Tax Incentives and Completed Fertility*

Do not cite without the author's permission.

Bence Szabó[1]

[1]bence.szabo@phd.unibocconi.it

October 25, 2022

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. To identify the reform effect, 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 less 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 that the policy would result in a qualitatively similar completed fertility effect 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, 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 Introduction

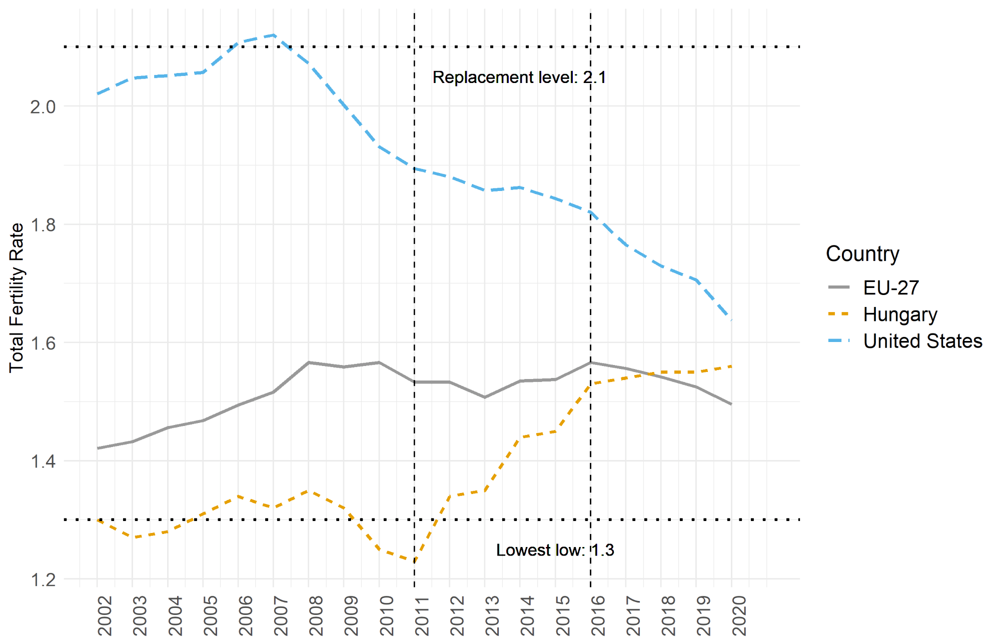
Persistent low fertility is one of the central policy issues in developed countries. Since the 1970s, the total fertility rate[1] in OECD countries have declined from around 2.84 to 1.61, below the replacement level of 2.1 that would maintain the population size. As Figure 1 shows, below replacement 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 fertility what is considered 'lowest-low', below 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 changing preferences of families regarding the number of children. However, evidence suggests that in many developed countries, the realized number of children indeed lags behind the preferred amount[3]. Figure 2 displays the difference between average completed fertility at age 40 and the ideal number of children in European countries, showing that in most countries, families end up with fewer children on average than what they would ideally prefer. Therefore in several countries, there is an aggregate economic interest aligned 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 child-bearing 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 in 45 years (Kohler et al., 2006).

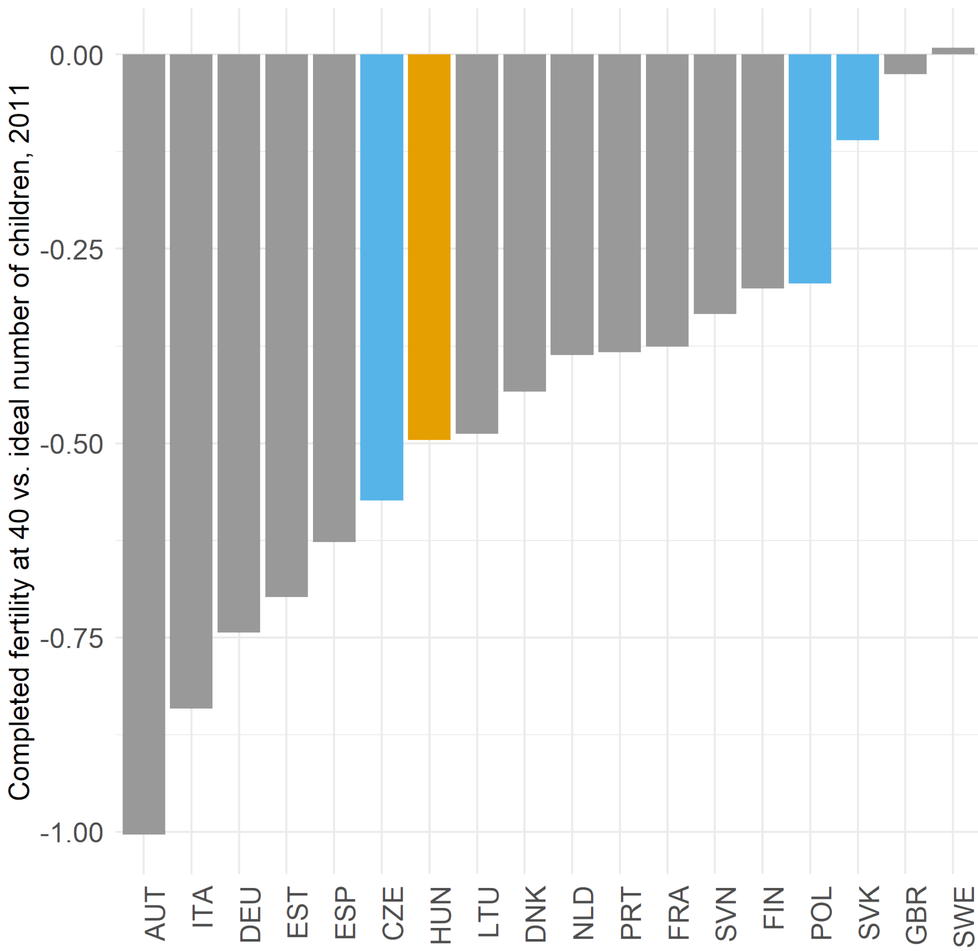
[3]Measuring fertility preferences, however, is a complicated issue, as described by (Spéder and Kapitány, 2009).

Figure 1: Evolution of Total Fertility Rates



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).

Figure 2: 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 focusing on Hungary. This developed European country experienced 'lowest-low' fertility levels for an extended period before 2011, rebounding to the EU average by 2016. The Hungarian Government introduced a large-scale expansion of the family tax break in 2011, increasing its magnitude from 0.05% of the GDP to 0.64% of the GDP, around 17% percent of all public spending on families (OECD, 2021). The policy's structure includes non-linearities due to a large jump at the third child and tax base deductions being non-refundable. 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 for a rich set of observable characteristics not affected by the policy. According to the labor market information, I impute the prime age prospective income for the adults in the household from 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 female employment.

The findings suggest the policy increased completed fertility for families with at least two children by around 0.76% and 5.64% for those with less than two children, with the weighted average being around 2.36%. The results are driven by religious, married couples where the age of the youngest child is less than three. Estimates regarding indirect maternal employment are not significant overall; however, there seem to be some negative employment effects for those segments where completed fertility increased. The estimates are robust to several changes in the specification. I also estimate a simple partial equilibrium model of household behavior to simulate counterfactual policies, and to examine possible longer-run effects regarding maternal employment, completed fertility, and income redistribution. A comparative statics exercise using the structural model indicates an around 3.25% higher completed fertility driven by middle-income families and around 4.8 percentage points lower employment driven by lower-income families due to the original policy. I also find that alternative policies might have led to slightly higher fertility effects without changing the overall public expenses on the policy, without regressive income distribution consequences.

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 theoretical 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 predictions about optimal policies as a function of underlying preference parameters. While theoretical papers focus on completed fertility and the change in the number of children by the end of the female fertility cycle, empirically, such effects are not easy to detect. Most policies are expected to induce two types of adaptation: a change in the timing of births and a change in completed fertility, or the so-called *tempo* vs. *quantum* effects. It is easy to see that several policies might induce families to bring their births forward in time to receive the benefits earlier. Still, in the end, the completed number of children might stay 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 unaffected by the policy (Parent and Wang, 2007; Kim, 2014). Adda et al. (2017) uses a dynamic structural model based on German data to study women's fertility decisions and career choices to differentiate between short-run and long-run consequences. They find that relatively large short-run effects of a one-time birth subsidy could be dominantly due to a change in the timing of births and mostly mitigated over the life-cycle of women.

There is large empirical literature on the effect of economic circumstances, such as unemployment or insecurity, on fertility (Currie and Schwandt, 2014; Sommer, 2016; Clark and Lepinteur, 2020), but also specifically on the role of tax-related incentives. One branch of literature made use of low-level aggregated data in quasi-panel settings mainly concerning the Earned Income Tax Credit in the United States (Moffitt, 1998; Baughman and Dickert-Conlin, 2003; Gauthier, 2007; Baughman and Dickert-Conlin, 2009; Thévenon and Gauthier, 2011), or for Hungary most recently Bördős and Szabó-Morvai (2021). 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). The last two decades more studies have been turning to individual-level data 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 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 can 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 and 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 2. In Sections 3 and 4, I detail the construction of the analysis dataset and establish the empirical strategy for identifying the reduced form results. In Sections 5 and 6, I present the results, run several robustness checks, and then continue in Section 7 to build and estimate a simple household model examining different counterfactual scenarios. Finally, I discuss the findings in Section 8 and conclude in Section 9.

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 only a timing effect occurs (Bongaarts and Sobotka, 2012). However, in Hungary, evidence points to the fact that, especially at the second birth parity, a decline in completed fertility also took place (Kapitány and Spéder, 2015). At the same time, childlessness has slowly become more prevalent for younger cohorts childlessness has also started to become more prevalent (Spéder, 2021).

Although some more minor interventions were affecting fertility behavior before 2011 as well (Gábos et al., 2009; Spéder et al., 2020), the governments after 2010 made fertility a focal point of economic policy. 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 main policy instruments, see Table A1). The overhaul of the family tax break was implemented as a part of personal income tax reform. Still, it played an instrumental part in major redistribution towards higher income families with more children (Tóth G. and Virovácz, 2013; Varga et al., 2020).

The policy instrument 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[4]. 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 similarly to the Child Tax Credit, there is no phasing out at higher levels of income (Hoynes,

[4]Note that since 2014, families not being able to claim the total amount based on personal income 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.

