscarriages, for the period 2004-2014, using the complete individual-level administrative data collected by the Hungarian Central Statistical Office. Note that due to miscarriage events often not being discovered or reported, I probably underestimated this probability. Also, note that I do not consider abortion in the model or undesired pregnancies.

Initial and terminal conditions

Households are assigned an inheritance in the form of home ownership, drawn from the distribution estimated from the available Household Budgetary Survey of Hungary for the period 2004-2014 (introduced later in more detail), along with house size, mortgage, and location distributions at the age of 25 of the female. Only those households can have mortgages at the start that are simultaneously assigned home ownership, following the model's logic. These households also start with negative savings corresponding to the down payment for the owned house, set at 50% of the value of the home priced at the 2004-2014 levels, with a hard cap of 10 million HUF to avoid poorer households not being able to repay their debt.

Households start employed, with their education and their desired number of children fixed. The education group, along with the number of initial children in the household, are drawn from their joint distribution estimated from the 2011 Census ($\mathbb{P}^{\mathrm{ed}^{M,F},N}$) (at female age of 25). The joint distribution captures assortative mating and selection into education tracks due to preferences, studied for instance in Adda et al. (2017). The desired number of children is drawn from population distribution $\mathbb{P}^{\nu}$, independently from the parents' educational attainment, as this information is unavailable.

The household's government subsidy uptake also impacts the terminal period savings. Households are allowed to violate the commitment to the government regarding children, but it comes with a penalty. If they decide not to respect the requirement of having the number of children they committed to, they are forced to pay back the remaining corresponding

[13] I also estimated the distributions w.r.t. child parity, but it did not change the probabilities significantly, so I kept the simpler version.

subsidy amount to the government as a lump sum at the end. If they have two instead of three children, they must pay 7,400,000 HUF (∼20,000 EUR). If less than two, they have to pay 10,000,000 HUF (∼30,000 EUR)[14]. Defaulting on the debt is not allowed in the model.

The terminal value $V_{i,T+1}^{\nu_i,\mathrm{ed}_i^M,\mathrm{ed}_i^F}$ captures an additional 20 years of life with the same housing conditions and the number of children they have at time T. The non-durable consumption value $\underline{c}_i$ is imputed into the post-terminal utility function, which is calculated the following way. Households are required to have a non-negative asset position by the end of the time horizon, set at 20 years. After checking the conditions for the number of children due to the government subsidy, the residual savings after calculating any penalties are checked to be non-negative. These savings are then consumed in equal shares for the following twenty years[15]. The net household income is also assumed to remain constant for future periods, and the probability of unemployment is set to 0. Specifying a terminal value is necessary from a modeling perspective as households would often sell their residence in the last period if they see no further utility in keeping it. It is equivalent to a particular type of bequest where the utility derived from the additional 20 years of owning the residence provides the bequest incentive.

Period utility function

Households derive instantaneous utility from non-durable consumption, the number of children, the crowdedness of the house, location, and home ownership in the following way:

$$u_{\mathrm{ed}^F}(c_{i,t}, N_{i,t}, H_{i,t}, L_{i,t}, O_{i,t}) =$$
$$\frac{(c_{i,t}-1)^{1-\gamma_{\mathrm{ed}^F}}-1}{1-\gamma_{\mathrm{ed}^F}} \exp\Big(- w^N_{\mathrm{ed}^F}(N_{i,t}-\nu_i)^2 - w^H \frac{N_{i,t}+1}{I[H_{i,t}>0]+1} - w^L L^2_{i,t} + w^O I[O_{i,t}>0]\Big)$$

where $I[.]$ denotes the indicator function. The utility function is multiplicative, reflecting complementarity between the components of the function. As mentioned earlier, ν_i denotes the desired number of children, which is the source of unobserved heterogeneity, responsible

[14] I abstract away from the penalty interest paid, as it does not change the magnitudes of the penalty even if they might be significant, but would introduce unnecessary complexity in the computations.

[15] There could be a future point of improvement by setting consumption to be based on the consumption Euler equation. In this version, I stayed with the simpler setup

for observationally identical households ending up with different choices. However, I censor the distribution to include only 0-3 preferred children, with values above three incorporated into the highest category. I also let agents have three children at maximum. Notice that households with lower and higher educated females are allowed to have different risk aversion parameters (γ_{ed^F}) and different preferences for children ($w^N_{\text{ed}^F}$). The latter concept is represented as suffering disutility from being further away from the desired number. Allowing for education-dependent parameters captures the previous self-selection into different education and career tracks according to preferences about children and consumption, as indicated by Adda et al. (2017). Consumption is constrained from below by 1 with a price of p^c, which denotes the living costs corresponding to the subsistence level based on the calculations of the Hungarian Central Statistical Office (HCSO, 2016)[16]. House size $H_{i,t}$ is coded as a binary variable, representing 50 and 100 m^2 apartment sizes as 0 and 1. However, the number of children also affects the crowdedness of the house (third term in the utility function), with disutility calculated as the number of children over the house size. House size is modeled such that the small one is half the size of the large apartment, be it owned or rented[17]. The household also derives utility/disutility from being further away from the city center, representing the city center amenities. Ownership status also provides positive utility, as it is generally assumed and found in the literature (Davis and Van Nieuwerburgh, 2015).

Budget constraint

$$p^c\left(1+\frac{N}{3}\right)c_{i,t} + S_{i,t+1} + \kappa(n_{i,t}, N_{i,t}) + p^H_t(h_{i,t}, l_{i,t}) + \mu(H_{i,t}, L_{i,t}, h_{i,t}, l_{i,t}, \mathrm{F}_{i,t}, M_{i,t}) +$$
$$\rho(H_{i,t}, L_{i,t}) \leq NW(W^M_{i,t}, W^F_{i,t}) + (1+ir)S_{i,t} + \mathrm{FHA}(f_{i,t}) + p^H_t(H_{i,t}, L_{i,t})$$

The budget constraint builds on the baseline setup of macroeconomic models with mortgage decisions (Davis and Van Nieuwerburgh, 2015). The first term is the value of the total household consumption $p^c\left(1+\frac{N}{3}\right)c_{i,t}$, with a child's consumption corresponding to about one-third of an adult couple, per the estimates of the Hungarian Central Statistical Office

[16] I do not normalize by this value to impute nominal prices more easily during the model development.
[17] Differently from Attanasio et al. (2012), renting larger houses is allowed for the households.

(HCSO, 2016). This term also introduces the flow cost of children in terms of consumption such that higher consumption of adults in the household also induces higher spending on children in a linear way. Therefore, we implicitly assume a structure of preferences regarding child quality, which is omitted from this model (discussed in Sommer, 2016). As discussed earlier, the next period's savings are denoted by $S_{i,t+1}$. The function $\kappa(n_{i,t}, N_{i,t})$ denotes the costs and benefits associated with children in the household. It includes child benefits, however, also the out-of-pocket spending for giving birth in Hungary, which might be substantial, even if not reported officially[18]. The function $p_t^H(h_{i,t}, l_{i,t})$ assigns a price to a house with house type $h_{i,t}$ and location $l_{i,t}$, while $\rho(H_{i,t}, L_{i,t}, O_{i,t})$ gives the rental cost or house service costs associated with the house type at t (the latter assumed to be 0 for now). Finally, the function $\mu(.)$ introduces moving costs, which gives incentives for agents to remain sedentary (as often used in the literature, per Davis and Van Nieuwerburgh, 2015). The moving costs are set up such that houses under mortgage commitment cannot be sold while being 10% of the house price in case of selling.

The revenue side of the budget is the following. The net income of the household, $NW(.)$, is the function of the gross incomes of the male and female in the household[19]. Unemployment spells are set to one year, during which the unemployed person receives lower wages[20]. In the case of a new child, the woman's salary is accounted for as two-thirds approximating the amount of the benefits maternity policies provide in Hungary (Makay, 2020).

Gross income is given by a Mincerian reduced form equation of education, employment status, location choice, and experience (here measured by the age of the household). Women gather less experience due to their number of children to address the opportunity cost of childbirth, and each child counts as two years less experience[21]. The equations are the following, where $e_{i,t} \in \{0, 1\}$ represents employment, and $W_{\min}$ denotes the amount representing

[18] Based on the blog of the Hungarian National Bank, in Hungarian: https://novekedes.hu/elemzesek/ne-kamuzzunk-sulyos-szazezrekbe-kerul-magyarorszagon-az-ingyenes-szules

[19] Since 2011, the number of children plays a more significant role in personal income taxation, however for simplicity in the model, I set taxes at 42%, consisting of a 15% personal income tax, and a 27% social contributions, reflecting the taxation of the era.

[20] Calibrated to the minimum wage as of 2011 at around 80,000 HUF a month

[21] Following maternity benefits design of Hungary, Makay (2020)

unemployment benefit:

$$W_{i,t}^{M} = \mathrm{e}_{i,t}\exp(\beta_0^{W,M} + \beta_1^{W,M}X_{i,t} + \beta_2^{W,M}X_{i,t}^2 + \beta_3^{W,M}\mathrm{educ}_i^M + \beta_4^{W,M}L_{i,t}) +$$
$$(1-\mathrm{e}_{i,t})W_{\min}$$
$$W_{i,t}^{F} = \mathrm{e}_{i,t}\exp(\beta_0^{W,F} + \beta_1^{W,F}\max\{X_{i,t}-2N_{i,t},0\} + \beta_2^{W,F}\max\{X_{i,t}-2N_{i,t},0\}^2 +$$
$$\beta_3^{W,F}\mathrm{educ}_i^F + \beta_4^{W,F}L_{i,t}) + (1-\mathrm{e}_{i,t})W_{\min}$$

The parameters are estimated by regressing the log of gross wages on experience, education, and location, separately for the two genders[22]. Note that these parameters should not be interpreted as causal effects of the variables; they should only be interpreted as coefficients of the best linear prediction.

Recursive form

The household problem can be summarized in a recursive form in the following way:

$$V_{i,t}^{\nu_i,\mathrm{ed}_i^M,\mathrm{ed}_i^F}(\Omega_{i,t}) = \max_{\substack{c_{it}\geq 1, n_{i,t}\in\{0,1\},\\ h_{i,t}\in\mathcal{H}, l_{i,t}\in\mathcal{L}, o_{i,t}\in\mathcal{O}\\ m_{i,t}\in\{0,1\}, f_{i,t}\in\mathcal{F}_{i,t},\\ S_{i,t+1}\geq S_{i,t}I[S_{i,t}<0]+\\ -m_{i,t}(1-\delta)p^H(h_{i,t},l_{i,t})I[S_{i,t}\geq 0]}} u_{\mathrm{ed}^F}(X_{i,t}) + \beta\mathbb{E}_{\sigma_{i,t+1}}[V_{i,t+1}^{\nu_i,\mathrm{ed}_i^M,\mathrm{ed}_i^F}(\Omega_{i,t+1})|\Omega_{i,t}]$$

$$X_{i,t} = (c_{i,t}, N_{i,t}, H_{i,t}, L_{i,t}, O_{i,t})$$

$$\Omega_{i,t} = (\sigma_{i,t}, H_{i,t}, L_{i,t}, O_{i,t}, S_{i,t}, M_{i,t}, \mathrm{F}_{i,t})$$

subj. to:

[22] I omitted here the variance parameter of the log-normal distribution, which would be the appropriate form for expected value.

Budget constraint:

$$p^c\Big(1+\frac{N}{3}\Big)c_{i,t}+S_{i,t+1}+\kappa(n_{i,t},N_{i,t})+p_t^H(h_{i,t},l_{i,t})+\mu(H_{i,t},L_{i,t},h_{i,t},l_{i,t},\mathrm{F}_{i,t},M_{i,t})+$$
$$\rho(H_{i,t},L_{i,t})\leq NW(W_{i,t}^M,W_{i,t}^F)+(1+ir)S_{i,t}+\mathrm{FHA}(f_{i,t})+p_t^H(H_{i,t},L_{i,t})$$

State transitions:

$$N_{i,t+1}=N_{i,t}+n_{i,t}\cdot\iota_{i,t},L_{i,t+1}=l_{i,t},H_{i,t+1}=h_{i,t},F_{i,t+1}=F_{i,t}+f_{i,t}$$
$$M_{i,t+1}=\begin{cases}0, & \text{if } M_{i,t}=1 \text{ and } S_{i,t+1}\geq 0\\ M_{i,t}+m_{i,t}, & \text{otherwise}\end{cases}$$

Initial conditions:

$$\nu_i\sim\mathbb{P}^{\mathrm{nu}},\iota_{i,t}\sim\mathbb{P}_t^{\mathrm{ed}_i^F},$$
$$(\mathrm{ed}_i^M,\mathrm{ed}_i^F,N_{i,0})\sim\mathbb{P}^{\mathrm{ed}^{M,F},N},(\mathrm{e}_{i,0}^M,\mathrm{e}_{i,0}^F)=(1,1),$$
$$H_{i,0}\sim\mathbb{P}^H,L_{i,0}\sim\mathbb{P}^L,O_{i,0}\sim\mathbb{P}^O,M_{i,0}\sim\mathbb{P}^M$$

$$\textit{Terminal value: } V_{i,T+1}^{\nu_i,\mathrm{ed}_i^M,\mathrm{ed}_i^F}(\Omega_{i,T+1})=\sum_{s=0}^{19}\beta^s u_{\mathrm{ed}^F}(\underline{\mathrm{c}}_i,N_{i,T+1},H_{i,T+1},L_{i,T+1})$$

where i denotes a household, t denotes periods which at this point correspond to years of age for the female adult of the household, $\sigma_{i,t}$ collects the random states of employment, house prices, and the number of children, that the agents take expectation over. $u(.)$ denotes the utility or period utility function, β denotes the discount factor. Interest rate (ir) is set as a baseline at 5% a year[23] reflecting the 2013-2019 mortgage interest rates of Hungary (HNB, 2019), and $\beta=\frac{1}{1+ir}$ is set accordingly. The down payment ratio (δ) is also fixed, as a baseline at 0.5[24]. Using fixed interest rates implicitly assumes fixed-rate mortgages (vs. adjustable-rate mortgages). This type of mortgage contract has been the most prominent option in Hungary in the last decade (HNB, 2019), and it constitutes the majority in the U.S. as well (Piazzesi and Schneider, 2016).

[23]In case of biannual periods, it is set as $ir=(1.05)^2-1$

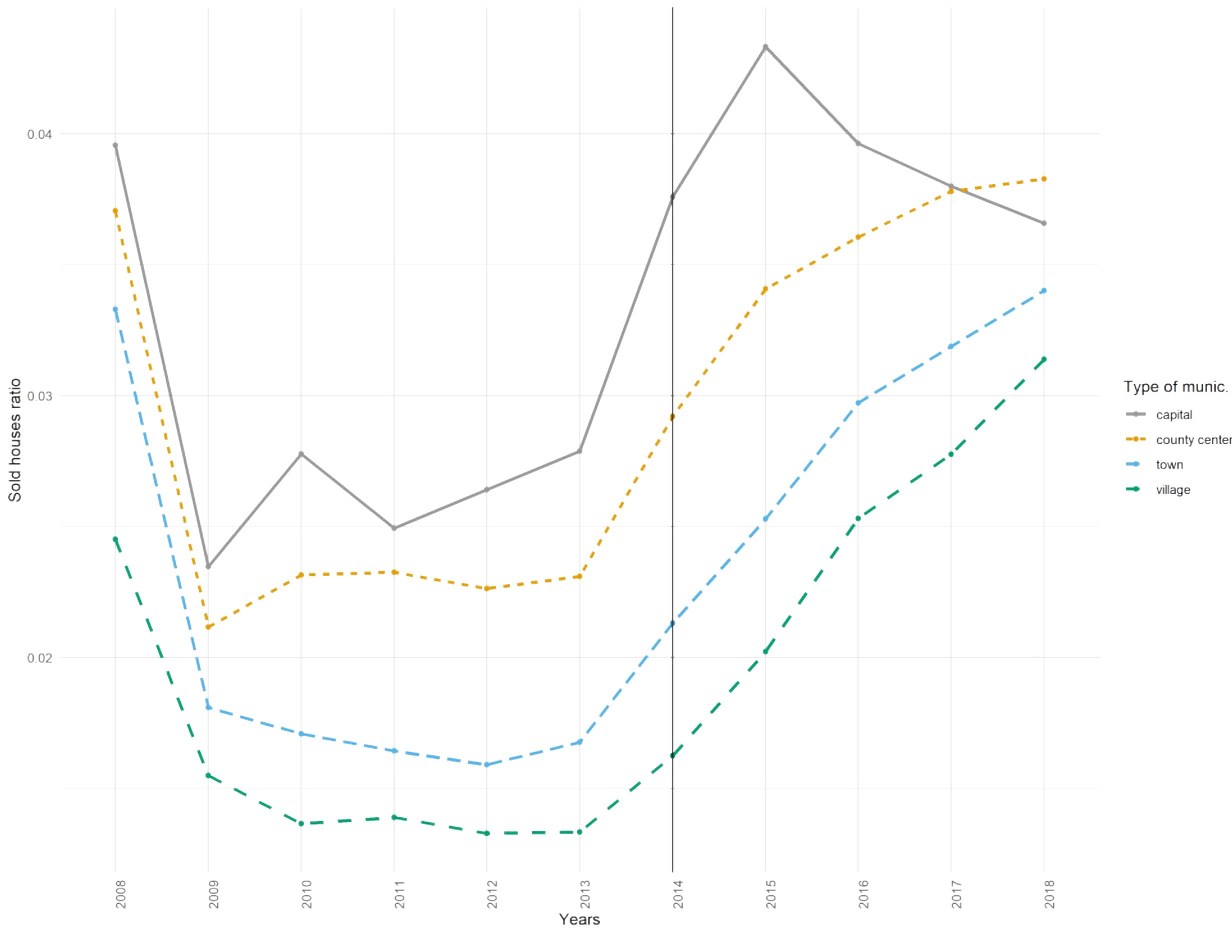
[24]According to the calculations of the website Bankmonitor.hu, specializing in mortgage loans, banks typically required a 0.4-0.5 down payment ratio in Hungary. Source in Hungarian:, accessed 29/07/2020.

