# Tax Incentives and Completed Fertility*

Do not cite without the author's permission.

Bence Szabó[1]

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

September 19, 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 major overhaul with 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, providing a rich set of individual parental pre-policy control variables, and observe the change in the number of children within families. To identify the reform effect, I use variation along the initial number of children in 2011 and prospective household income, while comparing cohorts just before and after the end of the fertility cycle. I argue that the policy induced 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 by around 5.6% for families with less than two. The findings are driven by religious and married couples with a young child. No significant overall 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 to find a qualitatively similar completed fertility effect, with an around 4.8 percentage points decrease in maternal employment rate due to the policy.

*Keywords*: Completed Fertility, Family Tax Break

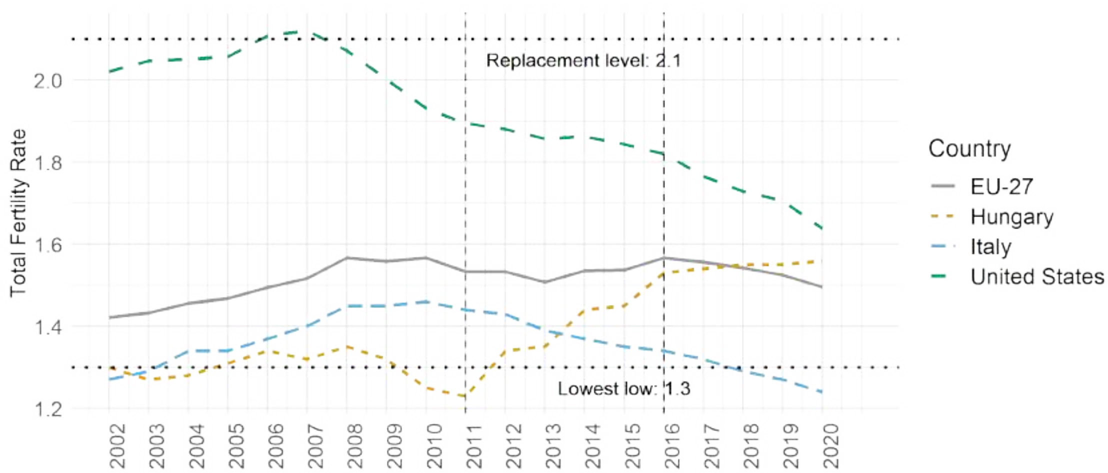
*JEL Classification*: J13, J18, H24

*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 to Guido Tabellini, Marco Di Maggio, Gábor Kertesi, János Köllő, Gábor Békés, Hedvig Horváth, Maria Sironi, Thomas Le Barbanchon, Blanka Imre, Bálint Ván, and Jacopo Lunghi. I am also grateful for the useful 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, and 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

One of the central policy issues in several developed countries is persistent low fertility. Since the 1970s, total fertility rates[1] in OECD countries have declined from around 2.84 to 1.61. Total fertility rate has remained more or less stable during the last two decades, below the population replacement level of 2.1, and reaching 'lowest-low' levels of below 1.3 in some countries of Southern, Central and Eastern Europe, at which birth cohorts would shrink to half in around 50 years (Kohler et al., 2006). In this paper I focus on one of these countries, Hungary, reaching its minimum total fertility rate of around 1.23 in 2011 (left panel of Figure 1). Countries with such fertility rates and without substantial migration are projected to experience worsening living standards in the long run (Lee and Mason, 2014), despite lower than replacement level fertility could be beneficial regarding consumption per capita.

Figure 1: Evolution of Total Fertility Rate in Hungary and comparable countries



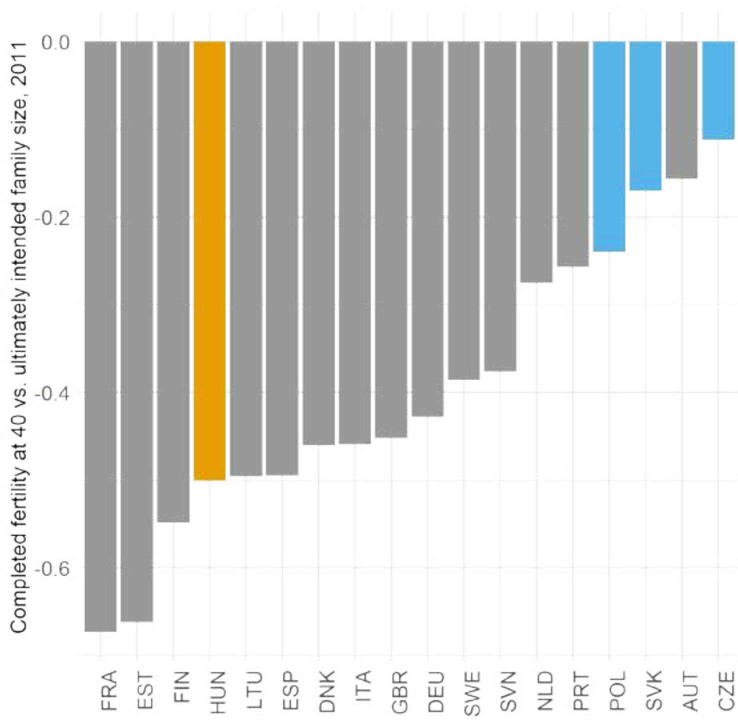

*Note:* Based on the data of the World Bank.