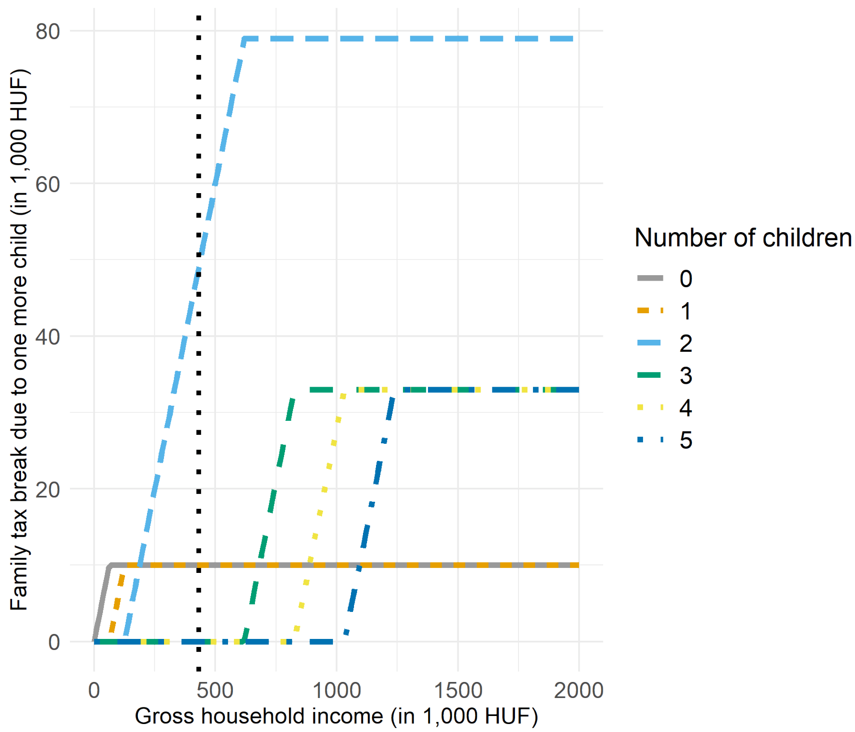
2019; Goldin and Michelmore, 2021), which results in regressive redistribution.

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[5]. If the family had less 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 could not claim the full amount of the tax break with three children; a family would need 45% more than the average income.

Figure 3 displays 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 regardless of income, as tax bases are always sufficiently large to deduct the maximum amount. However, for households with two children, one additional child could give around 80,000 HUF monthly increment if income is high enough, increasing at the slope 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 limits.

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

Figure 3: 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.

3 Data

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, and without a unique identifier, so we must create a new identifier on which we can match the observations. This identifier is based on the male or female self-identifying as the leader of the household and spouse of the leader of the household, allowing for the inclusion of non-married couples. 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 necessary, 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 the selection issues originating from being able to observe 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 (45 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 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.[6]

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 algorithms (or more advanced ones such as Abramitzky et al., 2020) or any other possibly endogenous variables. However, other than data errors, we can be quite confident that the families found in both Census datasets are the same. Table 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.

[6] The fraction of births to mothers forty or more years old was around 4.5% in 2016, according to the Central Statistical Office in Hungary.

Table 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 4 displays the matching success rates by cohorts (left panel), along with estimated statistically significant coefficients after 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 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.[7]

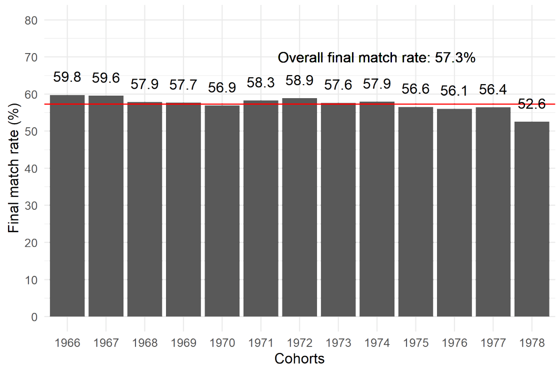
Linking families existing in 2016 to themselves in 2011 requires that the family unit was observed in both censuses. It is impossible if the family unit did not exist earlier, or it was not enumerated because of not being present in the country due to emigration, for instance, or they were not cohabiting officially. The right panel of Figure 4 helps to give an idea of which families were matched more successfully. 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. 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 constellation are the ones that might be intuitively more stable and longer-lasting. This exercise suggests that

[7] 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.

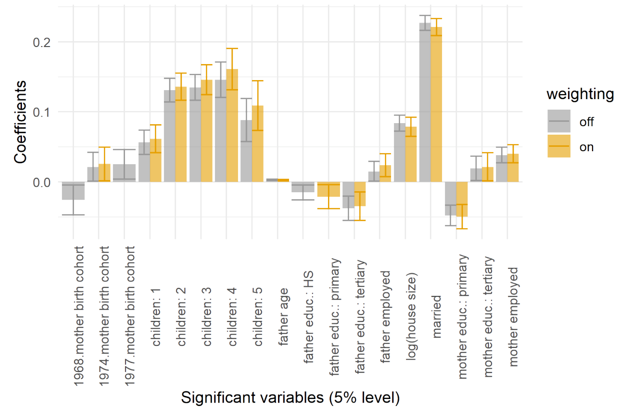
when we interpret the findings, we should keep in mind that validity might be limited to those couples that are intrinsically more stable.

Figure 4: 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.

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. It might be true as 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). Another advantage is that personal unobserved earning capabilities (potentially correlating with fertility decisions) within the same subset of people cannot play a role in inducing bias in the estimates and that 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)

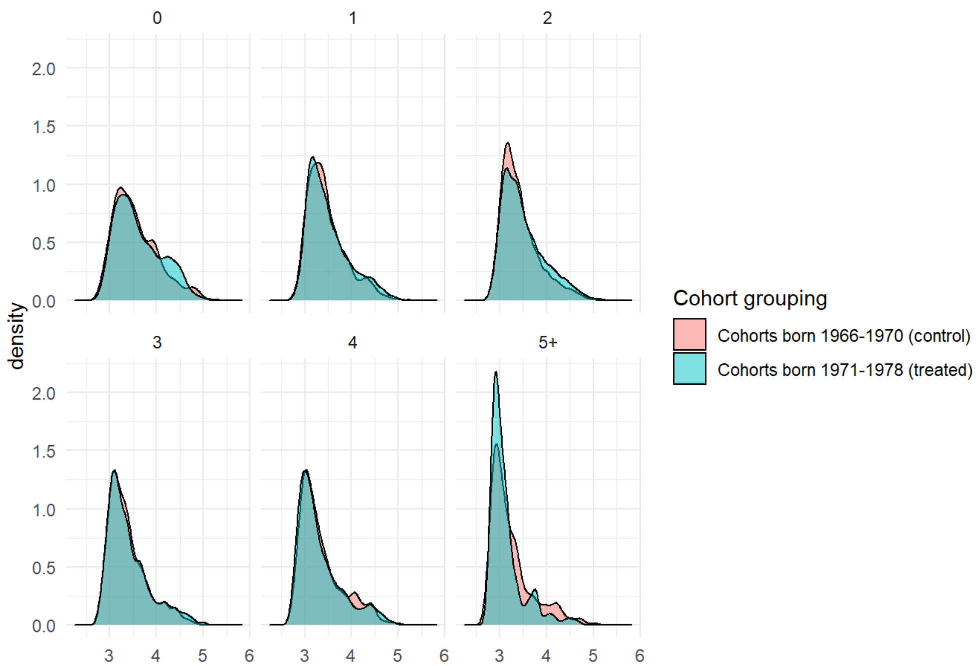
The salary match rates corresponding to the 2016 Microcensus are displayed in Table 2 (the rates for the 2011 Census are similar). We can see that more than 95% of the relevant sample could be matched at least using occupation information. In contrast, gender, region, municipality type, and education 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 5 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 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 5: Kernel density of the log of imputed prospective household income by cohort groups and the number of children in 2011



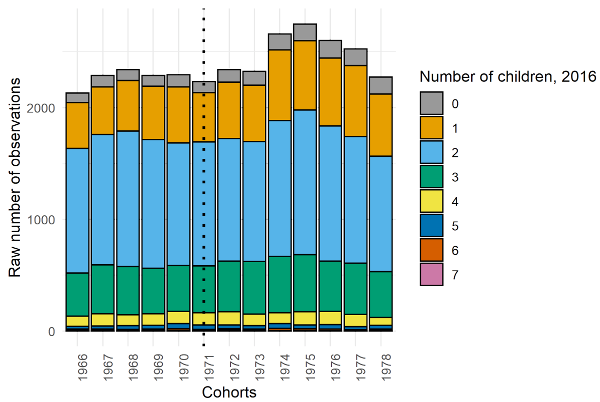
Note: the figure shows the estimated kernel densities of log household incomes (measured in 10,000 HUF units) by cohort groups and the number of children, based on the matched dataset of the 2016 Microcensus and the 2011 Census of Hungary.

3.3 Descriptive statistics

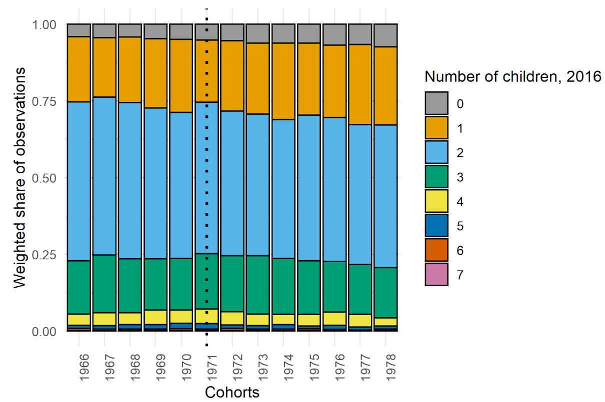
This section gives an overview of the final analysis sample. It shows the composition regarding cohorts and the number of children, the imputed prospective household incomes, and the change in the value of the outcome variables over time. Figure 6 displays the unweighted and weighted number of observations and their distribution by the number of children in 2016. The observation window spans 38-50 years for the relevant cohorts, implying that we can treat all except for the 1977 and 1978 cohorts at the end of their fertility cycle. 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 6: Observations in the analysis sample by the number of children in 2016

(a) Unweighted number of observations



(b) Shares using the 2016 Microcensus weights



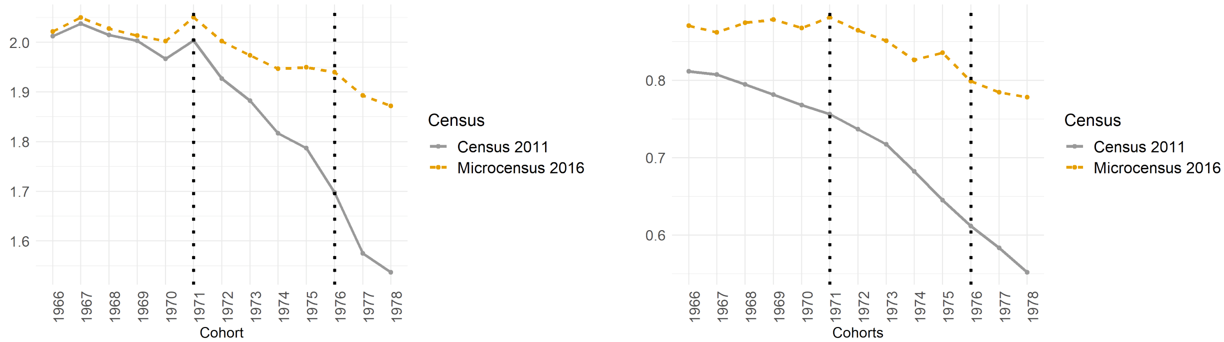
Note: the number of families in the sample, based on the matched dataset of the 2016 Microcensus and the 2011 Census of Hungary.