3.3.2 House supply

House supply introduces several complexities analytically and computationally. Hence, many in the literature treat house prices as exogenous to the household problem (Duranton and Puga, 2015). The price of housing is composed of two elements: the price of the land and the structure itself. Piazzesi and Schneider (2016) documents that the overwhelming majority of the historical variation of house prices can be attributed to land prices. It is easy to see that rising land prices can result from a sudden increase in demand, which an inelastic supply cannot follow fast enough. In my paper, I incorporate this phenomenon, which requires some level of equilibrium responses of prices. There are instances in the literature where supply is explicitly modeled, such as Glaeser et al. (2008); however, their model would not be feasible in this setting.

In my model, the supply of housing has three components. First, there is a constant per period available m^2 space for newly built houses, governed by calibration parameters for each location denoted as (α^0, α^1), representing a 'fixed stream of housing services' (Piazzesi and Schneider, 2016) often used in the literature. The revenue generated by selling these houses are not contributing to the household budgets, and I do not consider this income in the model. The available external supply of housing measured in m^2 is then calculated the following way for location l: $H_l^S = \alpha^l \cdot |\mathcal{I}| \cdot 50$ where $|\mathcal{I}|$ denotes the number of households in a cohort, in the model. So the parameter captures the relation between the amount of newly available space and the number of new households entering the housing market on the demand's size. This component has the most relevance in this policy context as the Family Housing Allowance originally aimed at purchasing newly constructed dwellings. In this version of the model, this component is not reacting to price changes corresponding to a fully inelastic housing supply; experimenting with additional elements to the supply side could provide valuable model extensions. However, as Figure 3.5 shows, the number of real estate transactions is less than 5% of the total available houses for all municipality types, suggesting that even with sharply increasing house prices, an overwhelming fraction of owners do not sell their property. Still, at the same time, as we have seen earlier, house transactions dominantly occur in the domain of owner-occupied properties.

Figure 3.5: Share of houses sold of the total, by type of municipality, 2007-2018



Note: the author's calculations, based on the publicly available data of the Hungarian Central Statistical Office.

Second, I introduce a set of 'old' households starting with homeownership and maximizing expected utility similar to the young households in a simplified model: without the ability to have children or to take up a mortgage or the government benefit. They react to price changes and present a finite stock of housing available on the housing market, responding to changing prices and decreasing as we move forward in time. However, in other model applications, this part of the supply could play a significantly more important role. And third, young households also constitute part of the housing supply if they change their residence. It is worth mentioning that I abstract away from several aspects which could play a role, such as the quality of the apartments, optimal portfolio decisions, AirBnB, and, most crucially, speculative or investment-focused house purchases are not considered here. These are all aspects that might be worth investigating further in the future.

3.3.3 'Pseudo-equilibrium' in the housing market

House prices are present in the model as state variables, which must have a small number of discrete levels due to constraints of feasibility. That setup does not allow for an actual temporary equilibrium price discussed in Piazzesi and Schneider (2016), which would clear the housing market for each location in each period. Instead, I employ the following algorithm that enables period-by-period 'pseudo-equilibrium' prices and allocations:

1. Fix period t

2. For each combination of possible location-m^2-prices collected into vector $p^H \in \mathbb{R}^2_+$, we calculate for each household i and for each location l the net demand for total space, denoted by $h^D_{i,l}(p^H)$

3. For each combination of location-m^2-prices, we also calculate the external net supply for each location l as a sum of fixed external stream of housing, added to the housing sold by the old households, denoted jointly by $h^S_{j,l}(p^H)$, j indicating different sources of supply

4. We sum up the net demand for space of the households, and the net supply from different sources, for each location

5. We calculate the squared difference of the total net demand from the total net supply and select the m^2-price yielding the least distance as the 'pseudo-equilibrium' price

So the temporary 'pseudo-equilibrium' price vectors for the housing markets can be defined in the following way:

$$(p^H)^* \in \text{argmin}_{p^H \in \mathbb{R}^2_+} \sum_{l \in \{0,1\}} \omega^l \left(\sum_{i \in \mathcal{I}} h^D_{i,l}(p^H) - \sum_{j \in \mathcal{J}} h^S_{j,l}(p^H) \right)^2$$

where $p^H \in \mathbb{R}^2_+$ denotes the m^2-price vectors, $h^D_{i,l}(p^H)$ denotes the net housing space demand of household i in location l given the m^2-price vector, and $h^S_{j,l}(p^H)$ denotes the net housing space supply provided by either an 'old' household or provided externally. The ω^l parameters govern how much weight we put on each location's deviation from its market clearing, which allows for asymmetric approximate market clearing conditions for different housing markets.

This algorithm would yield a price vector for each time period and location for the housing market, so it would close the gap between demand and supply as much as the discretization allows. If the temporary equilibrium price vectors exist, this mechanism with sufficiently fine price grids will find them. However, uniqueness is generally not guaranteed (Duranton and Puga, 2015), which also stands for this algorithm.

3.3.4 Model solution

The model can be solved by backward induction, yielding a unique optimum due to the strict concavity of the utility function in non-durable consumption. First, I solve the model for young and old households, generating value and policy functions. In this model, prices are in temporary 'pseudo-equilibrium', meaning that they are allowed to react to the exogenously given inelastic housing supply, interacting with the housing demand of young households following its life cycle optimization program. Therefore we must simulate the entire histories of all housing markets jointly over the life cycle of the agents, somewhat similar to how it is treated in the literature when exogenous price processes are separately estimated, such as Attanasio et al. (2012).

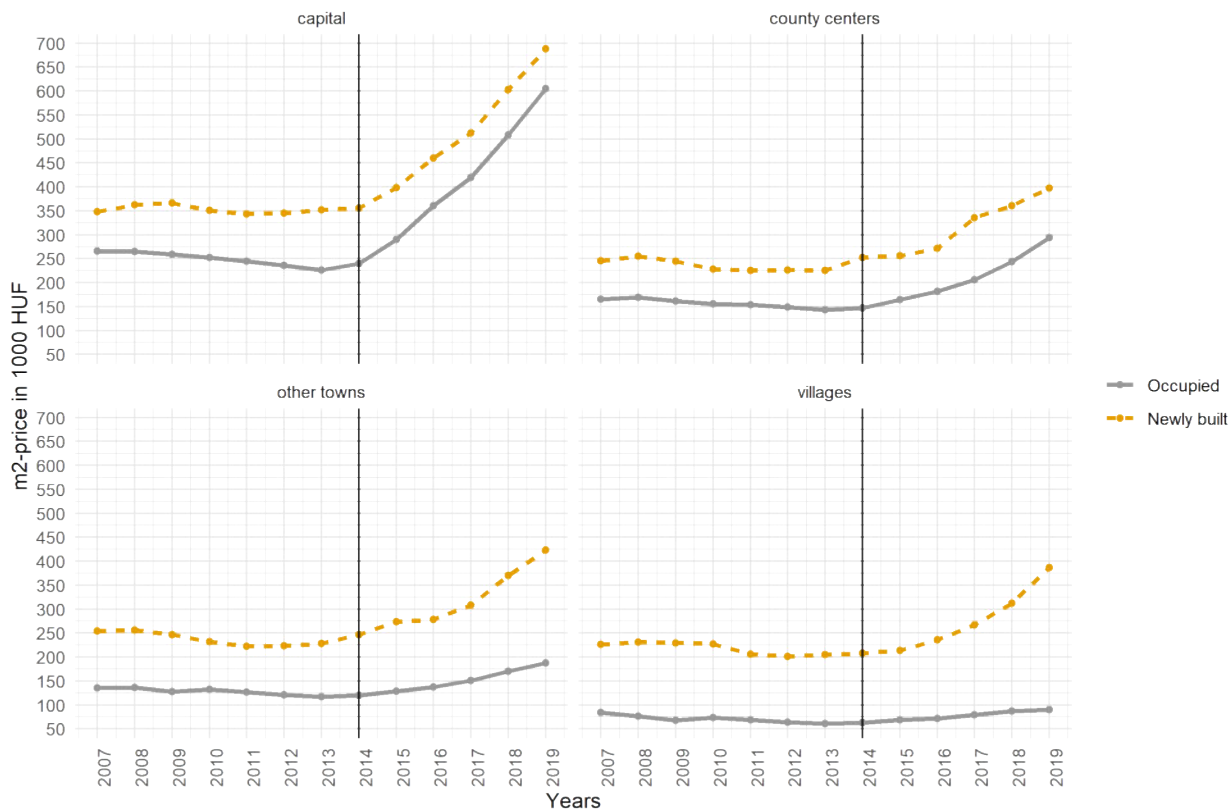
In the simulation, I include two cohorts of 150 'young' households (called Cohort-0 and Cohort-1) and 150 'old' households that act as price-sensitive supply. Cohort-0 households start at the beginning of the historical time of the simulation at age 25 and make optimal decisions according to the prescriptions of their policy functions in a price-taking manner. Cohort-1 households enter historic time two periods later, also at age 25, and make optimal decisions. However, they are lagging in their life cycle. This modeling choice allows for distinguishing between the effects of the policy on the first impacted cohort vs. future generations who already have to face the price impacts of the policy at an earlier stage in their life. It is the representation of the assessment of the Hungarian National Bank's comment on how the increasing house prices (if indeed the model generates that dynamic) might undermine the accessibility of housing for younger generations.

The m^2-price grids include six-six possibilities (Table 3.2) for each of the central and the rural locations, based on the historic price dynamics displayed in Figure 3.6, in the Hungarian context (HNB, 2019). We can see that for our estimation period, the central location (capital) prices for newly built dwellings are around 350,000 HUF, while the other locations are approximately 250,000 HUF. As I describe later, I use these m^2 prices to estimate the household demand parameters.

Table 3.2: Possible m^2-prices of houses

Central location	200	300	400	500	600	700
Rural location	50	100	150	200	250	300

Figure 3.6: Yearly average m^2-prices by type of municipality and age of the house, 2007-2019



Note: based on the house price estimates of the Hungarian Central Statistical Office.

The savings grids differ for each education group, which allows for representing the different credit limits given by their different earning powers. I could also set the savings grids finer for the lower educated group. Each grid contains 23 grid points, which are such that they reflect an annuity mortgage schedule for the relevant income groups, and are spaced evenly except around value 0, where it has two additional gridpoints. As I use the biannual model version instead of the annual one, the crudity of the savings grids plays an even less important role. As having mortgages could potentially result in significant indebtedness, it is vital to have a wide range of grid points for the savings accounts of households. Households are bound to choose a savings level from the grid point.

3.4 Estimation

I combine Hungarian survey and administrative datasets to estimate the model parameters using the 2004-2014 period when real estate prices did not experience significant turbulences.[25] I employ the simulated method of moments to estimate the preference parameters in the utility function. At the same time, I use the reduced-form Mincerian regressions for the wage parameters, which I treat simply as projection coefficients without any causal interpretation. I also estimate the probability distributions concerning the initial values of the state variables using the available information about households with cohabiting adults where females are 25-26 years old. The supply parameters (α) are then calibrated in a second step. Given the utility function parameter estimates, I simulate the model with endogenous prices and without the subsidy using different (α^0, α^1) supply parameters. Then I select the ones that produce the smallest squared deviation from the (350, 250) prices for the central and rural locations, respectively, that could be considered closest to the equilibrium prices for the 2004-2014 period. Figure 3.6 shows the average m^2-price levels for newly built and older houses as a reference point. We can see that the target prices for the second step are slightly higher for the central (Budapest, capital) location than their actual level. However, the price grids could not be extended further due to computational feasibility.

3.4.1 Data

The primary data source of the exercise is the Hungarian Household Budget Survey[26] , which is a yearly survey of private households, representative at the country level. The Hungarian Central Statistical Office collects detailed information on the population's consumption, income, housing, and several other demographic features, which is then used to calculate the product weights of the consumer price index. The information also contributes to national account calculations. The structure of the surveys has changed over the relevant period of 2004-2018, which required maintaining simple definitions of the variables that could map

[25]I am grateful for the help of the Databank at the Institute of Economics for providing access to the datasets used in this paper.

[26]In Hungarian: "Háztartási Költségvetési és Életkörülmény Adatfelvétel" abbreviated as HKÉF, the English description is available at
http://www.ksh.hu/apps/meta.objektum?p_lang=EN&p_menu_id=110&p_ot_id=100&p_obj_id=AEAA

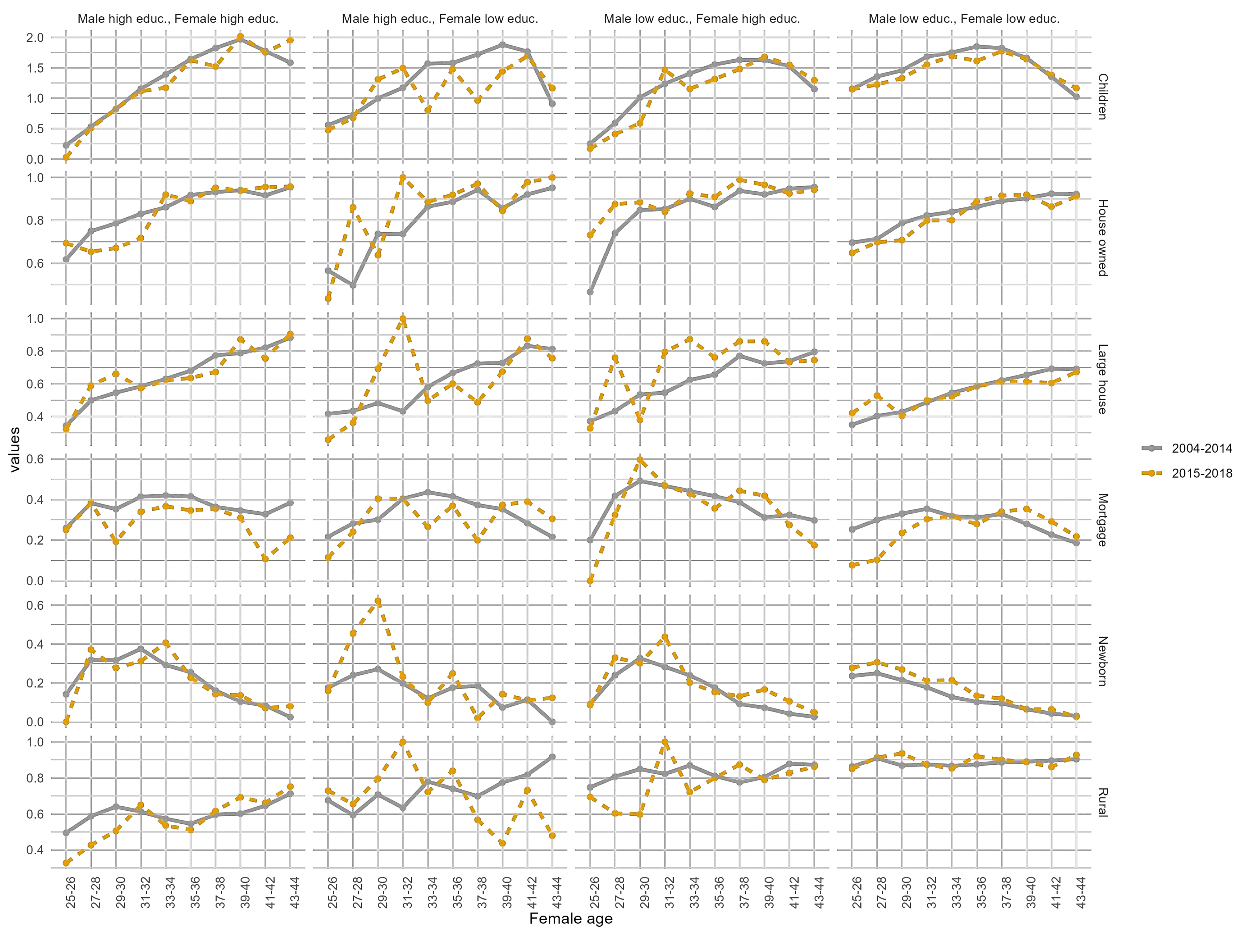
into the model. Since 2010, the survey has not included detailed nominal household income information. Therefore I used the Hungarian National Wage Survey data to impute expected net household wages conditional on the calendar year, age, gender, education, administrative region, and the type of municipality.