I am targeting two outcome variables in this paper: completed fertility and 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 and house size as validation exercises, which are displayed in the Appendix.

Figure 7 shows the averages of relevant 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 level. On average, around 0.25 additional children were born during these five years, around 13% of completed fertility. Shares of employed amongst the matched sample increased for both the fathers and the; however, for the latter, the relative change is substantially larger. 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.

Figure 7: Outcome variables in the analysis sample

(a) Number of Children (b) Female Employment



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

4 Empirical strategy

4.1 Identification

As mentioned earlier, this study relies on the 2011 policy change in Hungary introducing large tax base deductions as a function of the number of children and gross household income. The policy introduces non-linearities as shown earlier (Figure 3), 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 because 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 treatment (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 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 environment 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 should be 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 assume that unobserved heterogeneity denoted by $\eta_{i,c}$ is additively separable. Then for cohort c:

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

so 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 sample, and the conditional expectation of the unobserved heterogeneity.

It is easy to imagine that $\mathbb{E}_i[\eta_{i,c}|X_i, T_i] \neq 0$, 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: for instance, 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:

$$Y_{i,c} = \theta_c \mathbb{I}[i \in c] \times T_i + \alpha_c + \beta_1 T_i + \beta_2 \mathrm{W}_{i,2011} + \sum_{k \in \{0,\ldots,7\}} \gamma_k \mathbb{I}[N_{i,2011} = k] + \delta X_i + \varepsilon_{i,c}$$

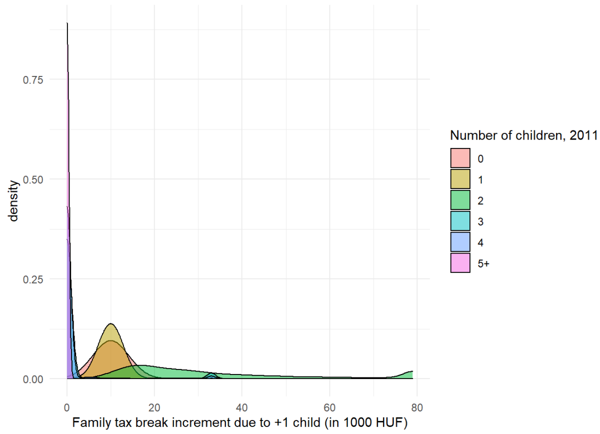
where for individual i in a biannual cohort c, $Y_{i,c}$ denotes the outcome, T_i denotes the amount of family tax break received as additional net household income, $W_{i,c,2011}$ the gross household salary income in 2011, $N_{i,c,2011}$ the number of children in 2011, and $X_{i,c}$ the additional controls.

4.2 Treatment definitions

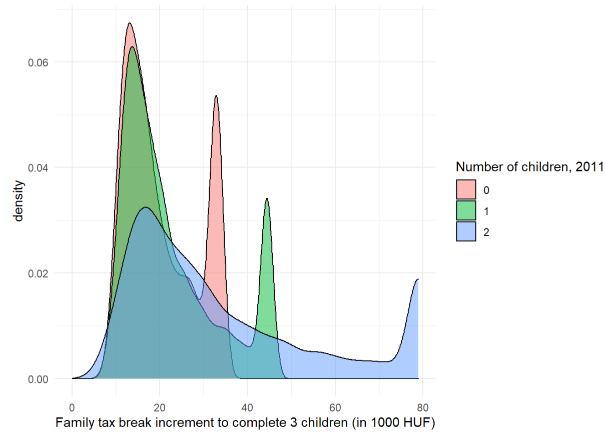
The policy creates two types of changes: on the one hand, it induces differential financial incentives to have one additional child for families with at least two children, and on the other hand, it alters incentives for families with less than two children to end up with three children. The more straightforward treatment variable we can construct is the family tax break increment for one additional child, which lowers the relative cost of one additional child. In this setting, most of the identifying variation comes at the third child, while there is almost no variation below that. 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. Ex-ante, a family might decide to increase the number of children from zero or one to three instead of two as a response to the policy incentives. Figure 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. 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 remains residual variation even after controlling for the number of children and income.

Figure 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 the extra money per month in 1,000 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).[8] Based on this setup, I define the two types of continuous treatment

[8]Rules of taxation over the period has changed slightly, after a significant overhaul in 2011 starting from 16% rate,

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:

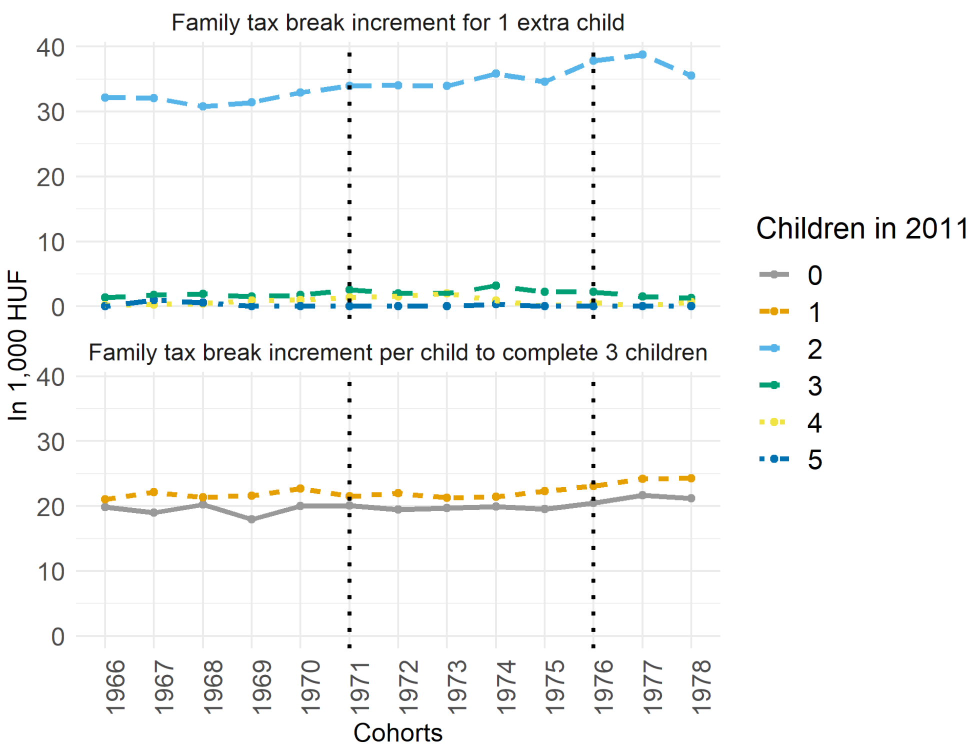
1. $T_{1,i} = M(N_{i,2011} + 1, S_{i,2011}) - M(N_{i,2011}, S_{i,2011})$: Additional money in 10,000 Hungarian Forints (HUF) per month for one additional child (examine separately also for families with at least two children, as it varies very little due to the low amount of deduction for zero and one 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 9 shows the average values of the family tax break variables and the imputed prospective household income by cohorts and number of children. By definition, the number of children and income jointly determine the number of tax breaks 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.

which decreased to 15% in 2016.

Figure 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). This value defines 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 ones, but we can conceptualize the treatment as the state financing a child's most basic needs.

5 Results

5.1 Completed fertility

I examine completed fertility (measured by the logarithm of the number of children in 2016 + 1)[9] and maternal employment as relevant outcomes of the policy.[10] 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 [11]. There are two treatment variables: the additional money received from one more child with meaningful variance for families with at least two children and the extra money received after completing three children for families with less 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 older families in the counterfactual group who are more similar to the fertile age cohorts, but as I show later 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 indeed not statistically different 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 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.[12] For

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

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

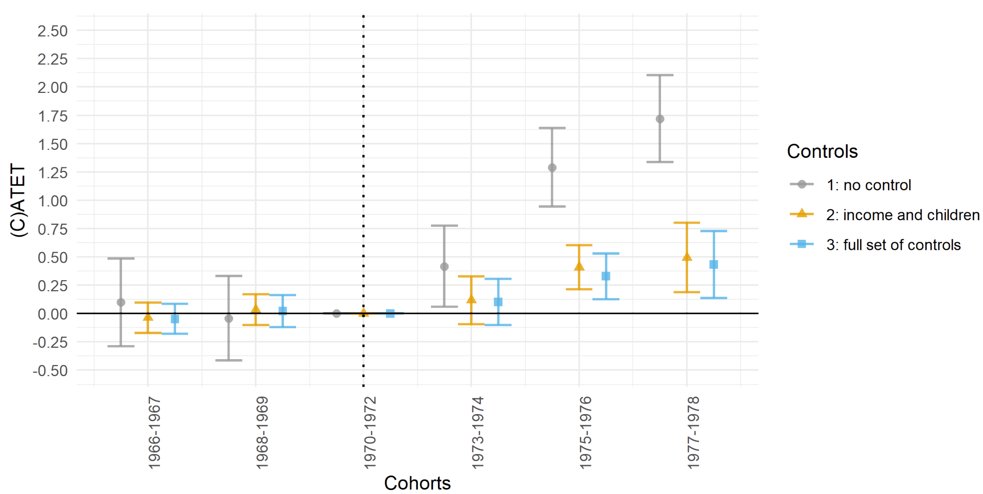
[11]The set of controls includes the following: number and gender composition of children; household income; parental education, employment, occupation, industry, log of prospective labor income, and age; home ownership status; living conditions (number of rooms, heating, comfort level, house size); 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.

[12]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.

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.

Figure 10 presents 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 0.76% completed fertility effect at these birth parities. There are no statistically significant effects for the 1973-1974 cohorts. Furthermore, the non-significant point estimates without controls for the older cohorts indicate that the assumption regarding unconfoundedness is supported, and we could interpret the estimates for the fertile cohorts causally.

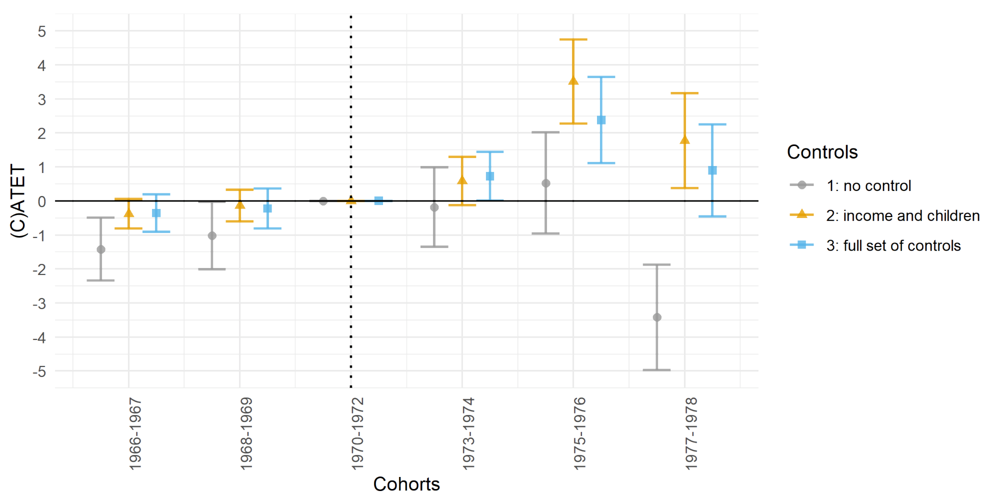
Figure 10: Effect of 10,000 HUF family tax break on completed fertility (in %), families with at least two children



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 less than two children are reported by Figure 11. While we see no significant effects without controls and some evidence for pre-trends in 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 2.38% per 10,000 HUF is estimated, which at the mean treatment of 2.37 implies a 5.64% effect at the lower birth ranks.

Figure 11: Effect of 10,000 HUF family tax break on completed fertility (in %), families with less than two children



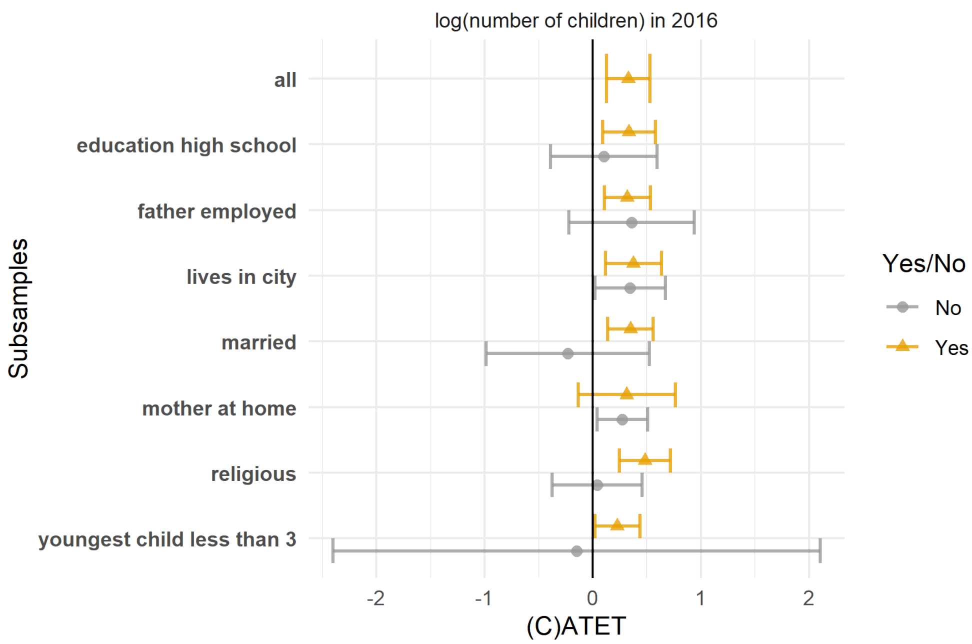
Note: based on regressing the log(1+number of children) in 2016 on cohort groups interacted with treatments, with a baseline of 1970-1972. The full set of controls includes the following from 2011 besides household income and number of children: gender composition of children; parental education, employment, 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 less 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 helps to study which segments of the population are more responsive to financial incentives regarding their completed fertility. To address this question, I run the same regressions for several relevant subgroups[13], and I display the estimates for the 1976 cohort in Figure 12 and Figure 13. For families with at least two children, there are three aspects with some differences. Although the confidence intervals overlap in all cases, the confidence bounds for religious, married couples, and where the youngest child is less than three years old are above

[13]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. '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. Religious indicates if both parents indicate belonging to a religious group. 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 point estimate of their counterpart. One-sided Z-tests[14] would suggest that religiousness is significant at the 5% level, while marital status is so at the 10% level. The picture is somewhat different for families with less than two children (Figure 13). While the youngest child being less than three years old again seems to matter (one-sided Z-test indicating significance at the 10% level), no other variable plays a role in creating heterogeneity in the effect of the policy.

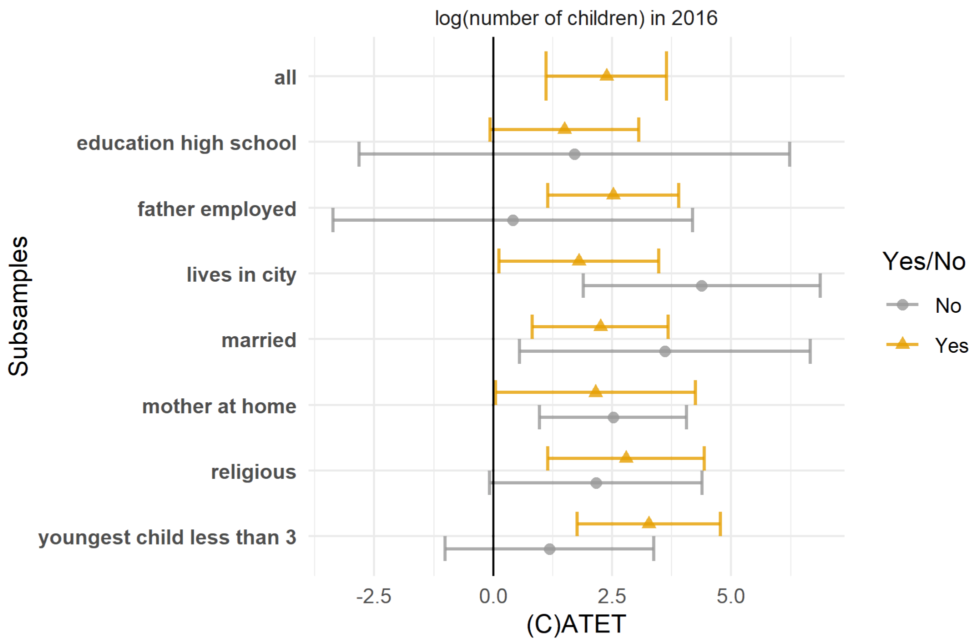
Figure 12: Effect of 10,000 HUF family tax break on completed fertility for the 1975-1976 Cohorts (in %), families with at least two children



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.

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

Figure 13: Effect of 10,000 HUF family tax break on completed fertility for the 1975-1976 Cohorts (in %), families with at least two children



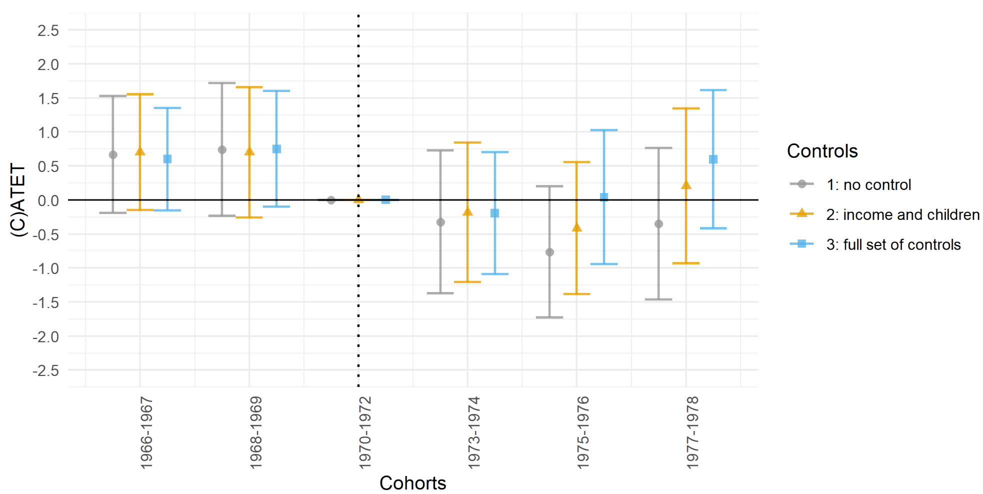
Note: based on regressing the log(1+number of children) in 2016 on cohort groups interacted with treatments for the 1975-1976 cohort, with a baseline of 1970-1972 for different subsamples. The full set of controls includes the following from 2011 besides household income and number of children: gender 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 less than two children by the mother's age of 35. Sample and treatment sizes vary between subsamples.

5.2 Mother's employment

As a secondary outcome, we examine the mother's employment status in the household. These results reflect a short-run labor market change that should be understood in connection to our previous findings on fertility. There are economic models describing how short-run and long-run labor supply decisions are intrinsically related to women's fertility choices due to high costs on their careers, which affect their sorting into occupation or education groups (Adda et al., 2017). As we only observe households for which only late fertility adjustments can be made, long-run labor market effects cannot be identified in this setup as all included cohorts are affected similarly. I use a simple model simulation exercise later to shed light on the longer-run effects in that regard. It is also important to note that the employment status of older cohorts should be unaffected by the change in the cost of children, as any measured effect should come as a consequence of new children so that we can use this as a validation exercise.