For each survey of 2004-2018, I kept in the sample those households which included two adults identified as male and female members of a cohabiting couple, not necessarily in marriage. Then I created the following household-level variables for the estimation of the model: consumption ratio (consumption as a fraction of net household wage income), house size category where houses with three or more rooms are categorized as large, location category based on residing in the capital (Budapest) as a proxy for the central location in the model, homeownership, mortgage, newborn children (children less than one) and the total number of children (children family status and less than 19 years of age). Finally, I pooled the pre-policy period of 2004-2014 and the post-policy period of 2015-2018 to increase the number of observations, as each survey contained only a small number of households after conditioning on education and female age.

A major deficiency of the dataset is that it was impossible to reconstruct the households' net savings position; hence, moments connected to this piece of information cannot be used in the estimation. Another potential problem is the selection to become a couple, which is not addressed by the model or the sampling. The latter issue appears both regarding calendar time with changing family structures over time and due to the actual effect of the policy on family formation. Addressing this selection problem could be a valuable addition to the present model.

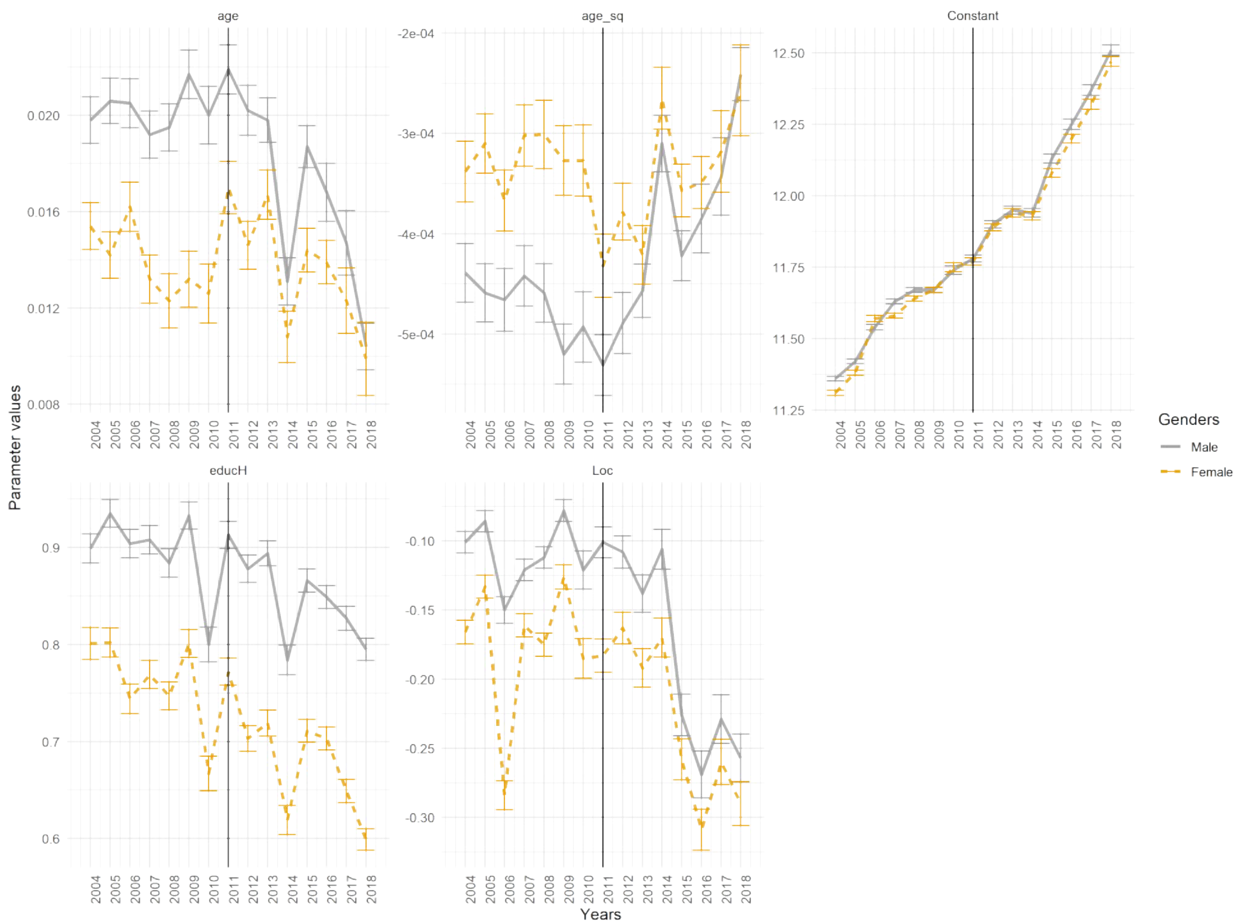
The following Figure 3.7 shows, by education categories and periods, the averages of the main choice and state variables as a function of the age of the female. We can immediately notice the greater variance in the time series for the shorter period due to the smaller time window. Otherwise, we can see that, on average, most variables, such as the total number of children, evolve over the life cycle more or less similar in both periods. The consumption ratio seems lower, which could be an unfortunate technical consequence of the wage imputation. The evolution of homeownership also appears to be similar, starting from a high level. However, we can notice a slight uptake in the ratio of larger houses in the latter period for younger households. Another thing to note is that at the older age of females, children start to leave the household resulting in a decrease in the number of children at the end, which is impossible in the model (hence I use the life cycle maximum of the average, instead of the terminal value).

Figure 3.7: Average values for the main choice and state variables, 2004-2014 vs. 2015-2018



For the parameters in the Mincerian wage regressions of the model, I used the data from the Hungarian Wage Survey, which contains demographic and salary information about the labor force of firms with more than five employees. I regressed gross monthly salaries on a constant, age, age squared, and dummies for higher education and rural location, separately for each gender and year. Figure 3.8 shows the point estimates with their 95% confidence intervals. We can see a substantial variance in the estimates of some parameters, such as age, throughout the periods. I use the 2011 estimates as parameter inputs for the model simulations.

Figure 3.8: Parameter estimates for the wage equation with 95% CI



Lastly, I estimated the distributions for the initial conditions of the state variables in the following way. I used the 2011 Census data of Hungary to estimate the joint probability distribution of households where the female's age is 25-26 by male education, female education, and the number of children. For the other state variables (home ownership, large house, location, mortgage), I used the univariate distributions estimated from the Household Budget Survey, as they would have been over-saturated for a multivariate estimation. The Census data held no information on mortgages.[27] Table 3.3 presents the household counts

[27]It is possible, however, to use the Census for a joint estimation regarding the other variables besides

and the initial moments for households by education group and period on which the imputed distributions are based.

Table 3.3: Averages of initial state variables by education and period, Household Budget Survey, 2004-2018

Female age	Male high educ.	Female high educ.	Period	Household count	Large house	Owned house	Rural	Mortgage	Children
25-26	0	0	2004-2014	661	0.35	0.70	0.86	0.25	1.15
25-26	0	0	2015-2018	110	0.42	0.65	0.85	0.08	1.15
25-26	0	1	2004-2014	119	0.37	0.47	0.75	0.20	0.25
25-26	0	1	2015-2018	6	0.33	0.73	0.69	0.00	0.17
25-26	1	0	2004-2014	56	0.42	0.57	0.68	0.22	0.56
25-26	1	0	2015-2018	15	0.26	0.44	0.73	0.11	0.48
25-26	1	1	2004-2014	138	0.34	0.62	0.49	0.26	0.23
25-26	1	1	2015-2018	16	0.32	0.69	0.33	0.25	0.03

Note: the author's calculation, based on the Hungarian Household Budget Survey, 2004-2018.

In the model, the unobserved heterogeneity originates from the latent desired number of children, for which the distribution is estimated in the demographic literature (Kapitány and Spéder, 2015). I display this distribution in Table 3.4.

Table 3.4: Estimated distribution of the desired number of children

k	0	1	2	3
$\mathbb{P}^{\nu}(k)$	0.02	0.12	0.65	0.21

Note: based on (Kapitány and Spéder, 2015).

3.4.2 Model fit

In this subsection, I examine the model fit diagnostics. I use the simulated method of moments (SMM) to estimate the parameters to match a set of moment conditions comprising

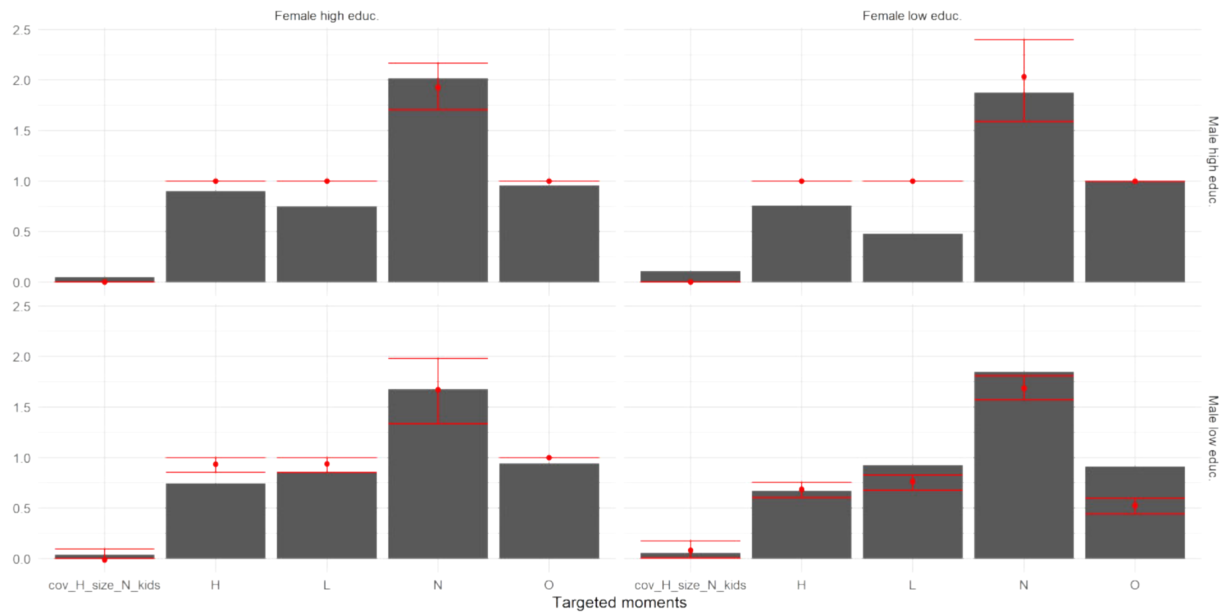
having a mortgage, although for some cells, the number of observations would be low even for the entire population. I decided against this option because according to the 2011 Census, around 80% of the households with 25-26-year-old females live in a house owned by them, which seems overestimated compared to the 46-70% ratios found in the Household Budget Survey. One possible explanation is that cohabiting couples might not appear as such in the census because, officially, many of them could have their permanent residence with their parents. So some cohabiting couples who live in a rented apartment might be accounted for as 'children' in the Census.

conditional expected values and covariances. The targets are derived from the Hungarian Household Budget Survey mentioned earlier. I use simple squared distance in the loss function without weights for the moments. I implemented the Cyclic Coordinate Search Algorithm (following Oswald, 2019) to estimate the demand parameters of the model. The algorithm fixes the previous guess of the parameter vector and then converges to a minimum changing only one of the parameters.

I used the following terminal period moments conditional on each combination of parental education of the first cohort: homeownership rate, the fraction of living in large houses, number of children, rural location rate, and covariance between house size and number of children. For the model fit, I used 200 Cohort-0 households without any other cohorts in 150 simulations, as I fixed the house prices at 350,000 HUF and 250,000 HUF for the central and the rural location, respectively, corresponding to the long-run average as we have seen in Figure 3.6 earlier. The simulations differ from each other as for each participating household and for each simulation, the unobserved heterogeneity of the ideal number of children, and the initial value for the other state variables (based on the joint distribution observed in the 2011 census) is redrawn, and then based on their optimal choices given each possible state of the world, I simulate forward the choice histories.

The following Figure 3.9 displays the in-sample fit of the model; the grey columns show the targeted data moments: covariance between terminal home size and the number of children, home size, location, number of children, and ownership status; while the red error bar shows the mean simulated moment with the 5th and 95th percentile, resulting from 150 simulations. We can see that the model has varying success across the moments. While fertility seems to be well-captured, other aspects often run into corner solutions and overshoot (house size or location for households of highly educated males), or in other instances, undershoot (low education home ownership) the target.

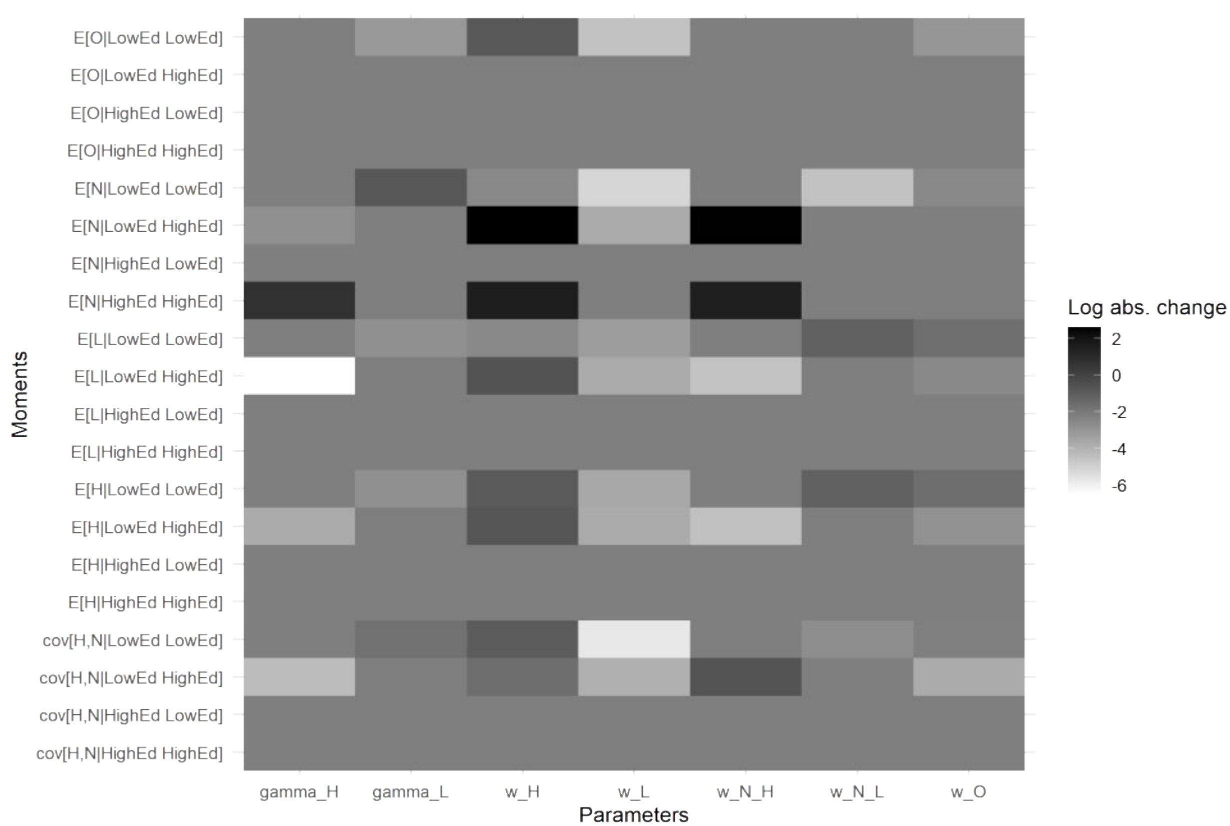
Figure 3.9: Fit of targeted moments



Note: model fit using 150 joint simulations of the 200 first-generation (Cohort-0) households, grey columns show the target, while the red error bars show the simulated moment with its 5th and 95th percentile.

Identification of the demand parameters can be summarized by the following heatmap of Figure 3.10, showing the intensity of change in moments responding to a marginal change in parameter values. The results are not surprising. We can see that the CRRA parameters (γ_{ed^F}) and the children parameters ($w^N_{\mathrm{ed}^F}$) react mostly to the expected number of children. House crowdedness parameter w^H to house size and the number of children, while location w^L and ownership w^O affect the moments more weakly but more uniformly.