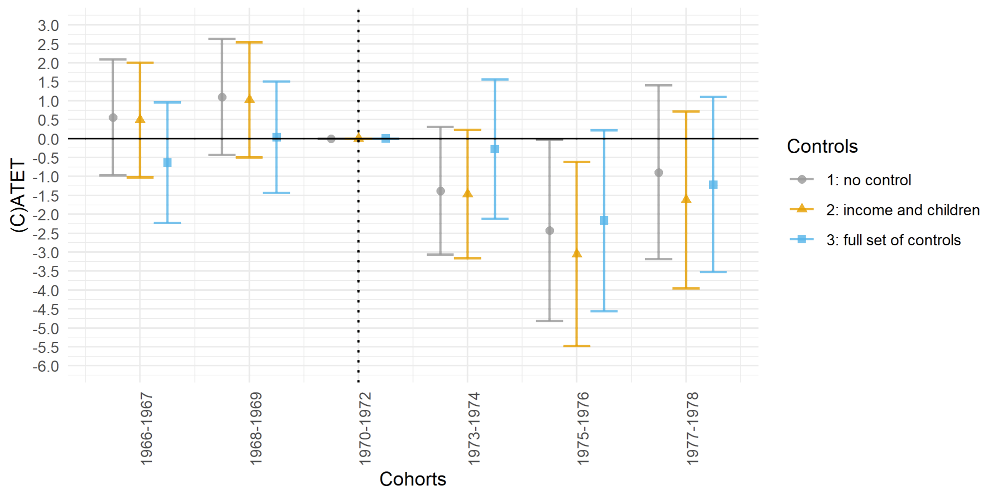
Figures 14 and 15 present the estimates for families with at least two, and less than two children respectively. We can see that no robust effects can be detected in either case, with some significant negative point estimates for families with less than two children if we do not include all controls. Reassuringly, the estimates for older cohorts are not statistically significant either, supporting the previous identifying assumptions.

Figure 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.

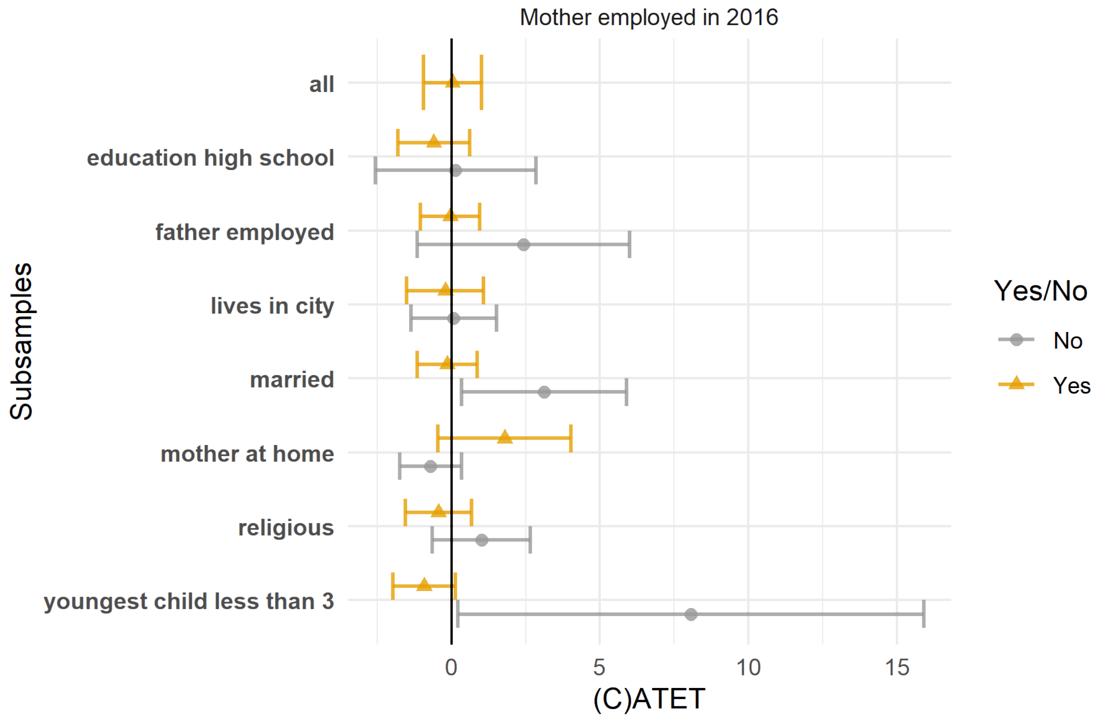
Figure 15: Effect of 10,000 HUF family tax break on maternal employment (in percentage points), families with less than two children



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 less than two children by the mother's age of 35, with a sample size of 10,243. The mean treatment size is 2.37.

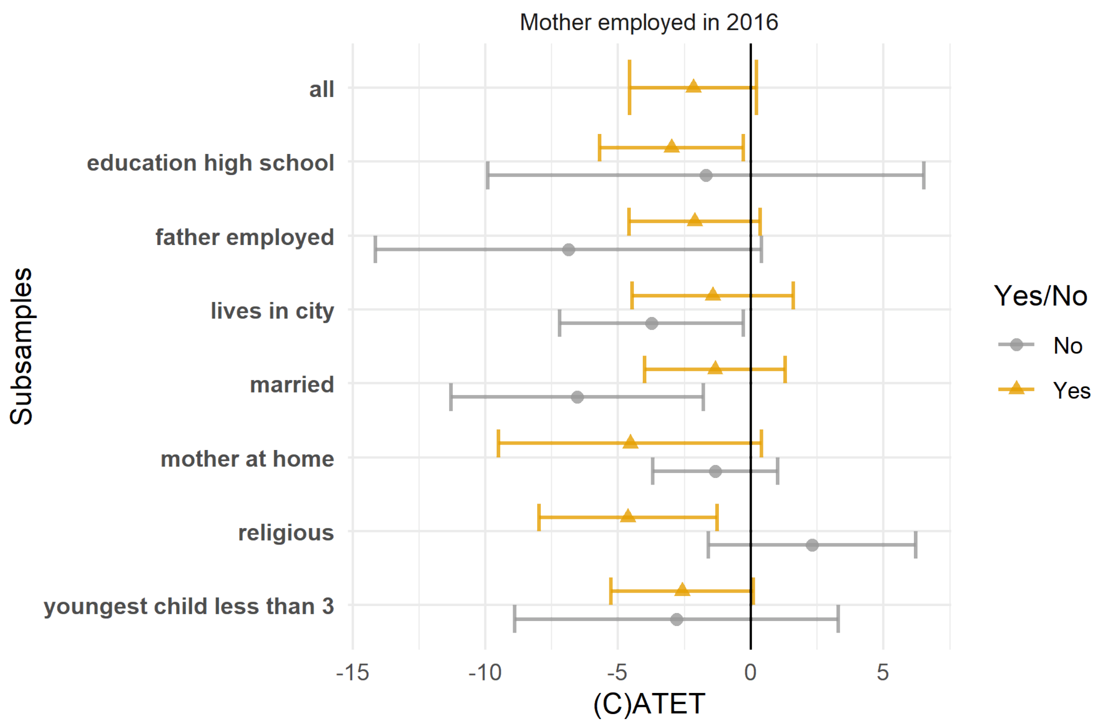
Finally, we can turn to the heterogeneity of the employment effects, displayed by Figures 16 and Figures 17, again using the 1970-1972 cohorts as a baseline. For families with at least two children (Figure 16), we can detect lower employment exactly where higher completed fertility is visible (religion, marital status, child's age), which is a natural consequence. The difference using two-sided Z-tests is statistically significantly negative for marital status, and the youngest child is less than three years old. However, we can see some statistically significant positive employment effects for mothers at home in 2011. For families with less than two children, Figure 17 suggests significant differences only along religiosity at the 5% level: maternal employment for more religious couples seems to be lower than their counterparts.

Figure 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.

Figure 17: Effect of 10,000 HUF family tax break on maternal employment (in percentage points) for 1975-1976 Cohorts, families with less 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 less than two children by the mother's age of 35. Sample and treatment sizes vary between subsamples.

6 Robustness of the results

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 3 and 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.

First, let us consider the families with at least two children (Table 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 outcome at about 2.49, an around 0.9% 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 0.63% effect of the policy, slightly smaller than the main specification.

Table 3: Robustness checks: completed fertility effect estimates for families with at least two children

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 4, which reports the robustness checks for families with less 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 in a family due to a 10,000 HUF tax break, which scaled up to the 2.37 mean treatment size and compared to the around 1.06 baseline outcome leads to a large, around 15.6% effect size. It is quite large compared to the main specification. Finally, using a discrete treatment definition, I find about 10% point higher completed fertility for a tax break higher than 25,000 HUF for the third child, which is present at 30% of the population, pointing to an around 3% 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. To conclude, the robustness tests regarding the results for families with less than two children show that these results are more sensitive to specification.

Table 4: Robustness checks: completed fertility effect estimates for families with less than two children

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. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

6.2 Robustness tests based only on the 2016 Microcensus data

To alleviate concerns about the bias the data linking might create, 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 36 and 35 years old at the policy's start and 41 and 40 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.

Table 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[15]. 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 0.18% effect size is detected, or 0.36% 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 0.46% (0.94% at the mean treatment), consistent with the fact that we expect higher short-run effects for cohorts not reaching completed fertility yet.

[15] 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 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 6 displays the estimates for families with less 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 1.22% effect (2.23% 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 2.07% (3.89% at the mean treatment).

Table 6: Robustness checks: completed fertility effect estimates for families with less 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 robustly detect completed fertility for the 1975-1976 cohorts, which this paper focuses on, with some evidence for higher short-run effects for the younger cohorts. However, this 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.

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 of 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 household 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, for lower-income households, it provides additional incentives to increase household income to the level where the entire tax break can be used. Besides comparing alternative policies, I use this model to validate the previous empirical results while shedding light on the policy's potential female labor supply effects that could not be identified.

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. I use this model for comparative statics to retrieve testable predictions regarding the effect of different tax policies.

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

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

$$s.t.$$

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

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

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 0 or 1 values, and W^M and W^F denote the prospective maximum household gross labor income given exogenously to the household. The fraction of gross income paid in taxes and adjusted for family allowance is given by $\tau(W^M,W^F,l^F,N)$, a function of the maximum prospective gross income of the household[16], labor

[16] I use the flat 16% personal income tax rate along with 17.5% rate of contributions, along with the family allowance

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 her 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 following the calculations 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 utility from parental consumption and children's. 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.

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.[17]

Table 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 Figure A17 shows that the found parameter values of $\alpha = 1.8838$ (SE: 0.0087), $\beta = 0.9087$ (SE: 0.0186) indeed minimize the loss function.

amounts based on the website of the Hungarian State Treasury.

[17]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 different preferences.

Table 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 8 compares non-targeted empirical and simulated moments. We can see that despite the estimation not directly matching 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 result from the independence assumption between income and the ideal number of children (indeed, these families might have, on average lower ideal number of children 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 higher overall completed fertility effect of the policy than the main estimates; however, the magnitudes of the two exercises are reasonably close.

Table 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%.