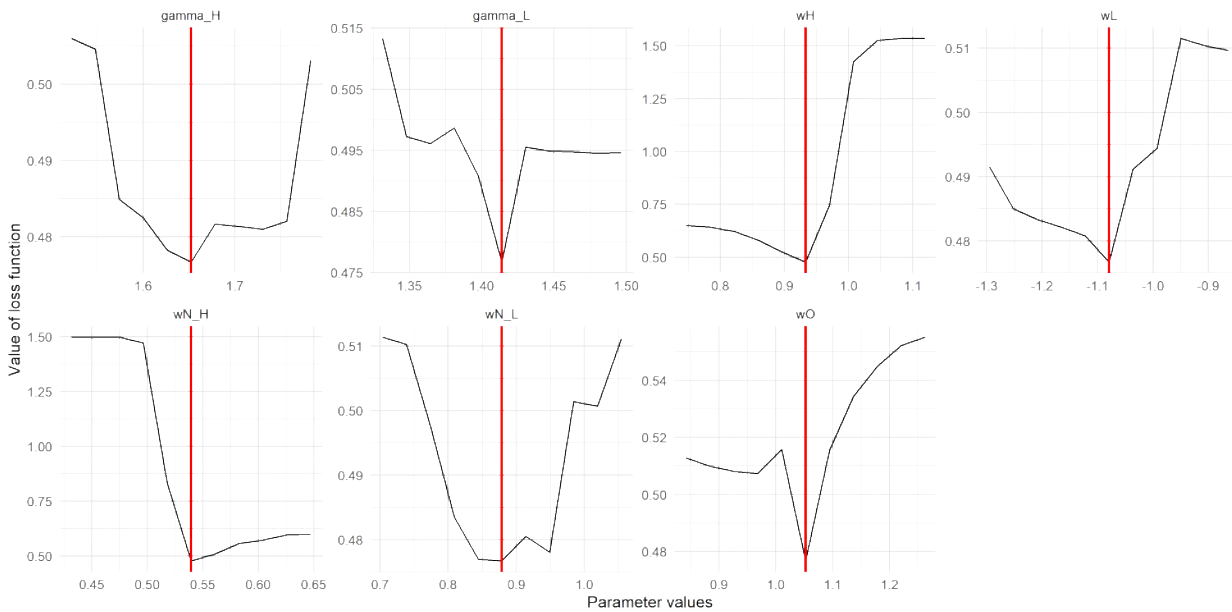
Figure 3.10: Identification of demand parameters



Note: 150 joint simulations of the 200 first-generation (Cohort-0) households.

Finally, Figure 3.11 shows the values of the loss function as a function of the parameters, indicating that, indeed, we have reached a local minimum with the combination of the parameter estimates in an adequately wide neighborhood of the minima.

Figure 3.11: Loss function by parameters



Note: 150 joint simulations of the 200 first-generation (Cohort-0) households.

3.4.3 Parameter estimates

Table 3.5 presents the parameter estimates and calibrations used for the model simulations. The parameters in the utility function are estimated by the simulated method of moments discussed earlier, while the wage regression parameters are estimated using reduced-form regressions. As discussed earlier, the supply parameters are chosen to minimize the distance from the price levels treated as equilibrium prices under the flexible price regime.

We can see that the γ parameters of the CRRA function are close to the value commonly found in the literature (as discussed by Attanasio et al. (2012), around 1.5). I also find that this parameter is higher for households of females with higher education. The w^N-parameters governing the disutility of being distant from the desired number of children is higher for females with lower education, consistent with the findings of Adda et al. (2017) that early career choices incorporate future fertility plans for females. It might manifest here that those who want more children select into career trajectories with lower wages, hence end up with lower education. The preference parameters regarding housing conditions indicate that given the model is correct, crowdedness (w^H), location (w^L), and homeownership (w^O) all play important roles in the life cycle behavior of households; however, the estimate for the location parameter is not statistically significant. Households are also found to appreciate homeownership and dislike crowdedness.

The regression parameters of the wage regressions show that wage profiles by age follow the usual quadratic shape on average for both genders. There is a substantial premium for higher education, and also for being located in the central area[28].

Table 3.5: Parameter estimates and calibrations

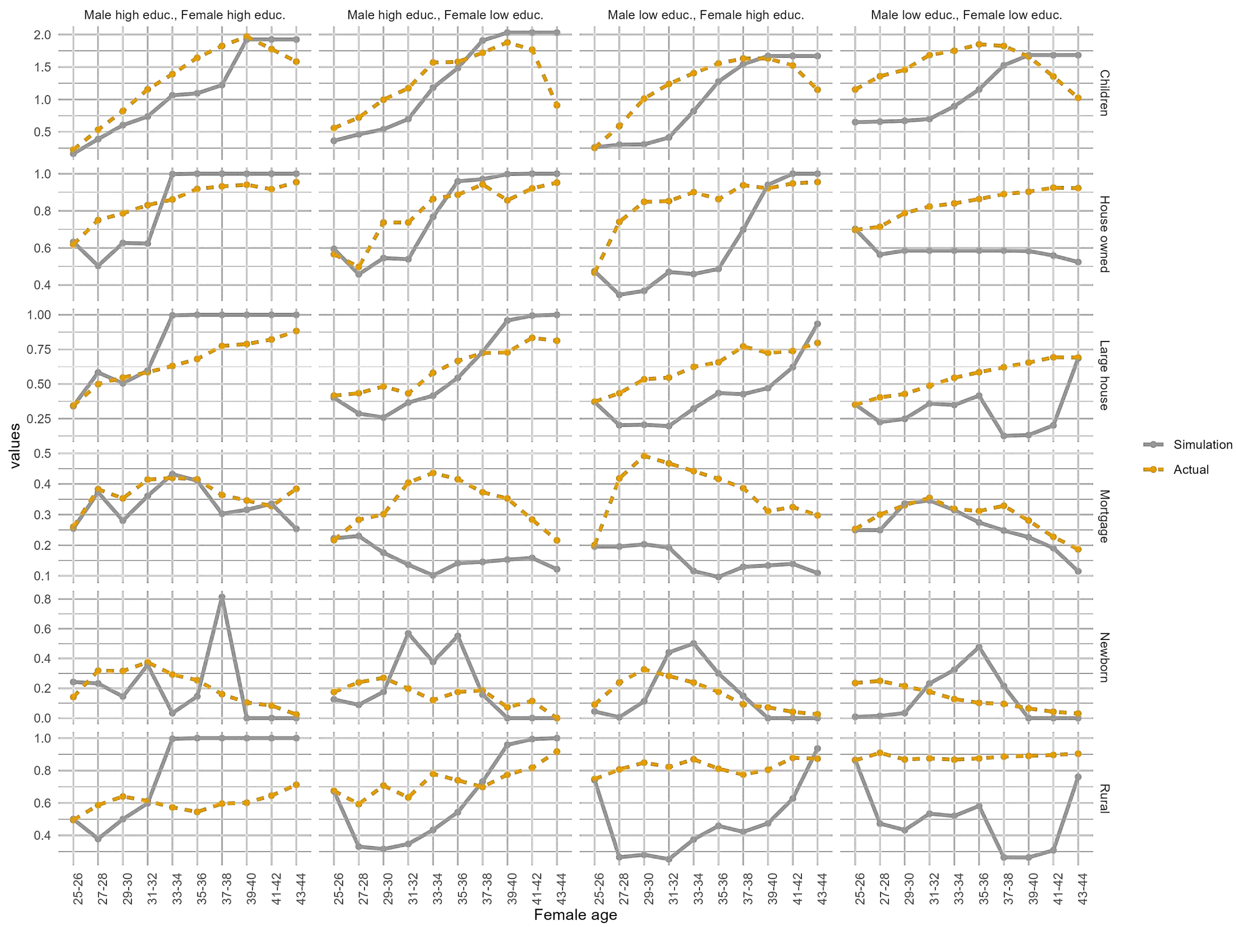
Parameter	Value, Estimate (S.E.)	Source
Utility function		
β	0.91	Set manually (biannual case)
γ_l	1.4140 (0.0599)	Estimated (SMM)
γ_h	1.6524 (0.0299)	Estimated (SMM)
w_l^N	0.8790 (0.1021)	Estimated (SMM)
w_h^N	0.5394 (0.0225)	Estimated (SMM)
w^H	0.9330 (0.0212)	Estimated (SMM)
w^L	-1.0790 (0.8978)	Estimated (SMM)
w^O	1.0522 (0.1242)	Estimated (SMM)
Budget constraint		
ir	0.10	HNB (2019) (biannual case)
δ	0.50	Set manually
ρ	0.05	Set manually
$\beta_0^{W,M}$	11.78 (0.0062)	Estimated (reduced form)
$\beta_1^{W,M}$	0.0219 (0.0005)	Estimated (reduced form)
$\beta_2^{W,M}$	-0.0005 (1.54e-05)	Estimated (reduced form)
$\beta_3^{W,M}$	0.913 (0.0071)	Estimated (reduced form)
$\beta_4^{W,M}$	-0.101 (0.0057)	Estimated (reduced form)
$\beta_0^{W,F}$	11.77 (0.0066)	Estimated (reduced form)
$\beta_1^{W,F}$	0.0170 (0.0006)	Estimated (reduced form)
$\beta_2^{W,F}$	-0.0004 (1.60e-05)	Estimated (reduced form)
$\beta_3^{W,F}$	0.772 (0.0071)	Estimated (reduced form)
$\beta_4^{W,F}$	-0.183 (0.0062)	Estimated (reduced form)
p_c	1,552.5 (in 1,000 HUF)	HCSO (2016)[29]
$W_{\min}$	960 (in 1,000 HUF)	Based on minimum wage in Hungary
Supply parameters		
α^0	0.0274	Calibrated to 2004-2014 house prices
α^1	0.0397	Calibrated to 2004-2014 house prices

[28]Note that we abstract away from commuting, which could potentially be very important, here, the estimates refer to the location of the firm sites.

3.4.4 Model validation

As we fit the model on the moments of the terminal values, we can use the previous periods to validate how well the model captures the dynamics of the key variables over the life cycle. Figure 3.12 displays how our main variables compare with their actual counterparts. We can see some weaknesses, most importantly concerning rural vs. central location choice, which could explain why the parameter estimate governing that aspect is not statistically significant. While in actuality, highly educated households choose to reside in the central location more than lower education households, in this model, we get the opposite. The central location provides higher earnings, so lower-education households are forced live in the central location comparatively more often. At the same time, the central location also gives disutility to households according to the parameter estimates. It results in households with higher education choosing to reside in the rural area without exception in the model.

Figure 3.12: Evolution of key variables over the life cycle, simulation vs. actual



Note: 150 joint simulations of the 200 first-generation (Cohort-0) households, and values based on the HKÉF data of 2004-2014.

3.5 Results of the policy simulations

Using the parameter estimates, I simulated the model for six different scenarios 150 times, with 150 Cohort-0, Cohort-1, and older households over ten biannual periods. House prices are endogenously evolving in all of them, possibly counteracting the intended policy effects. Then I study how fertility and housing variables evolve under these scenarios and show comparisons of household welfare in a partial equilibrium setting. The scenarios were the following:

1. No allowance (baseline)
2. Allowance available (Family Housing Allowance for larger houses)
3. Extra supply in housing without allowance (+200% in the α parameters)
4. Lower interest rate without allowance (3.5% instead of 5%)
5. Lower down payment requirement without allowance (40% instead of 50% required)
6. Lower interest rate with allowance (3.5% instead of 5% rate, with Family Housing Allowance)

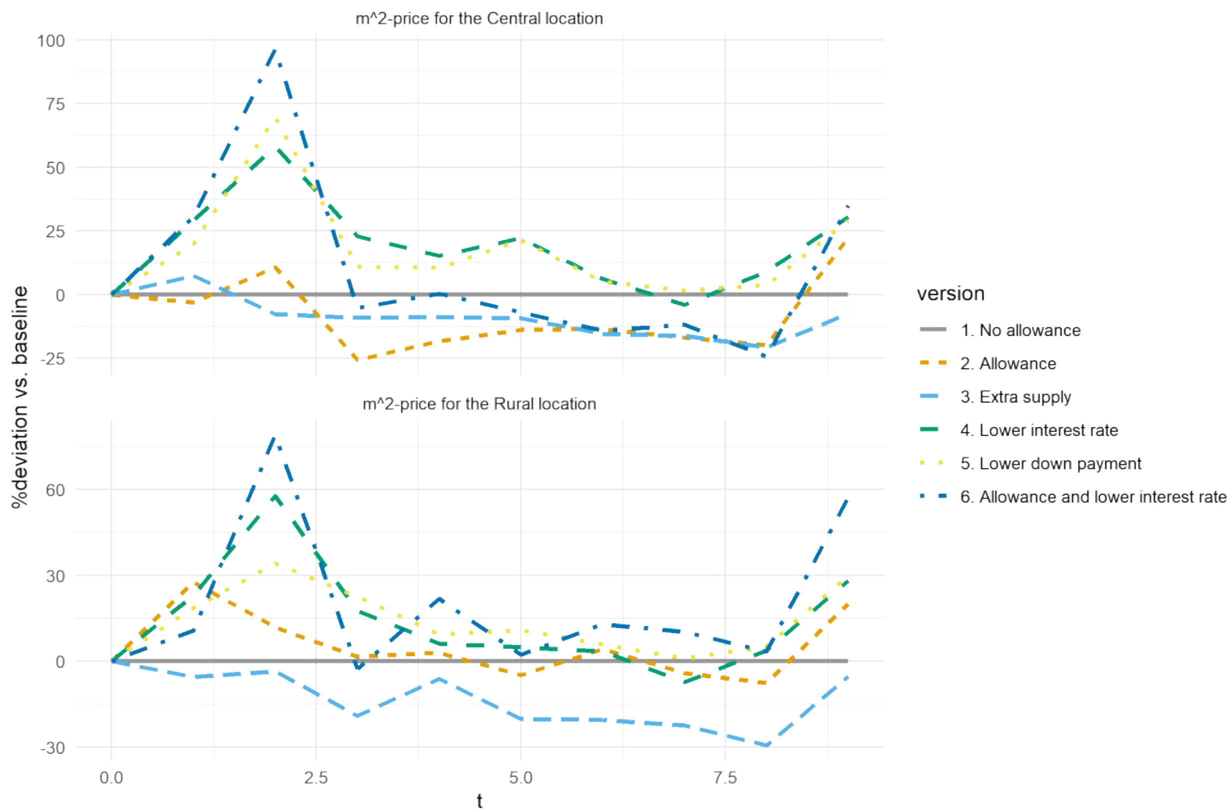
I included a combined monetary and fiscal policy scenario, 'Lower interest rate with allowance', to reflect that the Hungarian National Bank decreased interest rates substantially starting in 2013. So instead of the 5% interest rate assumed during the model estimation, which is more relevant for the pre-2014 period, I use 3.5% to represent the post-2014 period more accurately. Scenarios 'Allowance' and 'Lower interest rate' with and without the government allowance enable us to disentangle the potential effects of the policy on housing and fertility. The scenarios 'Extra supply' and 'Lower down payment' represent alternative housing market policies that directly target the housing market without fertility incentives, allowing us to examine the effects of simply alleviating housing-related constraints on the households.

3.5.1 Effect on the housing market

In this exercise, I assess the evolution of house prices compared to the baseline scenario. Note that even in the baseline scenario, I cannot reproduce constant prices at the target values described earlier with only two parameters (α); only the average over the 10-period (20-year) frame will be close to those values. Otherwise, there are fluctuations driven by the young households' life cycle (in Appendix, Figure 3A8). However, life cycle fluctuations are similar for all scenarios, so the differences can be interpreted as the effect of the policies.

First, let us consider the evolution of house prices under the proposed scenarios in Figure 3.13, showing %-deviation compared to the baseline. We can immediately observe that lower interest rates and down payment ratios will result in higher house prices in both the central and the rural locations (50-80% for the central and 30-60% for the rural) for the first four years (two periods). More importantly, we can see that besides an around 30% short-term increase in rural area house prices, the government allowance alone does not induce such significant changes close to what we would observe in the actual data. However, the allowance combined with a lower interest rate does produce around 100% increase in house prices for the central and 75% for the rural areas, which are somewhat close to the actual data points. However, after three to four periods, house prices will not differ much from the baseline scenario values. So the deviation is only temporary in this model (possibly due to only being able to simulate two cohorts, implicitly enforcing this adjustment mechanism). On the contrary, with the extra supply scenario, we can naturally drive down the prices even in the long run. Still, even an around 200% permanent increase in new housing decreases house prices by only 20-30%.