7.3 Policy scenarios

In this section, I use the estimated model to examine the 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[18]

[18] The original version of the policy had 62,500 HUF deduction for families with less than two children, and 206,250

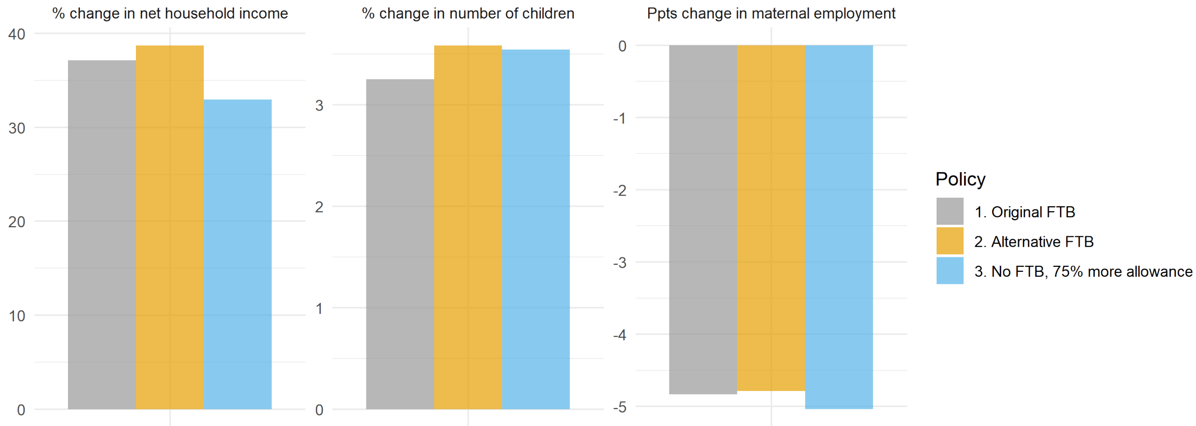
- No family tax break, but a 75% increase of the family allowance

The exercise aims to see whether the model is consistent with the estimated long-run fertility effects to add predictions regarding long-run direct female employment effects. Additionally, it helps to identify which income groups react to the policy according to the model and how alternative, more straightforward policies would fare. While the 'Alternative family tax break' simulates a simpler version keeping the incentives to work, the increase of the family allowance decreases the cost of children. I also group households into three income categories of 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 look at redistributive effects.

Figure 18 shows the aggregate changes due to different policies compared to the baseline in the number of children (%), net household income (%), and maternal employment (in percentage points). We can see that the model suggests a 3.25% increase in the number of children due to the original family tax break policy, along with a 4.8 percentage points lower maternal employment. The change in completed fertility is greater but reasonably close to what the reduced form estimates indicate, yielding around 2.36%, taking the sample size weighted sum of the two estimates. The alternative policies suggested are comparable in government expenditure (affected households end up with around 30-35% more net income), so they can be considered good comparisons. The alternative 'flat' family tax break would lead to slightly higher fertility (3.59%), with a similar effect on maternal employment rate (-4.8%) with approximately the same change in net income, while the no tax break scenario results in higher fertility effects (3.55%), slightly lower maternal employment (-5%), but around 5% less government expenditure.

HUF for families with at least two children.

Figure 18: Policy scenarios: aggregate changes in outcomes compared to the no-policy baseline

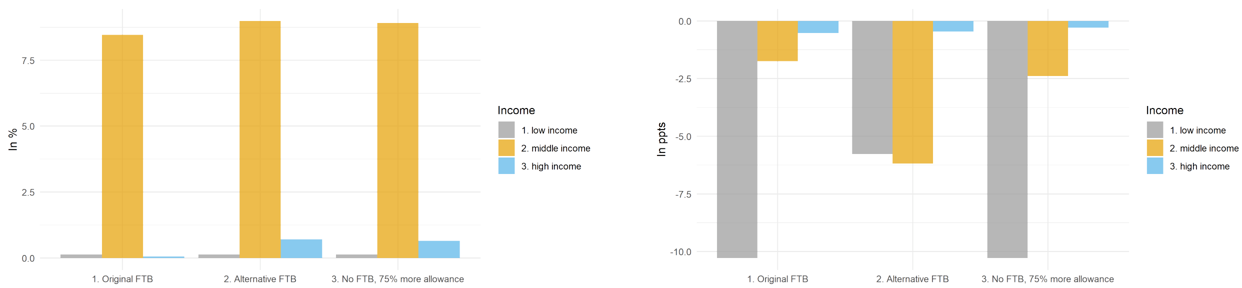


Note: the figure reports the model simulations' aggregate outcomes under different policies of the original family tax break (FTB), the alternative family tax break, and the no tax break scenario with elevated family allowance.

Figure 19 shows the effect of the different policy scenarios on the number of children and maternal employment by 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. In contrast, the impact on the two other groups is minimal. Both alternative policies result in higher fertility for middle and high-income families leaving low-income families unaffected. The effect on maternal employment, however, is different. While the original family tax break and the elevated family allowance decrease maternal employment of the lower-income families by around 10% (which by design leads to higher 'quality' of children), middle and higher-income families are only slightly affected. 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.

Figure 19: Policy effects compared to the baseline by income groups

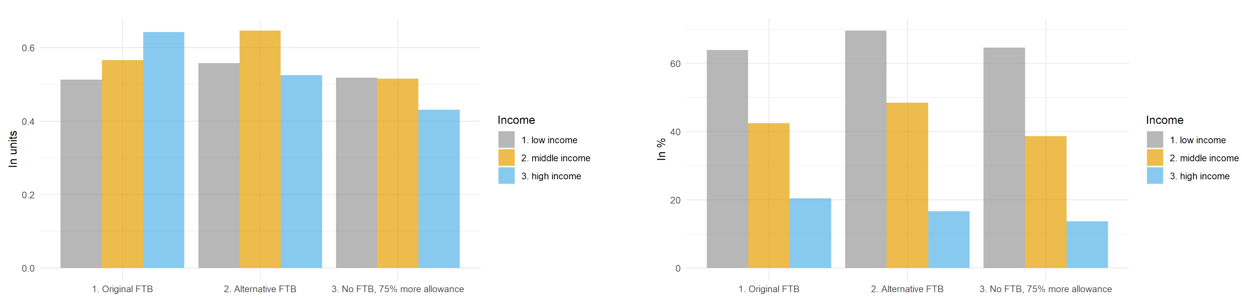
(a) Number of children (b) Maternal employment



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 20 reports the change in net income on the left-hand side in absolute terms (units of poverty line consumption), while it is expressed in percentage terms on the left-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 policies enrich middle-income families the most while providing the least to high-income households. In percentage terms, all the examined policies give more to the poorer households because they start with a lower income base.

Figure 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%.

8 Discussion

In this section, I put the findings into perspective: discuss the internal and external validity and place them into an international context based on 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, we find no statistically significant estimates indicating 'placebo' effects. Further robustness checks corroborate these remarks, as different measurements and sample restrictions yield qualitatively similar results. For the results of families with less 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 of the results.

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 estimates presented here suggest that for families with at least two children, the policy resulted in an around 0.76% increase in completed fertility. A less robust, around 5.64%

effect is identified for families with less than two children. Regarding maternal employment, there is no evidence of any effects overall; however, there seem to be some segments of the population that reacted more sensitively to the policy. For an overall policy effect, we could consider the sample size weighted sum of the two effects, yielding a value of around 2.36% increase in completed fertility. With all the precautions in mind about the results, the out-of-sample model simulation produces a somewhat higher, around 3.23% increase in completed fertility, supporting the validity of the magnitudes presented here. Additionally, the model predicts an around 4.8 percentage point lower level of maternal employment due to the policy.

Table 9 puts the findings of this paper into context by collecting the findings of some relevant papers in the literature. The results are lower than the short-run effects in the literature, consistent with the estimates not simultaneously capturing the timing effects. However, they are quite higher than the previously completed fertility estimates. One explanation could be that I examine a country where the additional resources might make a difference in the life of a family, the treatment size is large, and the policy is a recurring 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 in light of the changed tax incentives. The comparative statics exercise presented here suggests a larger impact on completed fertility. As the last point, I mention that the magnitudes presented here are in line with previous aggregate or quasi-panel results for Hungary (Gábos et al., 2009; Bördős and Szabó-Morvai, 2021).

Table 9: Comparable results in the literature

Countries	Studies	Year	Policy	Size	Short-run	Completed fertility
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.

9 Conclusion

In this paper, I study the effects of tax incentives on completed fertility (and to a lesser degree on maternal employment) 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 at 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 study how additional prospective income for having additional children affects the completed fertility of women.

I find that the policy led to 0.76% higher completed fertility for families with at least two children and 5.64% for those with less than two children (weighted average of around 2.4%). The findings are driven by religious, married couples where the age of the youngest child is less than three. Results regarding maternal employment are not significant overall. I also fit a simple partial equilibrium model of household behavior to simulate counterfactual policies, and to examine possible longer-run effects regarding maternal employment, completed fertility, and income redistribution. The model, among other results, suggests around 3.23% higher completed fertility (larger than the main estimates but close in magnitude) and around 4.8 percentage points lower employment attributed to the policy.

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.

ORCID

Bence Szabó ORCID: 0000-0002-8959-8912

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.

A Appendix

Table A1: 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.