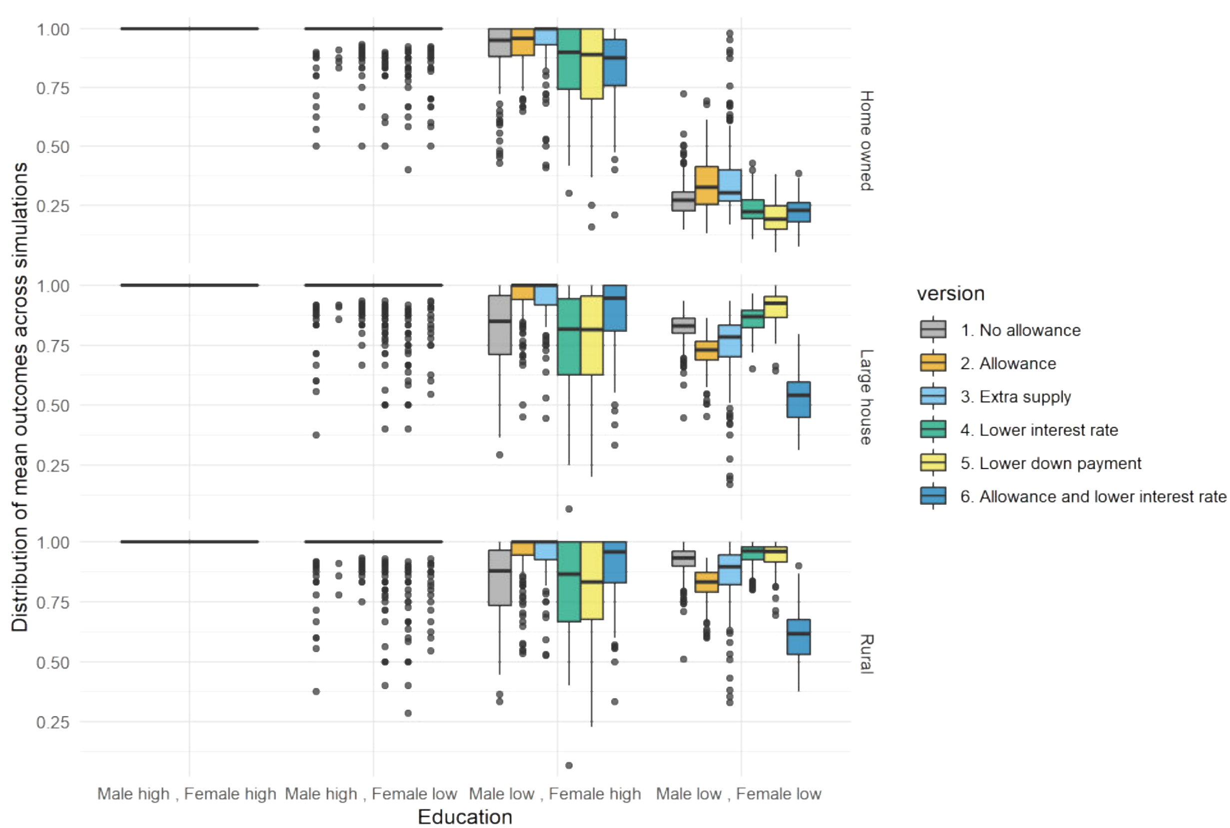
Figure 3.13: Evolution of house prices compared to the baseline scenario



Note: using 150 simulations of 150 Cohort-0, Cohort-1 and 'old' households. One period on the horizontal axis represents two years.

The housing-related terminal outcomes of households are displayed in Figure 3.14 (the evolution of the variables can be found in the Appendix, Figure 3A5). It is easy to see that there is no effect for households with high male education, as we run into a corner solution for most of these households. However, it suggests that housing outcomes in this model are related to household income primarily driven by male income. The results are mixed for households with low male education; I will focus on the most numerous low-male and low-female education groups. Regarding ownership, we can see that the homeownership ratio does not change much due to the scenarios. While with allowance and extra supply, it is slightly higher, in the other scenarios, it is somewhat lower. The fraction living in a large house does change substantially due to the policies. The Housing Allowance counterintuitively seems to decrease the fraction of these poorer households living in a larger house, especially if interest rates are lower, which induces the greatest elevation in house prices. Living in a large house is higher than the baseline if down payment requirements are lower, while they are around the same for the other scenarios. The fraction living in the rural location also decreases due to the allowance, even more under lower interest rates. This finding reflects the elevated house prices and that households can afford only smaller homes when moving into the central location to receive higher salaries.

Figure 3.14: Effect of scenarios on end-of-life-cycle housing outcomes



Note: using 150 simulations of 150 Cohort-0, Cohort-1, and 'old' households, based on the results of Cohort-0, the results for Cohort-0.

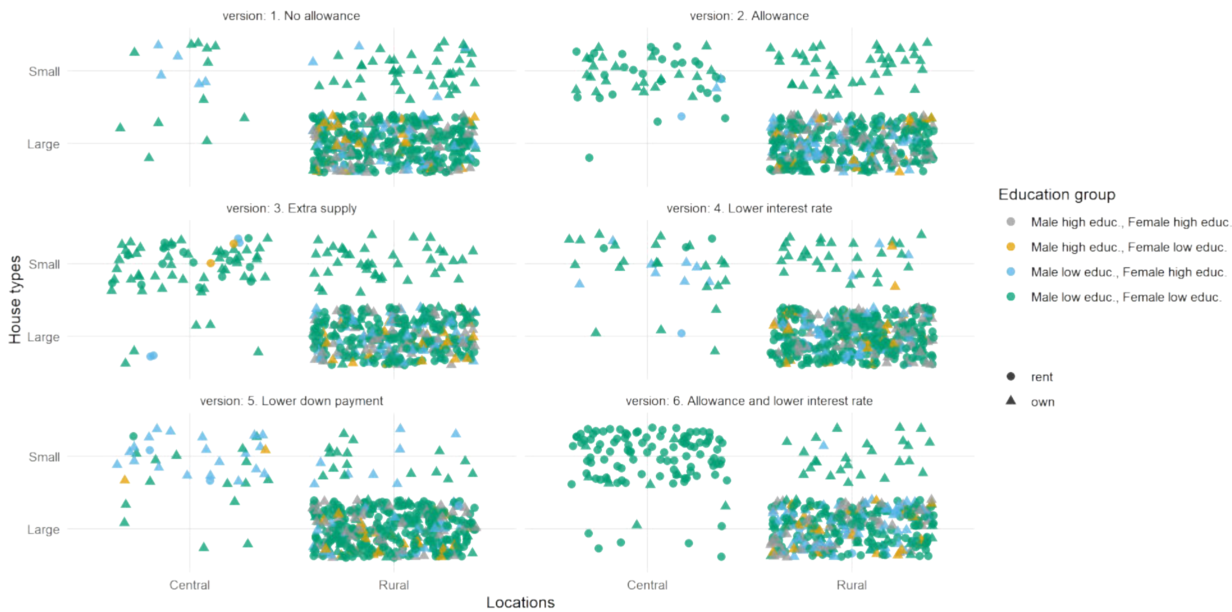
We can examine the sorting into house sizes and locations of households under our scenarios by education groups. Table 3.6 shows how households are distributed on average within their education group across sizes and locations, while Figure 3.15 illustrates the distribution of households with information on ownership status as well. We can again see that high-male education households live dominantly in large rural houses under any scenario, while the distribution of low-male education households varies by the circumstances. Most importantly, under the 'Allowance with lower interest rate' scenario, many more of the low male and low female education households live in small central area houses than under the other scenarios, which spot is the least desired according to the estimated preference parameters. At the same time, our figure shows that these households rent and do not own apartments in the center. This suggests the Allowance policy combined with low interest rates harms the poorest households' housing conditions, while neither would have such an impact.

Table 3.6: Household sorting into house size and location, by education

Size	Location	Education	Baseline	Allowance	Extra supply	Lower interest rate	Lower down payment	Allowance and lower interest rate
Large	Rural	Male high educ., Female high educ.	1.00	1.00	1.00	1.00	1.00	1.00
Large	Central	Male high educ., Female low educ.	0.00	0.00	0.00	0.01	0.01	0.01
Large	Rural		0.97	1.00	0.98	0.94	0.94	0.97
Small	Central		0.03	0.00	0.01	0.04	0.04	0.02
Small	Rural		0.00	0.00	0.01	0.01	0.01	0.00
Large	Central	Male low educ., Female high educ.	0.01	0.02	0.01	0.01	0.02	0.03
Large	Rural		0.81	0.92	0.93	0.76	0.75	0.86
Small	Central		0.16	0.04	0.04	0.17	0.17	0.08
Small	Rural		0.03	0.02	0.01	0.06	0.06	0.04
Large	Central	Male low educ., Female low educ.	0.01	0.01	0.01	0.01	0.01	0.01
Large	Rural		0.80	0.71	0.73	0.85	0.89	0.52
Small	Central		0.07	0.17	0.13	0.05	0.05	0.38
Small	Rural		0.12	0.11	0.13	0.10	0.05	0.09

Note: the author's calculation based on 150 simulations of the model Cohort-0.

Figure 3.15: Household sorting into house size, location by ownership status and education

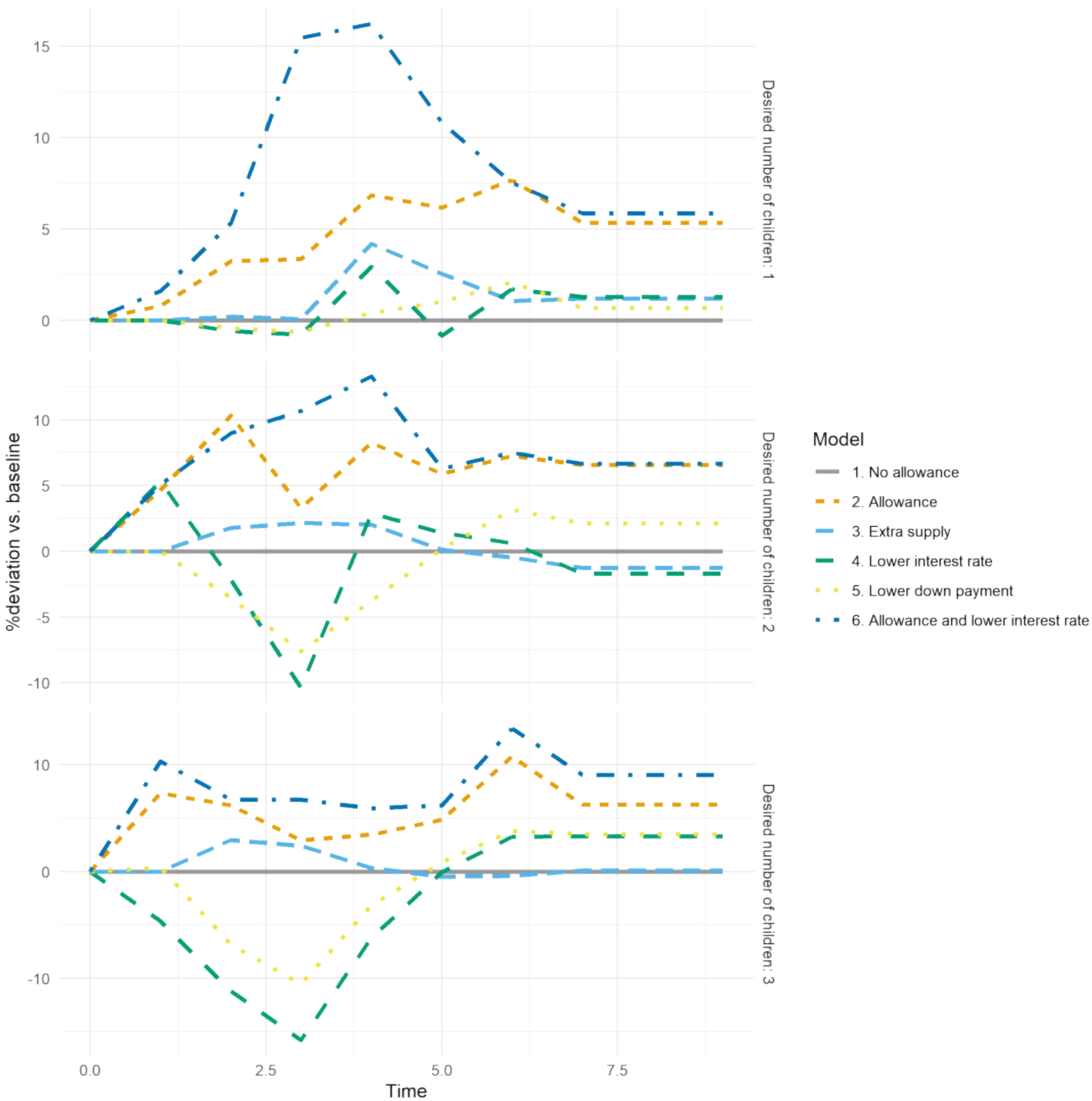


Note: the author's calculations based on three simulations of 150 Cohort-0, Cohort-1, and 'old' households, to allow for visibility of the results.

3.5.2 Effect on fertility and welfare

One point of focus in this analysis is to evaluate the Family Housing Allowance policy's potential long-run effects on the completed fertility outcomes of households vs. timing effects. Figure 3.16 shows the %-deviation compared to the baseline scenario, in the number of children for Cohort-0's life cycle, by the number of desired children. In this model setting, the Allowance results in an approximately 5-10% increase in completed fertility by the end of the life cycle. The policy also alters timing, as we can see that children are born earlier in the life cycle under the Allowance policy (hence the larger difference at the earlier periods). Other policies also affect fertility but to a lower degree. Lower down payment seems to cause an around 2-5% increase in completed fertility. A lower interest rate increases fertility for households that want one or three children but decreases fertility for those that want two children. The extra supply scenario causes a slight rise in fertility for households that want one child but a small decrease for those that want two, with no change for those that would like three.

Figure 3.16: Evolution of the total number of children compared to the baseline scenario, by the desired number of children



Note: using 150 simulations of 150 Cohort-0, Cohort-1 and 'old' households, the results of Cohort-0. One period on the horizontal axis equals two years.

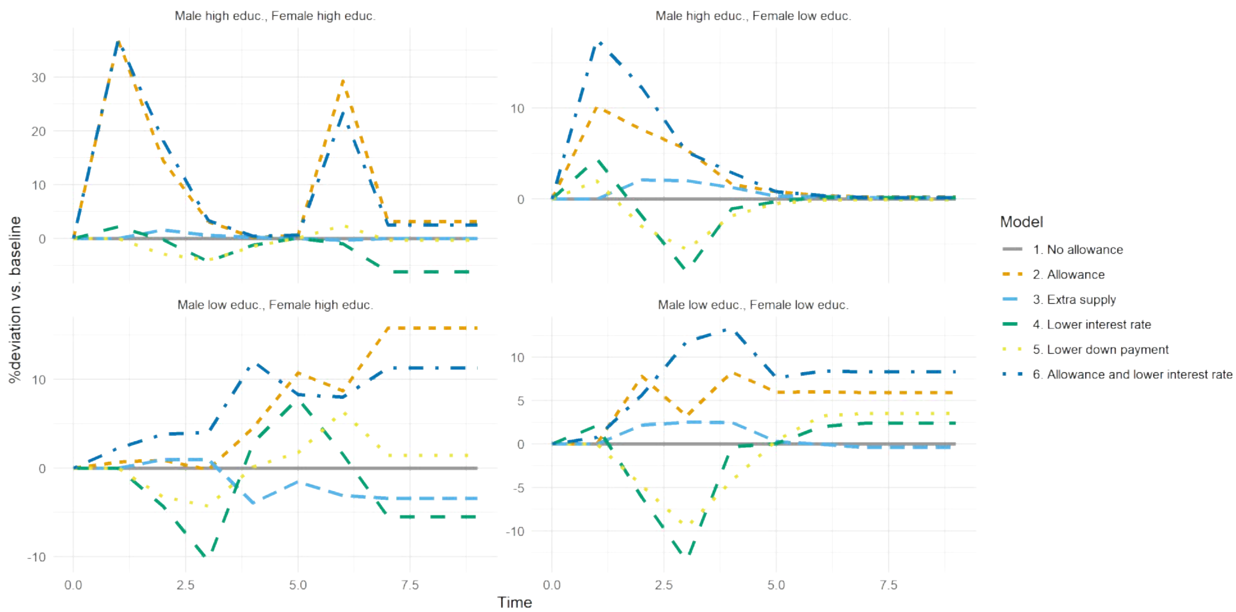
We can also examine which education groups change their fertility due to different policies — Table 3.7 and Figure 3.17 display these results. Households with high education increase their completed fertility by 0.05 (around 2.5%) child per family as a response to either allowance scenario. In contrast, their fertility decreases substantially (-0.12 child per family) under the lower interest rate scenario. Households with lower education react somewhat differently: they increase their completed fertility more as a reaction to the allowance policies (0.11 and 0.15 child per family, or around 6-8% respectively). However, their fertility is unchanged or slightly changed under all other policies. Changes in the timing of births depend on household education as well. Households with high male education bring their births forward in time due to the Allowance, but as we saw, it does not imply much in completed fertility. In comparison, lower education households increase their completed fertility more with the allowance, while their reaction in timing is lower.

Table 3.7: Average completed fertility by education

Education group	Baseline	Allowance	Extra supply	Lower interest rate	Lower down payment	Allowance and lower interest rate
Male high educ., Female high educ.	1.9282	1.9886	1.9279	1.8090	1.9215	1.9768
	(0.0126)	(0.0123)	(0.0126)	(0.0137)	(0.0128)	(0.0126)
Male high educ., Female low educ.	2.0596	2.0636	2.0620	2.0620	2.0580	2.0620
	(0.0172)	(0.0173)	(0.0173)	(0.0172)	(0.0173)	(0.0174)
Male low educ., Female high educ.	1.5697	1.8176	1.5160	1.4834	1.5920	1.7471
	(0.0154)	(0.0153)	(0.0155)	(0.0155)	(0.0154)	(0.0156)
Male low educ., Female low educ.	1.8219	1.9301	1.8146	1.8663	1.8860	1.9739
	(0.0061)	(0.0055)	(0.0062)	(0.0061)	(0.0060)	(0.0054)

Note: using 150 simulations of 150 Cohort-0, Cohort-1 and 'old' households, the results of Cohort-0. One period on the horizontal axis equals two years.