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

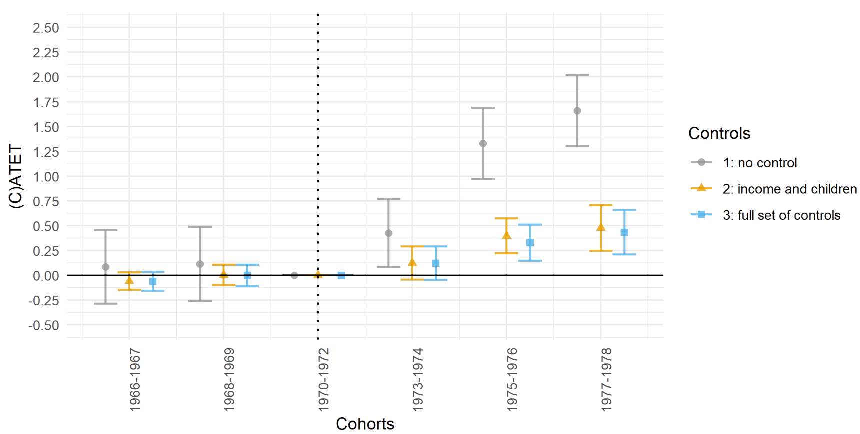


Figure A2: Fertility: families with less than two children, unweighted

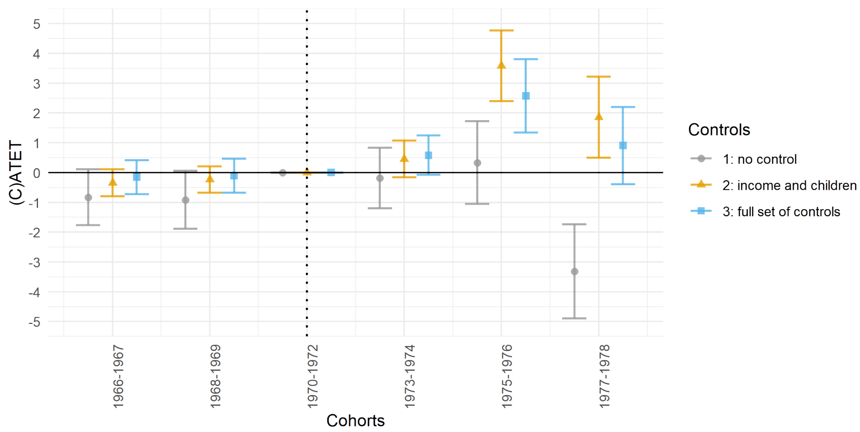


Figure A3: Fertility: families with at least two children, counts

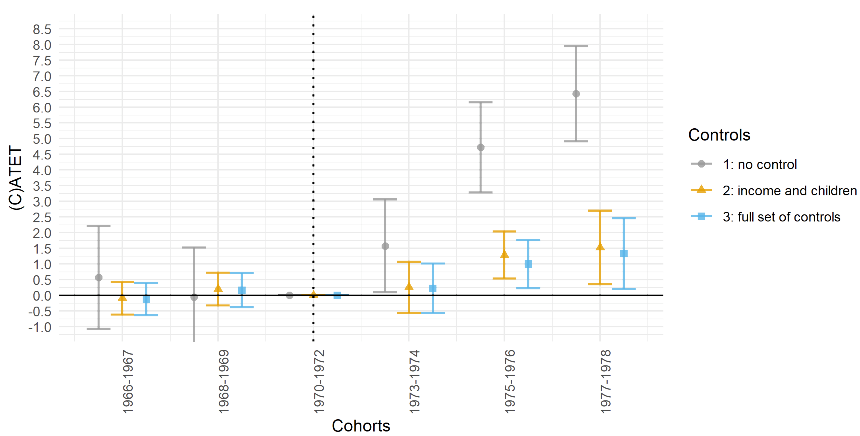


Figure A4: Fertility: families with less than two children, counts

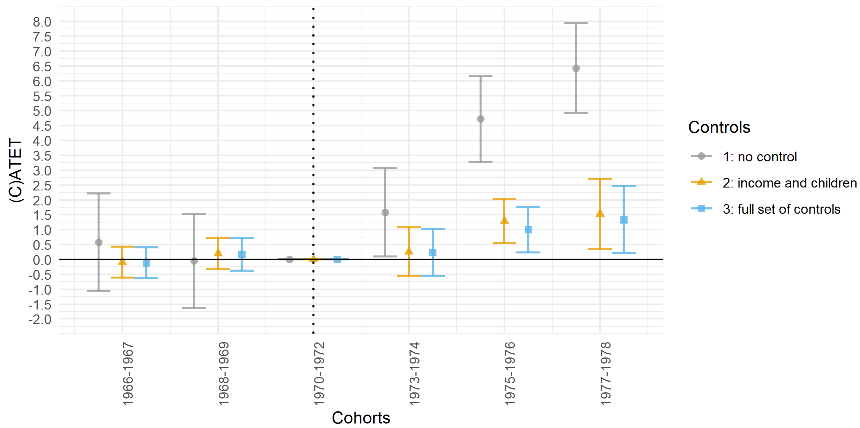


Figure A5: Fertility: families with at least two children, other spec.

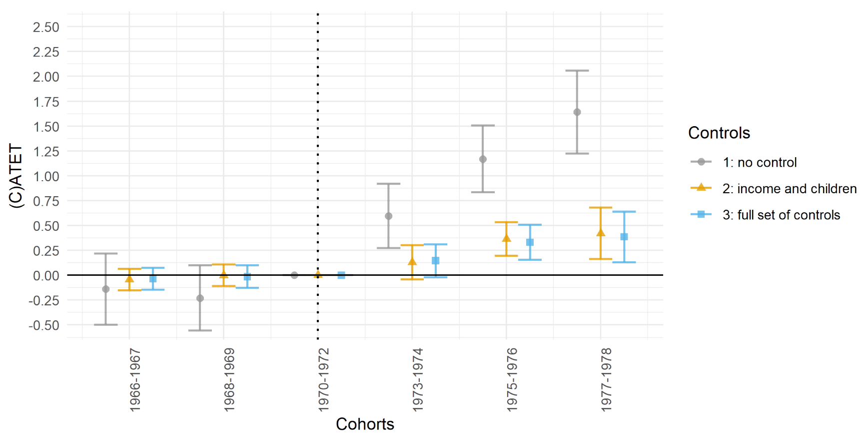


Figure A6: Fertility: families with less than two children, other spec.

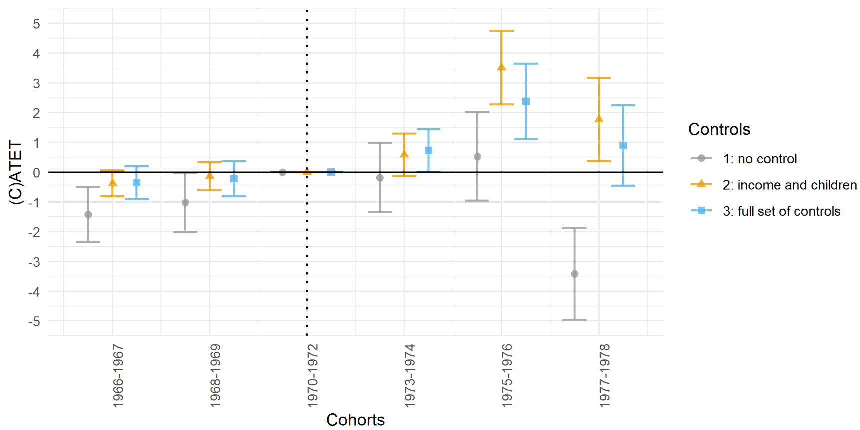


Figure A7: Fertility: families with at least two children, discrete

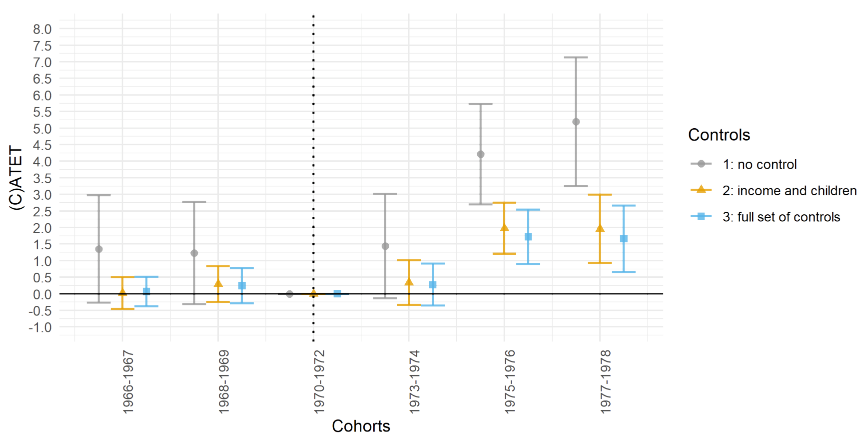


Figure A8: Fertility: families with less than two children, discrete

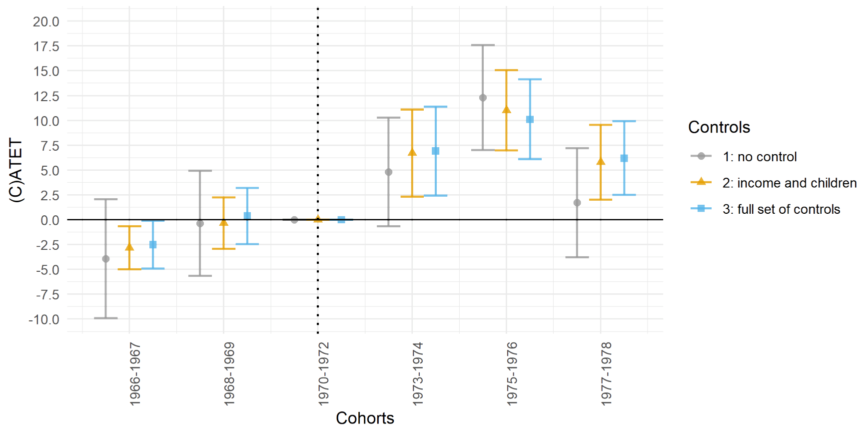


Figure A9: Male employment: families with at least two children

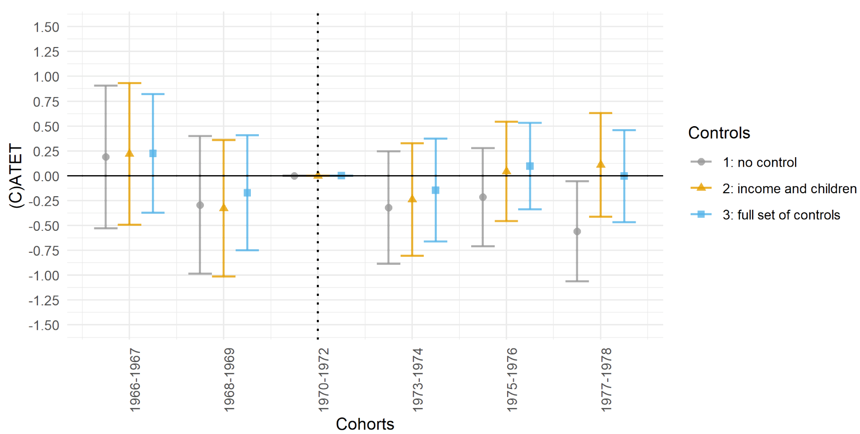


Figure A10: Male employment: families with less than two children

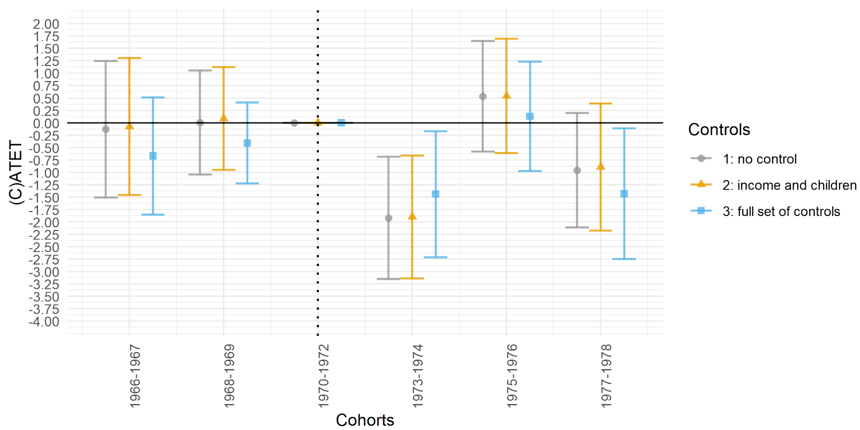


Figure A11: Male salary: families with at least two children

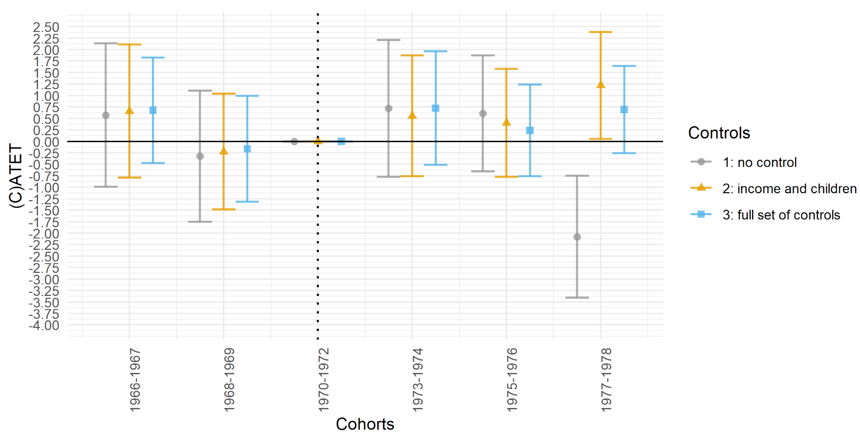


Figure A12: Male salary: families with less than two children

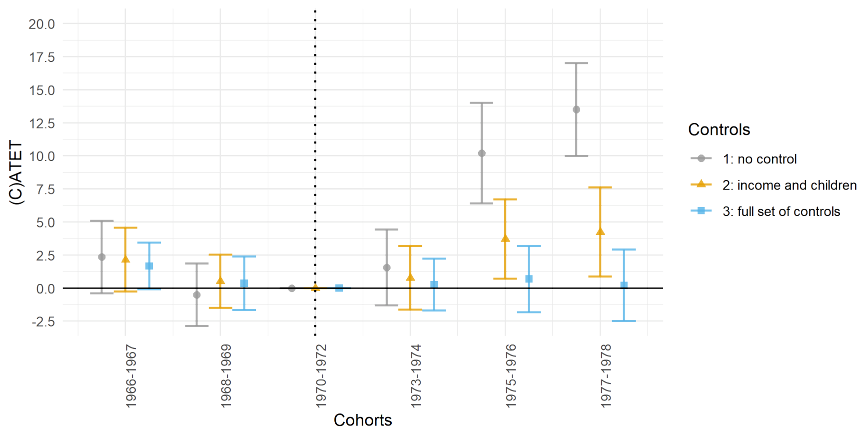


Figure A13: Female salary: families with at least two children

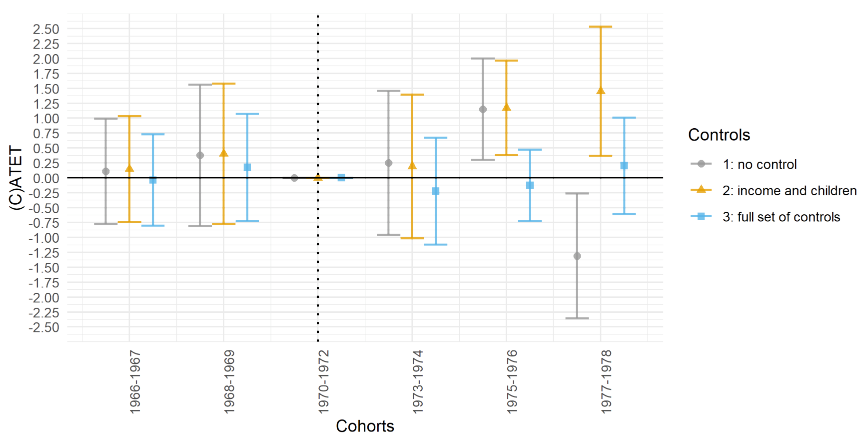


Figure A14: Female salary: families with less than two children

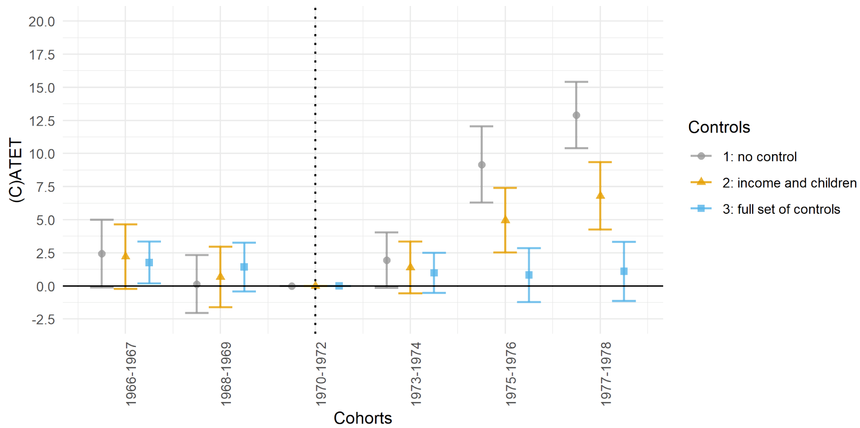


Figure A15: House size: families with at least two children

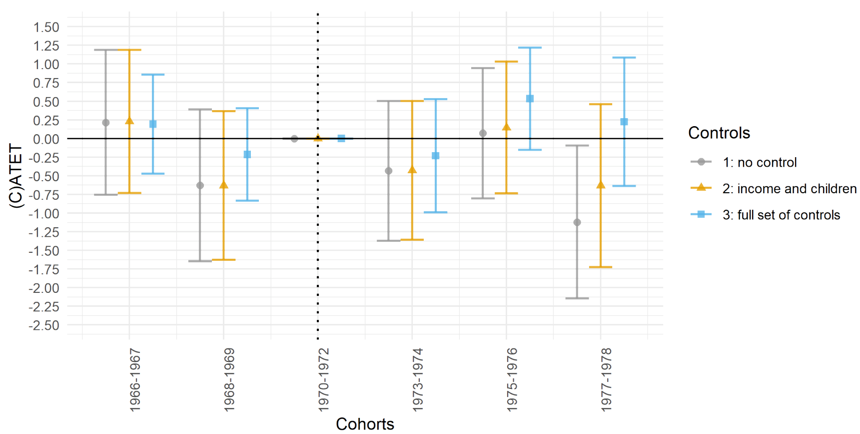


Figure A16: House size: families with less than two children

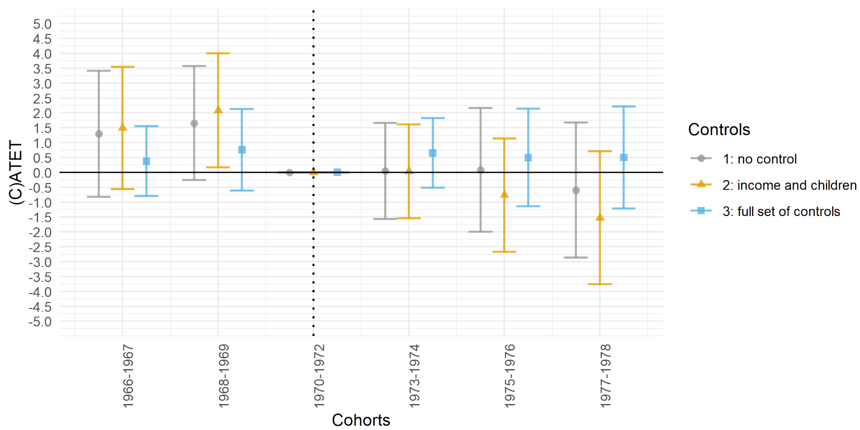
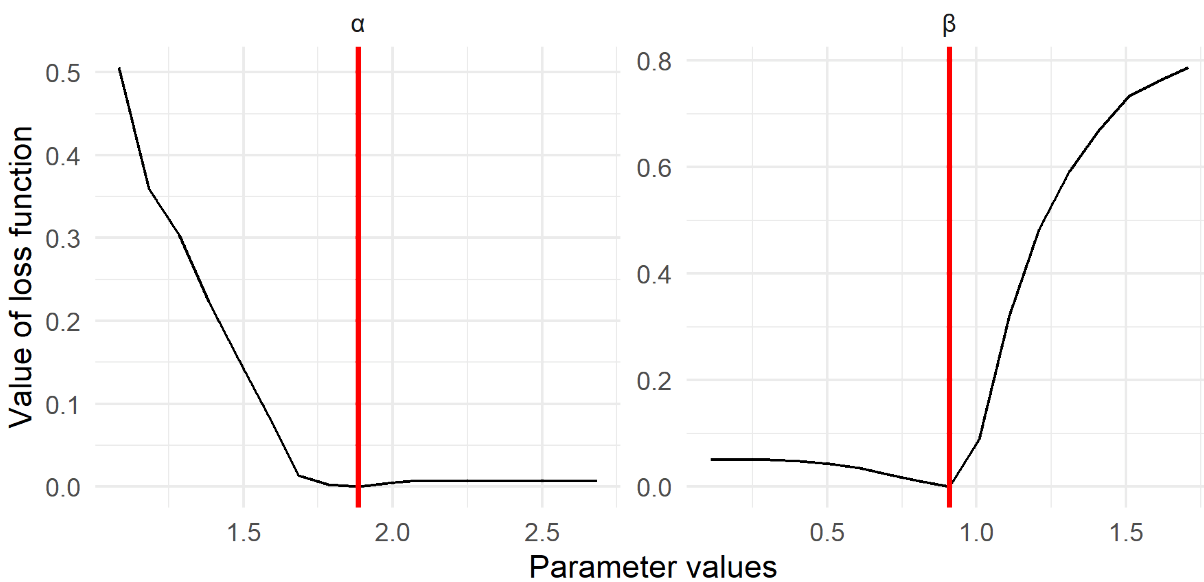


Figure A17: Loss function in the neighborhood of the estimated parameter values



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