Figure 3.17: Evolution of the total number of children compared to the baseline scenario, by education



Note: using 150 simulations of 150 Cohort-0, Cohort-1 and 'old' households, the results of Cohort-0. One period on the horizontal axis equals two years.

We can also examine the sorting of families into different housing conditions based on their fertility outcomes: Table 3.8 shows the distribution within the final number of children across house sizes and locations. In contrast, Figure 3.18 displays the sorting based on the households' level of education.[30] We can see that policies do not affect the housing condition of families without or with three children. While families without children settle in large or small rural houses, families with three live in large rural houses almost exclusively. The policies, however, change where families with one or two children end up living. Compared to the baseline, in the scenarios with the Family Housing Allowance, more of them reside in small central location houses (30 percentage point increase), and most of those would otherwise have lived in large rural dwellings in the baseline scenario. It suggests that according to this model, those families who do not end up with three children are somewhat worse off regarding their housing conditions. The figure also indicates that these families that end up

[30]The evolution of the variables can be found in the Appendix Figure 3A6

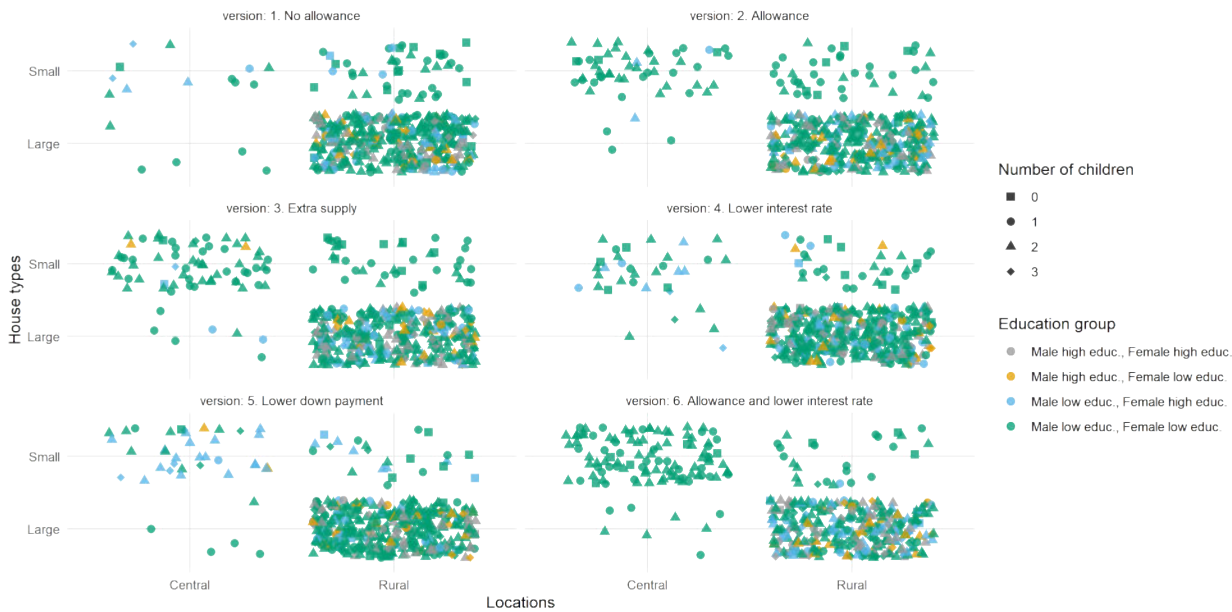
in the central location are households with low education.

Table 3.8: Household sorting into house sizes and location, by number of children

Size	Location	Children	Baseline	Allowance	Extra supply	Lower interest rate	Lower down payment	Allowance and lower interest rate
Large	Central	0	0.02	0.01	0.02	0.01	0.02	0.01
Large	Rural		0.70	0.65	0.73	0.73	0.70	0.69
Small	Central		0.05	0.05	0.05	0.04	0.04	0.04
Small	Rural		0.23	0.29	0.20	0.22	0.25	0.26
Large	Central	1	0.02	0.02	0.02	0.01	0.03	0.03
Large	Rural		0.72	0.47	0.66	0.80	0.83	0.40
Small	Central		0.09	0.29	0.14	0.05	0.06	0.38
Small	Rural		0.17	0.22	0.18	0.14	0.08	0.18
Large	Central	2	0.01	0.00	0.00	0.01	0.01	0.01
Large	Rural		0.89	0.85	0.86	0.89	0.92	0.64
Small	Central		0.06	0.10	0.08	0.06	0.05	0.32
Small	Rural		0.05	0.04	0.05	0.05	0.02	0.04
Large	Central	3	0.00	0.00	0.00	0.01	0.00	0.00
Large	Rural		0.91	0.99	0.90	0.90	0.90	0.97
Small	Central		0.08	0.00	0.08	0.08	0.08	0.02
Small	Rural		0.01	0.00	0.01	0.02	0.02	0.01

Note: the author's calculation based on 150 simulations of the model to allow for visibility of the results.

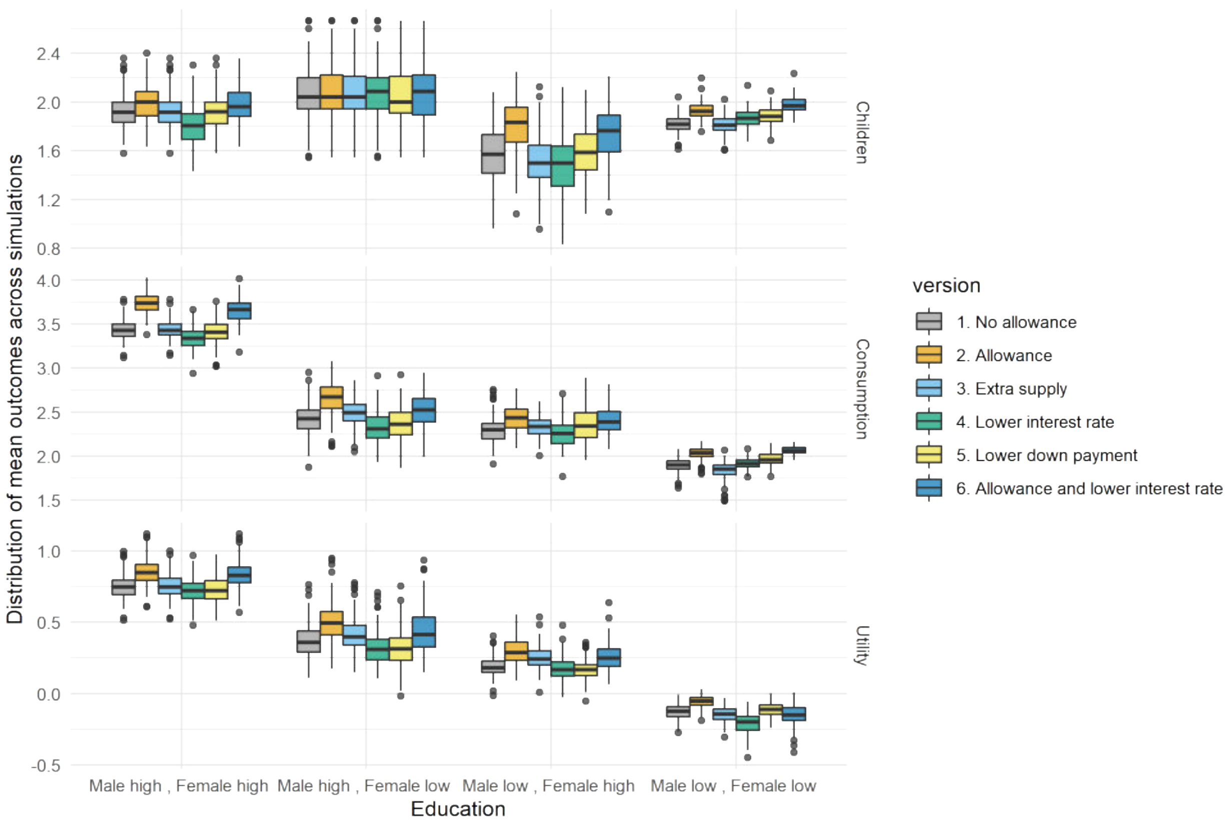
Figure 3.18: Household sorting into house size, location by the number of children and education, under three scenarios



Note: the author's calculations based on three simulations of 150 Cohort-0, Cohort-1, and 'old' households, to allow for visibility of the results.

Finally, we can also compare the scenarios directly by calculating the average discounted utilities realized over the life cycles of the households. Figure 3.19 shows these utilities, along with showing average lifetime consumption and then again the completed fertility, now as the distribution of averages over simulations. We can also see that for households with high education for at least one parent, the Allowance improved average utilities. In contrast, all other policies leave it more or less unaffected. However, lower interest rates seem to decrease utility substantially for households where both males and females have lower education. So much so that even with the Allowance, the net effect is not positive (while Allowance without lower interest rates seems to have a positive utility effect). I have already discussed that housing conditions of low-education households worsen under lower interest rates, which is then compensated by increased consumption and the number of children to have a net neutral effect. The figure shows that for all groups, non-surprisingly, the available government subsidy will result in higher consumption, while the other policies leave it unaffected.

Figure 3.19: Effect of scenarios on non-housing outcomes



Note: the author's calculations based on 150 simulations of 150 Cohort-0, Cohort-1 and 'old' households, the results for Cohort-0.

3.6 Discussion and conclusion

In this paper, I constructed a life cycle model of household behavior focusing on fertility, location, and housing choices, with unobserved heterogeneity in the preferences for the desired number of children and endogenously evolving house prices. Using the period of 2004-2014, I estimated the parameters characterizing the preferences of the households and then used them to analyze the effects of different policy scenarios on the housing conditions and completed fertility of households. I concentrated on the potential impact of the Hungarian Government's Family Housing Allowance program, which has been running since 2015.

Without additional changes, the Allowance program itself cannot explain the house price increases observed in Hungary after its introduction. However, combining it with lower interest rates (as these policies indeed coexisted in the relevant time frame) induces a house price evolution relative to the baseline scenario comparable to the one observed after 2015 in Hungary. According to this model, the combination of the fiscal and monetary policy jointly could be responsible for the house price increases, possibly negating some of the welfare effects of the policy.

The model suggests that poorer households (represented by both parents having attained a non-tertiary education level) were affected ambivalently by the policy, mainly driven by the counteracting negative impact of higher house prices. While consumption and the number of children seem to respond positively to the policy incentives, their housing conditions worsen in the long run. It occurs because many have to move to the less desirable central location and rent small apartments instead of living in self-owned, larger rural houses. Comparatively, households with higher education seem to benefit from the policy, seemingly without significant drawbacks. These findings do not seem to contradict the recent brief assessment of HNB (2021), which suggests that the policy helped families with three children but did not raise housing availability for other family types.

As future improvements and extensions of this model, I will attempt to resolve the problems of model fit regarding rural and central locations across households with different educational backgrounds. Furthermore, a drawback of the approach I chose is that when allowed to fluctuate endogenously, house prices, even in the baseline scenario, are far from stable, which would further strengthen the conclusions of this study.

References

Adda, J., Dustmann, C., and Stevens, K. (2017). The Career Costs of Children. *Journal of Political Economy*, 125(2):293–337.

Attanasio, O. P., Bottazzi, R., Low, H. W., Nesheim, L., and Wakefield, M. (2012). Modelling the demand for housing over the life cycle. *Review of Economic Dynamics*, 15(1):1–18.

Banai, A., Danila, P., Fábián, G., and Nagy, T. (2019). How CSOK changed the domestic housing market? (In Hungarian: Hogyan alakította át a CSOK a hazai lakás-és hitelpiacot?). Technical report, Hungarian National Bank.

Barakova, I., Calem, P. S., and Wachter, S. M. (2014). Borrowing constraints during the housing bubble. *Journal of Housing Economics*, 24:4–20.

Becker, G. S. (1991). *A Treatise on the Family*. Harvard University Press, Cambridge, Mass, enl. ed edition.

Billari, F. and Kohler, H.-P. (2004). Patterns of low and lowest-low fertility in Europe. *Population Studies*, 58(2):161–176.

Brueckner, J. K. and Zenou, Y. (1999). Why is central Paris rich and downtown Detroit poor? An amenity-based theory. *European Economic Review*, page 17.

Chambers, M., Garriga, C., and Schlagenhauf, D. E. (2009). The loan structure and housing tenure decisions in an equilibrium model of mortgage choice. *Review of Economic Dynamics*, 12:444–468.

Chan, S., Haughwout, A., and Tracy, J. (2015). How Mortgage Finance Affects the Urban Landscape. In *Handbook of Regional and Urban Economics*, volume 5, pages 987–1045. Elsevier.

Chetty, R., Hendren, N., and Katz, L. F. (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.

Chudnovskaya, M. (2019). Housing context and childbearing in Sweden: a cohort study. *Housing Studies*, 34(3):469–488.

Cloyne, J., Huber, K., Ilzetzki, E., and Kleven, H. (2019). The Effect of House Prices on Household Borrowing: A New Approach. *American Economic Review*, 109(6):2104–2136.

Cocco, J. F. (2005). Portfolio Choice in the Presence of Housing. *Review of Financial Studies*, 18(2):535–567.

Combes, P.-P., Duranton, G., and Gobillon, L. (2019). The Costs of Agglomeration: House and Land Prices in French Cities. *The Review of Economic Studies*, 86(4):1556–1589.

Davis, M. A., Fisher, J. D. M., and Whited, T. M. (2014). Macroeconomic Implications of Agglomeration. *Econometrica*, 82(2):731–764.

Davis, M. A. and Ortalo-Magné, F. (2011). Household expenditures, wages, rents. *Review of Economic Dynamics*, 14(2):248–261.

Davis, M. A. and Van Nieuwerburgh, S. (2015). Housing, Finance, and the Macroeconomy. In *Handbook of Regional and Urban Economics*, volume 5, pages 753–811. Elsevier.

Day, C. and Guest, R. (2016). Fertility and female wages: A new link via house prices. *Economic Modelling*, 53:121–132.

Dettling, L. J. and Kearney, M. S. (2014). House prices and birth rates: The impact of the real estate market on the decision to have a baby. *Journal of Public Economics*, 110:82–100.

Di Maggio, M., Kermani, A., Keys, B. J., Piskorski, T., Ramcharan, R., Seru, A., and Yao, V. (2017). Interest Rate Pass-Through: Mortgage Rates, Household Consumption, and Voluntary Deleveraging. *American Economic Review*, 107(11):3550–3588.

Duranton, G. and Puga, D. (2015). Urban Land Use. In *Handbook of Regional and Urban Economics*, volume 5, pages 467–560. Elsevier.

Eckstein, Z., Keane, M., and Lifshitz, O. (2019). Career and Family Decisions: Cohorts Born 1935-1975. *Econometrica*, 87(1):217–253.

Ermisch, J. (1999). Prices, Parents, and Young People's Household Formation. *Journal of Urban Economics*, 45(1):47–71.

Ermisch, J. and Steele, F. (2016). Fertility expectations and residential mobility in Britain. *Demographic Research*, 35:1561–1584.

Esping-Andersen, G. and Billari, F. C. (2015). Re-theorizing Family Demographics. *Population and Development Review*, 41(1):1–31.

Famiglietti, M., Garriga, C., and Hedlund, A. (2020). The Geography of Housing Market Liquidity During the Great Recession. *Federal Reserve Bank of St. Louis Review*, 102(1).

Fischer, M. and Khorunzhina, N. (2019). Housing Decision with Divorce Risk. *International Economic Review*, 60(3):1263–1290.

Fisher, J. D. M. and Gervais, M. (2011). Why Has Home Ownership Fallen Among the Young. *International Economic Review*, 52(3):883–912.

Flynn, L. (2017). Delayed and depressed: from expensive housing to smaller families. *International Journal of Housing Policy*, 17(3):374–395.

Flynn, L. B. (2020). The young and the restless: housing access in the critical years. *West European Politics*, 43(2):321–343.

Gambaro, L., Joshi, H., and Lupton, R. (2017). Moving to a better place? Residential mobility among families with young children in the Millennium Cohort Study. *Population, Space and Place*, 23(8):e2072.

Glaeser, E. L., Gyourko, J., and Saiz, A. (2008). Housing supply and housing bubbles. *Journal of Urban Economics*, 64(2):198–217.

Green, R. K. and Lee, H. (2016). Age, demographics, and the demand for housing, revisited. *Regional Science and Urban Economics*, 61:86–98.

HCSO (2016). A létminimum-számítás megújítása [The Reform of the Calculation of the Minimum Subsistence Figure]. Technical report.

Hedlund, A. (2018). Credit constraints, house prices, and the impact of life cycle dynamics. *Economics Letters*, 171:202–207.

HNB (2019). Lakáspiaci jelentés 2019 november [Real Estate Report of the Hungarian National Bank, November 2019]. Technical report, Hungarian National Bank.

HNB (2021). Lakáspiaci jelentés 2021 május [Real Estate Report of the Hungarian National Bank, May 2021]. Technical report, Hungarian National Bank.

Holly, S., Pesaran, M. H., and Yamagata, T. (2010). A spatio-temporal model of house prices in the USA. *Journal of Econometrics*, 158(1):160–173.

Hungarian Government (2016a). 16/2016. (ii. 10.) korm. rendelet az új lakások építéséhez, vásárlásához kapcsolódó lakáscélú támogatásról [16/2016 (ii.10.) government decree about the housing subsidy concerning the construction and purchase of new dwellings]. https://net.jogtar.hu/jogszabaly?docid=A1600016.KOR.

Hungarian Government (2016b). 17/2016. (ii. 10.) korm. rendelet a használt lakás vásárlásához, bővítéséhez igényelhető családi otthonteremtési kedvezményről [17/2016 (ii.10.) government decree about the family housing allowance concerning the purchase and extension of owner-occupied dwellings]. https://net.jogtar.hu/jogszabaly?docid=A1600017.KOR.

Jordà, , Knoll, K., Kuvshinov, D., Schularick, M., and Taylor, A. M. (2019). The Rate of Return on Everything, 1870–2015. *The Quarterly Journal of Economics*, pages 1225–1298.

Kapitány, B. (2016). Lakáshelyzet és gyermekvállalás: fontos, de nem a legfontosabb. A csok termékenységi hatásairól [Housing situation and fertility: Important, but not the most important. The impact of CSOK on fertility]. *Korfa*, 16(1):1–4.

Kapitány, B. and Spéder, Z. (2015). Fertility. In Monostori, J., Őri, P., and Spéder, Z., editors, *Demographic Portrait of Hungary 2015.*, pages 41–55. HDRI, Budapest.

Kapitány, B. and Spéder, Z. (2020). Fertility. In *Demographic Portrait of Hungary 2018*, pages 49–66.

Kilgarriff, P., Charlton, M., Foley, R., and O'Donoghue, C. (2019). The impact of housing consumption value on the spatial distribution of welfare. *Journal of Housing Economics*, 43:118–130.

Kulu, H. and Steele, F. (2013). Interrelationships Between Childbearing and Housing Transitions in the Family Life Course. *Demography*, 50(5):1687–1714.

Kulu, H. and Vikat, A. (2007). Fertility differences by housing type: The effect of housing conditions or of selective moves? *Demographic Research*, 17:775–802.

Li, W., Liu, H., Yang, F., and Yao, R. (2016). Housing Over Time and Over the Life Cycle: a Structural Estimation. *International Economic Review*, 57(4):1237–1260.

Li, W. and Yao, R. (2007). The Life-Cycle Effects of House Price Changes. *Journal of Money, Credit and Banking*, 39(6):1375–1409.

Lin, P.-S., Chang, C.-O., and Foo Sing, T. (2016). Do housing options affect child birth decisions? Evidence from Taiwan. *Urban Studies*, 53(16):3527–3546.

Love, D. A. (2010). The Effects of Marital Status and Children on Savings and Portfolio Choice. *Review of Financial Studies*, 23(1):385–432.

Makay, Z. (2020). The Family Support System and Female Employment. In *Demographic Portrait of Hungary 2018*, pages 85–105. Hungarian Demographic Research Institute.

Merlo, A., Ortalo-Magné, F., and Rust, J. (2015). The Home Selling Problem: Theory and Evidence. *International Economic Review*, 56(2):457–484.

Mian, A., Rao, K., and Sufi, A. (2013). Household Balance Sheets, Consumption, and the Economic Slump. *The Quarterly Journal of Economics*, 128(4):1687–1726.

Mulder, C. (2013). Family dynamics and housing: Conceptual issues and empirical findings. *Demographic Research*, 29:355–378.

Mulder, C. H. (2018). Putting family centre stage: Ties to nonresident family, internal migration, and immobility. *Demographic Research*, 39:1151–1180.

Mulder, C. H. and Billari, F. C. (2010). Homeownership Regimes and Low Fertility. *Housing Studies*, 25(4):527–541.

Mulder, C. H. and Lauster, N. T. (2010). Housing and Family: An Introduction. *Housing Studies*, 25(4):433–440.

Ortalo-Magné, F. and Prat, A. (2016). Spatial Asset Pricing: A First Step. *Economica*, 83(329):130–171.

Ortalo-Magné, F. and Rady, S. (1999). Boom in bust out: Young households and the housing price cycle. *European Economic Review*, 43:755–766.

Ortalo-Magné, F. and Rady, S. (2006). Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints. *Review of Economic Studies*, 73(2):459–485.

Oswald, F. (2019). The effect of homeownership on the option value of regional migration. *Quantitative Economics*, 10(4):1453–1493.

Parent, D. and Wang, L. (2007). Tax incentives and fertility in Canada: quantum vs tempo effects. *Canadian Journal of Economics/Revue canadienne d'économique*, 40(2):371–400.

Piazzesi, M. and Schneider, M. (2016). Housing and Macroeconomics. In *Handbook of Macroeconomics*, volume 2, pages 1547–1640. Elsevier.

Sommer, K. (2016). Fertility choice in a life cycle model with idiosyncratic uninsurable earnings risk. *Journal of Monetary Economics*, 83:27–38.

Spéder, Z., Murinkó, L., and Oláh, L. S. (2020). Cash support vs. tax incentives: The differential impact of policy interventions on third births in contemporary Hungary. *Population Studies*, 74(1):39–54.

Ström, S. (2010). Housing and First Births in Sweden, 1972–2005. *Housing Studies*, 25(4):509–526.

Szalma, I. and Takács, J. (2015). Who Remains Childless? Unrealized Fertility Plans in Hungary. *Czech Sociological Review*, 51(6):1047–1076.

Sági, J., Tatay, T., Lentner, C., and Neumanné, I. V. (2017). Certain Effects of Family. *Public Finance Quarterly*, 62(2):171–187.

Todd, P. E. and Wolpin, K. I. (2006). Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility. *American Economic Review*, 96(5):1384–1417.

Vidal, S., Huinink, J., and Feldhaus, M. (2017). Fertility Intentions and Residential Relocations. *Demography*, 54(4):1305–1330.

Vignoli, D., Rinesi, F., and Mussino, E. (2013). Fertility Intentions and Housing Conditions in Italy: A Home to Plan the First Child? *Population, Space and Place*, 19(1):60–71.

Yao, R. and Zhang, H. H. (2005). Optimal Consumption and Portfolio Choices with Risky Housing and Borrowing Constraints. *Review of Financial Studies*, 18(1):197–239.

Zeman, K., Beaujouan, E., Brzozowska, Z., and Sobotka, T. (2018). Cohort fertility decline in low fertility countries: Decomposition using parity progression ratios. *Demographic Research*, 38:651–690.

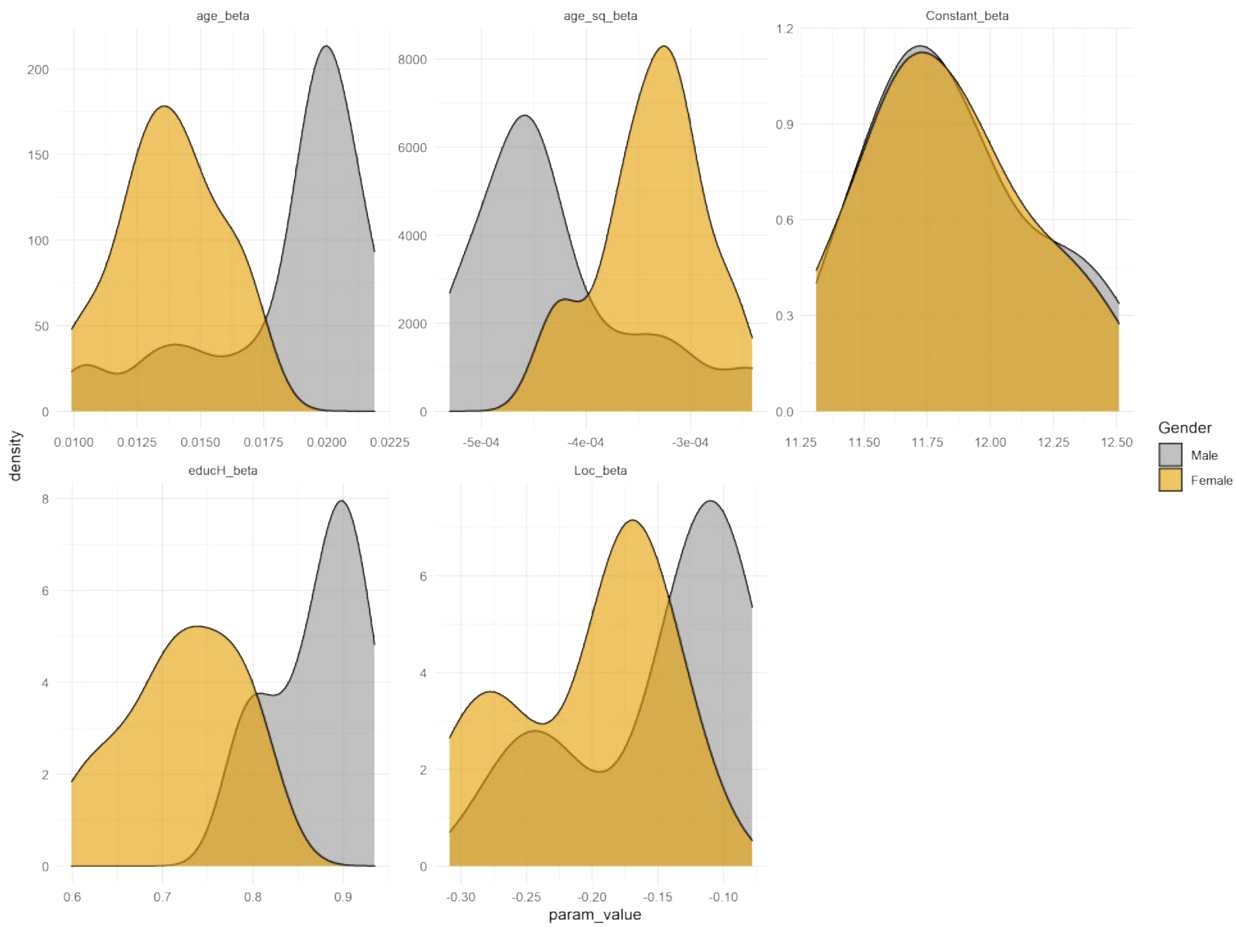
Öst, C. E. (2012). Housing and children: simultaneous decisions? – a cohort study of young adults' housing and family formation decision. *Journal of Population Economics*, 25(1):349–366.

Öst, C. E. and Wilhelmsson, M. (2019). The long-term consequences of youth housing for childbearing and higher education. *Journal of Policy Modeling*, 41(5):845–858.

Appendix

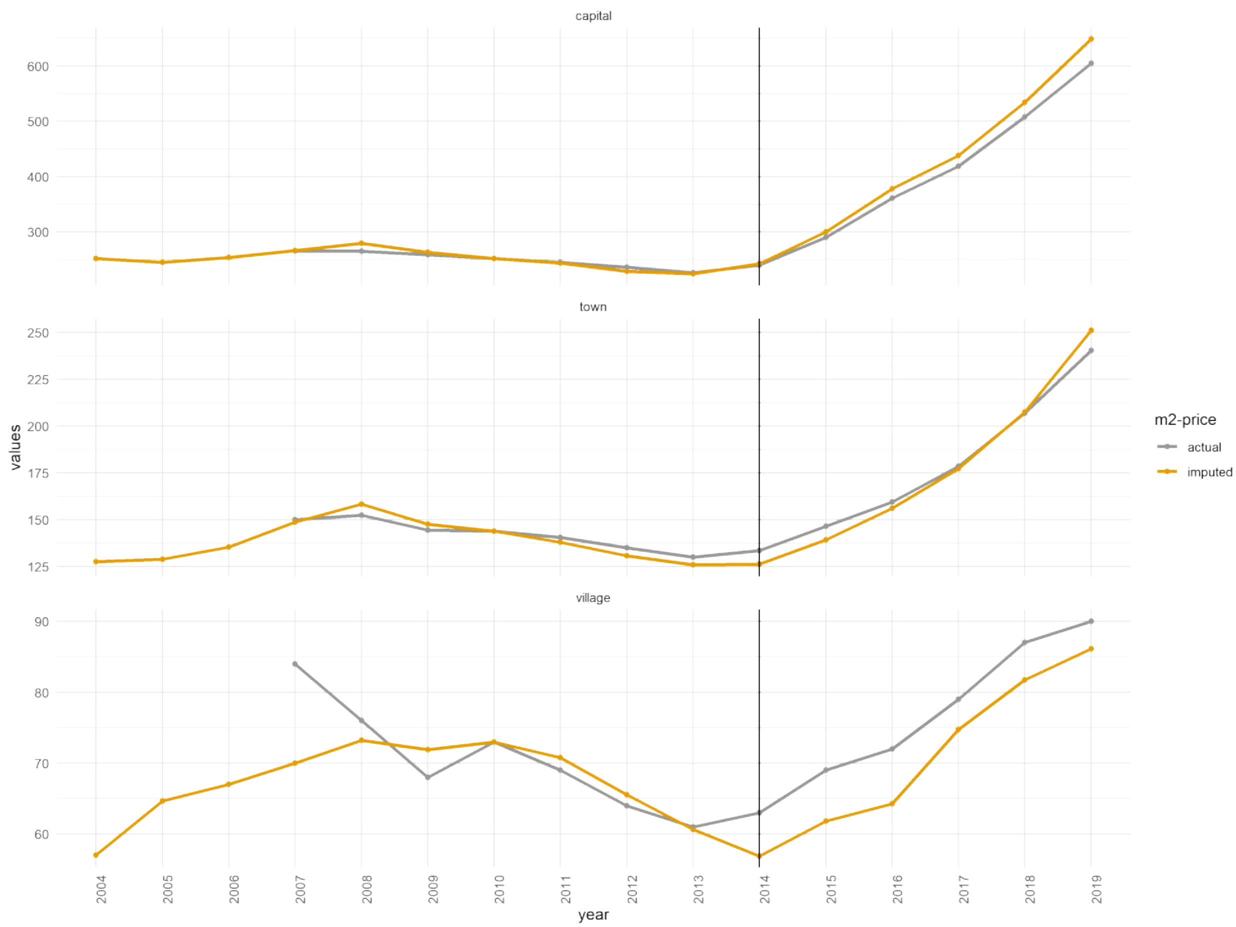
All additional figures and programs are available upon request.

Figure 3A1: Estimated densities of the estimated wage regression parameters, 2004-2018



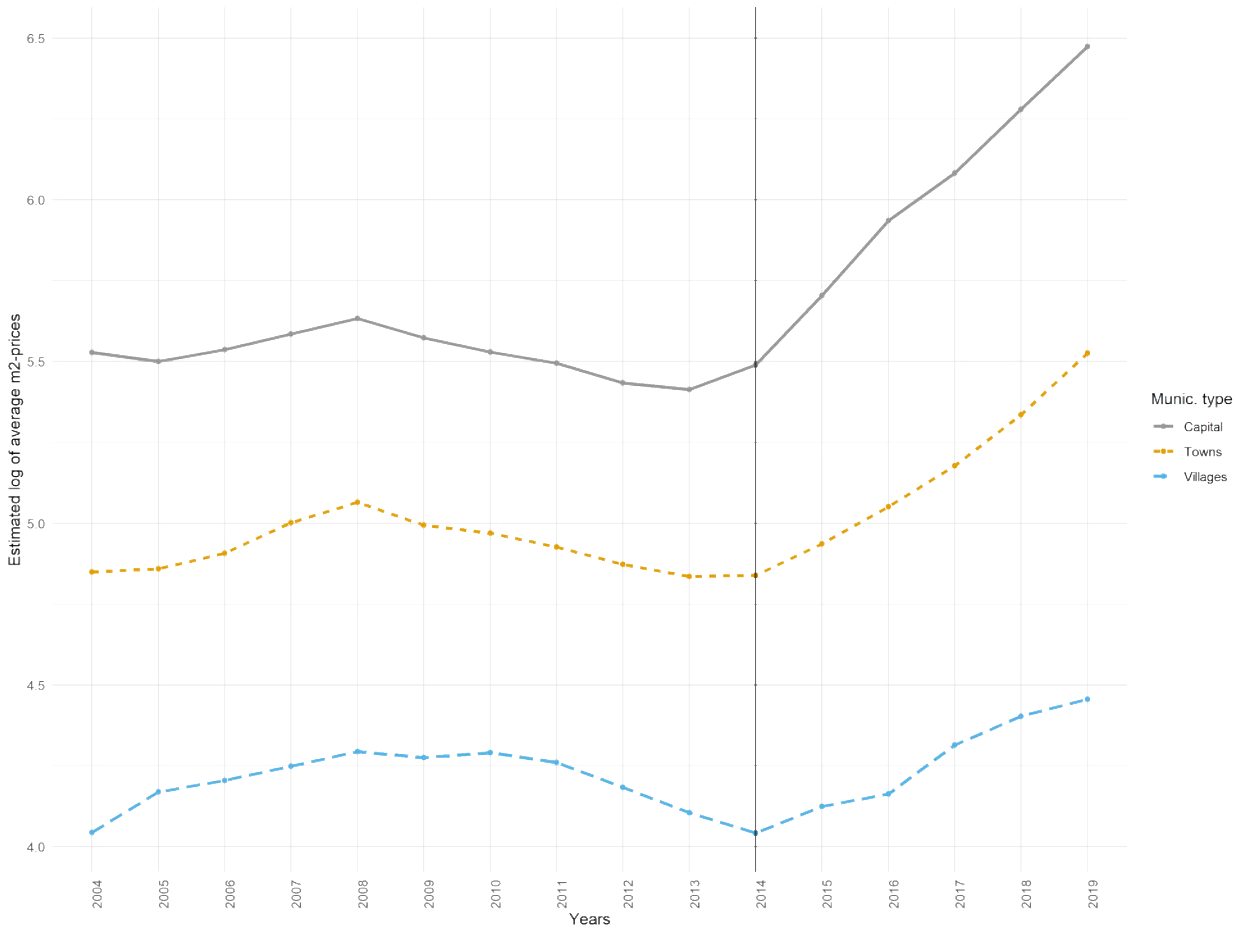
Note: the author's calculations based on the regression estimates from the Wage Survey data of Hungary, 2004-2018.

Figure 3A2: Imputed m^2-prices compared to available estimates



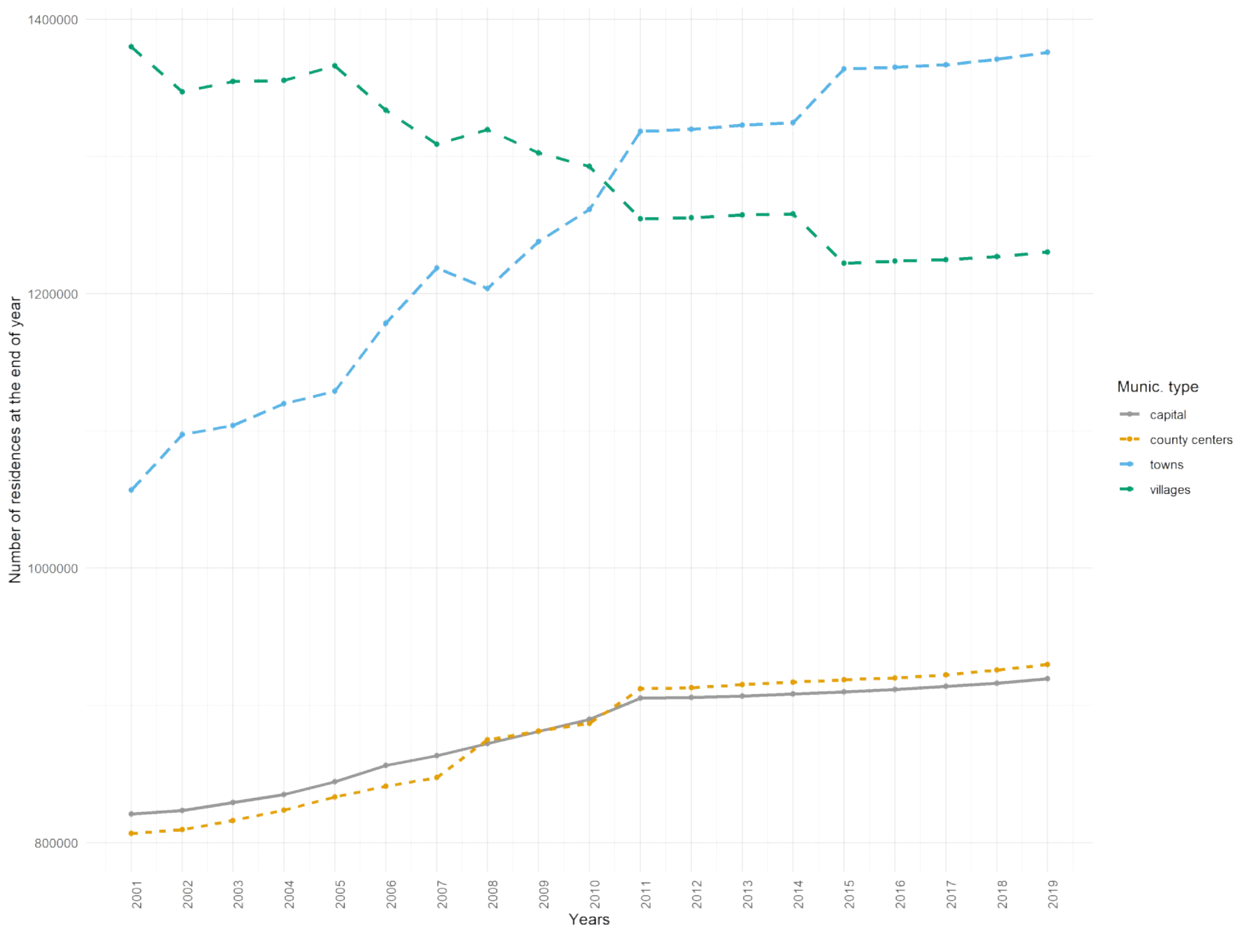
Note: the author's calculations based on the publicly available data of the Hungarian National Bank (HNB) and the Hungarian Central Statistical Office (HCSO). The earliest average m^2-price estimates of the HCSO are only available from 2007, while the HNB publishes nominal and real house price indices for earlier years. I use the two sources of information to estimate the pre-2007 prices, using 2010 as the baseline.

Figure 3A3: Logarithm of yearly average m^2-prices in 1000s of HUF, by type of municipality, 2004-2019



Note: the author's calculations based on the nominal house price index of the Hungarian National Bank and house price estimates of the Hungarian Central Statistical Office.

Figure 3A4: Number of houses at the end of the year, by type of municipality, 2001-2019



Note: based on the publicly available data of the Hungarian Central Statistical Office.

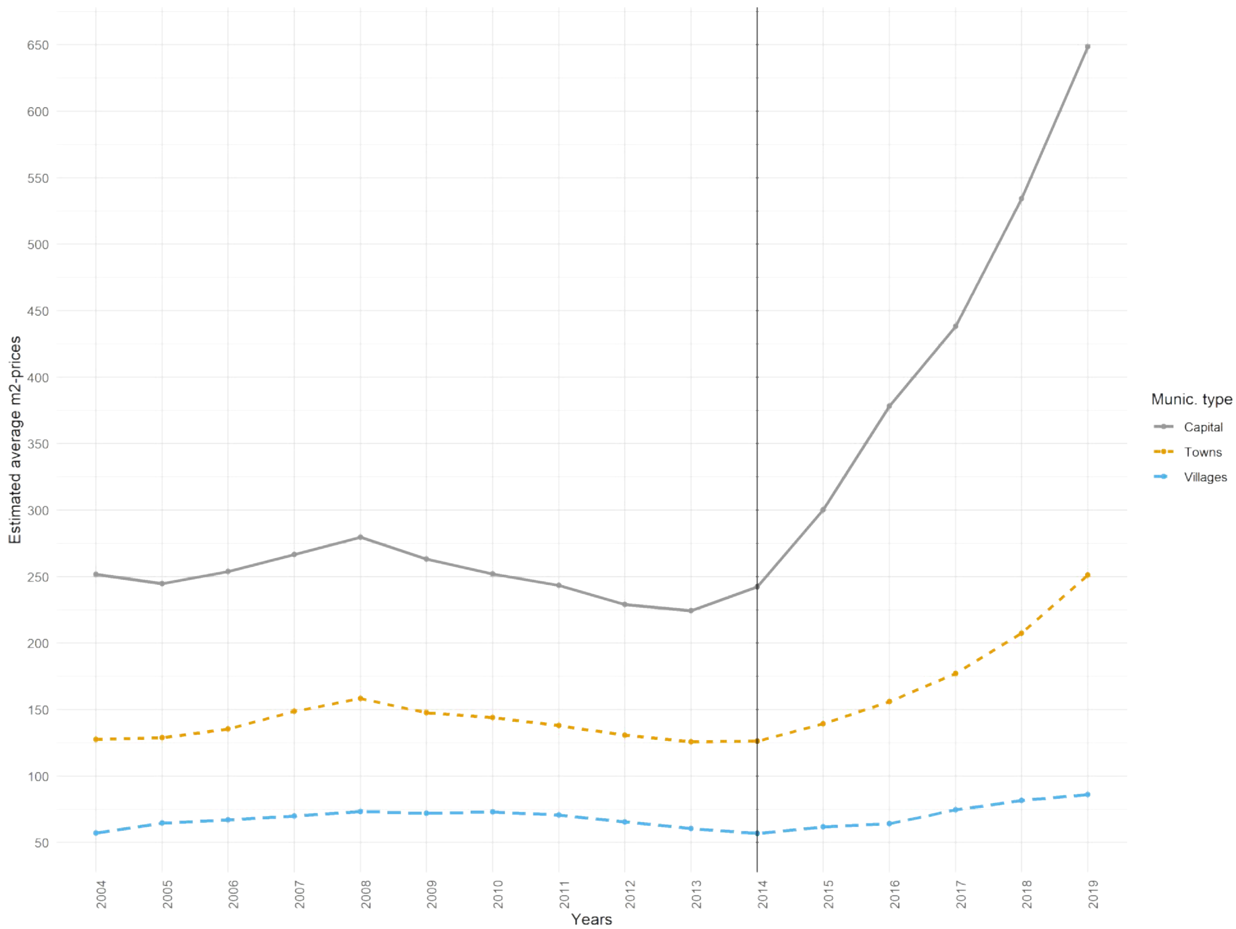
Figure 3A5: Evolution of housing variables under different scenarios

Note: the author's calculations based on 150 simulations of the model. One period on the horizontal axis equals two years.

Figure 3A6: Evolution of non-housing variables under different scenarios

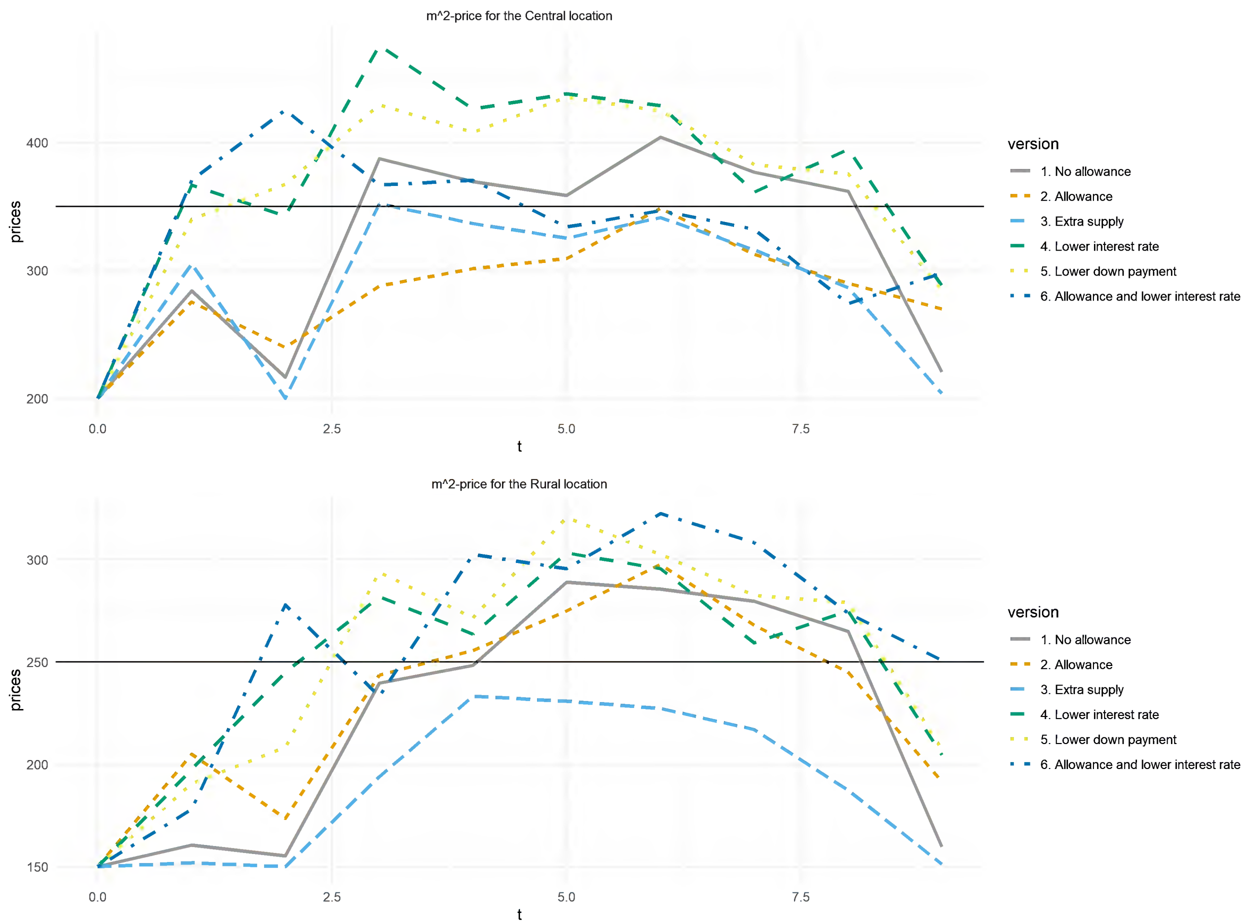
Note: the author's calculations based on 150 simulations of the model. One period on the horizontal axis equals two years.

Figure 3A7: Average m^2-prices in 1000s of HUF, by type of municipality, 2004-2019



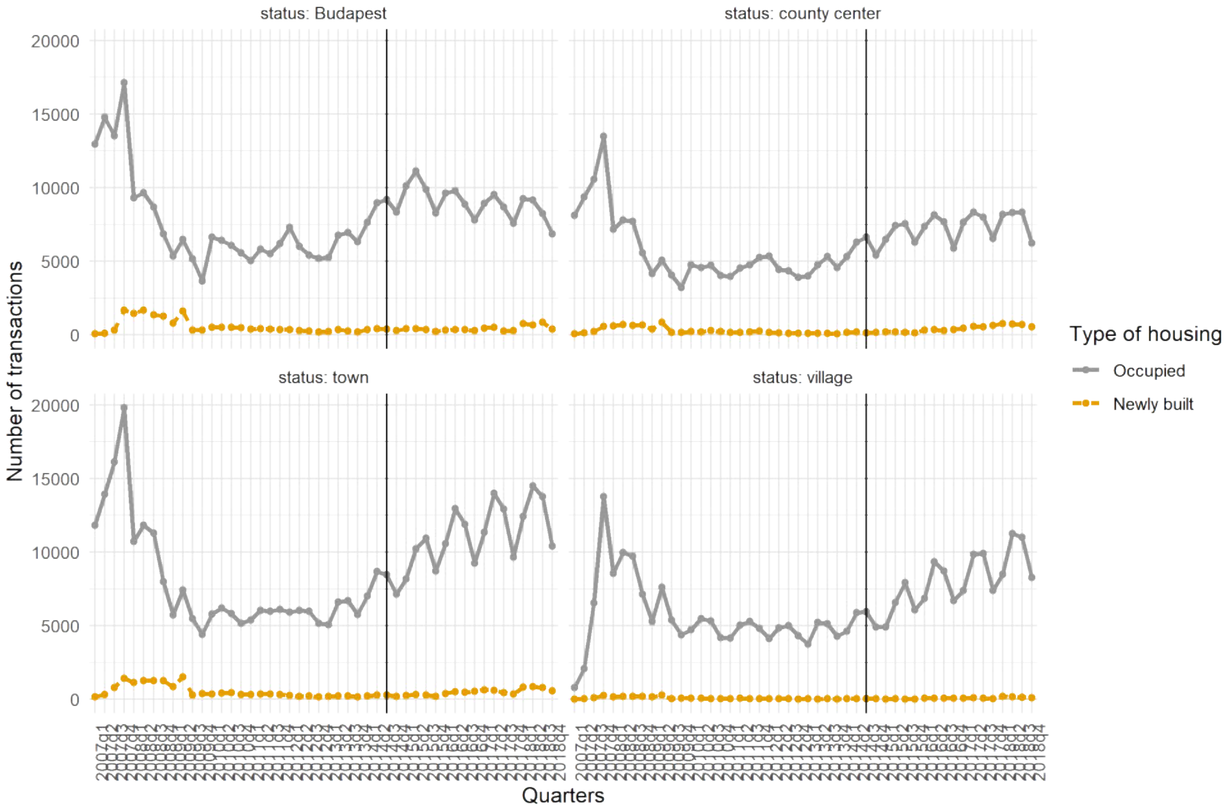
Note: the author's calculations based on the nominal house price index of the Hungarian National Bank and house price estimates of the Hungarian Central Statistical Office.

Figure 3A8: Average m^2-prices for the scenarios



Note: the author's calculations based on 150 simulations of the model. One period on the horizontal axis equals two years. Targets of the prices are indicated with the horizontal lines at 350 for the central and 250 for the rural area.

Figure 3A9: Number of housing transactions by type of municipality and house, 2007q1-2018q4



Note: based on the publicly available data of the Hungarian Central Statistical Office.