Financial incentives might be of limited use if fertility outcomes are predominantly driven by household preferences. The demographic literature has identified several other drivers (Billari and Kohler, 2004; Frejka and Sardon, 2004; Billari et al., 2009; Goldstein et al., 2009; Esping-Andersen and Billari, 2015; Greenwood et al., 2017; Zeman et al., 2018; Seltzer, 2019), such as the shifting

[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 totalling the age-specific fertility rates as 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

societal role of women in the household, or their increased labor force participation.[2] However, as Figure 2 shows, a comparison of the ultimately intended number of children and completed fertility from 2011 suggests that in many countries the realized number of children at the end of the fertility cycle lagged behind the intentions, especially in Hungary compared to countries with similar background such as Czechia, Poland or Slovakia. Although measuring fertility preferences is difficult (Spéder and Kapitány, 2009), there is evidence that in 2011 it was reasonable to believe that the government could intervene successfully. More globally speaking, for countries with low fertility and limited migration it is important to understand the role taxation policies can play as fertility incentives.

Figure 2: Cross-country relation of completed fertility and ultimately intended family size




*Note:* Based on the data of the Eurobarometer, and the Human Fertility Database.

In this paper I examine the causal effect of taxation related financial incentives on completed fertility, by utilizing the family tax break reform of a low-fertility developed country, Hungary. The Hungarian Government introduced a large-scale expansion of family tax breaks in 2011, increasing its magnitude from 0.05% of the GDP to 0.64% of the GDP, around 17% percent of all public

[2] Although the cross-country correlation between labor force participation and fertility has changed over time from negative to positive (Kohler et al., 2006).

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 (quantum effects), without mixing the timing effects into the estimates. To achieve this, I construct a novel 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 this prospective household income, I calculate the financial incentives introduced by the policy, and in a difference-in-differences design I 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 the same change on average as it did for those that could still increase their number of children. Along with completed fertility, I also examine maternal employment while exploring the heterogeneity of the results, and build a simple partial equilibrium structural model to study counterfactual scenarios and long-run effects.

The findings suggest that the policy increased completed fertility for families with at least two children by around 0.76%, and by around 5.64% for those with less than two children, the sample size 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 maternal employment are not significant overall, however, there seem to be some negative employment effects for those segments where completed fertility increased. 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 model indicates an around 3.25% higher completed fertility, and around 4.8 percentage points lower employment due to the policy.

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), traditionally modeled jointly as part of the production of market and household goods, while trading-off the quantity and quality of children. There is a large literature on the microeconomic theory of tax policies affecting fertility studying various setups (Cigno, 1986, 2001; Fraser, 2001; Balestrino and Cigno, 2002; Cigno and Pettini, 2002; Docquier, 2004; Meier and Wrede, 2013; Baudin et al., 2015), with predictions about optimal policies as a function of underlying parameters.

While theoretical discussions focus on completed fertility, 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 in order to receive the benefits earlier, but 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 fertility decisions and career choices of women to differentiate between short-run and long-run consequences, and finds that relatively large short-run effects of a pro-fertility policy 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 due to gradual demographic change the examined subsets or cells within the population might alter in composition, so the generally small positive fertility effects estimated by this literature still leaves 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, but similarly to the quasi-panel evidence these papers also identify short-run treatment effects mostly with innate limitations to address timing vs. completed fertility in a convincing manner. 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 demographic economics, public economics, and to a smaller extent labor economics the following way. I take advantage of the unique structure of a tax policy targeting fertility, enabling causal identification of the quantum effect of a large-scale government intervention in a developed country, avoiding the problem of usual short-run estimates that might jointly capture timing effect as well. The constructed novel family-level linked dataset provides individual-level evidence with rich pre- and post-policy observable characteristics enabling to study the heterogeneity of the treatment effects. 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, and compare relevant policies taking into account the redistributive impact as well.

The structure of the paper is as follows. I introduce the context of the family tax break extension, and describe the details of policy in Section 2. In Section 3, I present the construction of the analysis dataset. Section 4 presents the empirical strategy for the identification of the reduced form results reported in Section 5. In Section 6 follows the building and estimation of a simple household model examining different counterfactual scenarios, while Section 7 discusses the findings, and Section 8 concludes.

## 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 shrank, primarily due to the low fertility rate (Őri and Spéder, 2020). In this period, fertility behavior can be characterized by a postponement of the birth of the first child, 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), while slowly for younger cohorts childlessness has also started to become more prevalent (Spéder, 2021).

Although there were some smaller sized interventions 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 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 both the circle of eligibility, and the amount of the available subsidy. Since then, families with less than three children also became eligible to a smaller degree, but the tax break includes a large jump at 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, but it played an instrumental part resulting in major redistribution towards higher income families with more children (Tóth G. and Virovácz, 2013; Varga et al., 2020).

The policy instrument itself is a non-refundable tax base deduction that can be exercised by either one of the parents or jointly regardless of marital status. However, in its initial version it is implicit that the tax base deduction cannot exceed the available taxable income of the family[3]. This results in profound consequences regarding the tax break change for an additional child based on the income of the family. 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) compared to 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, 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[4]. In case the

[3]Note that since 2014, families not being able to claim the full 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.

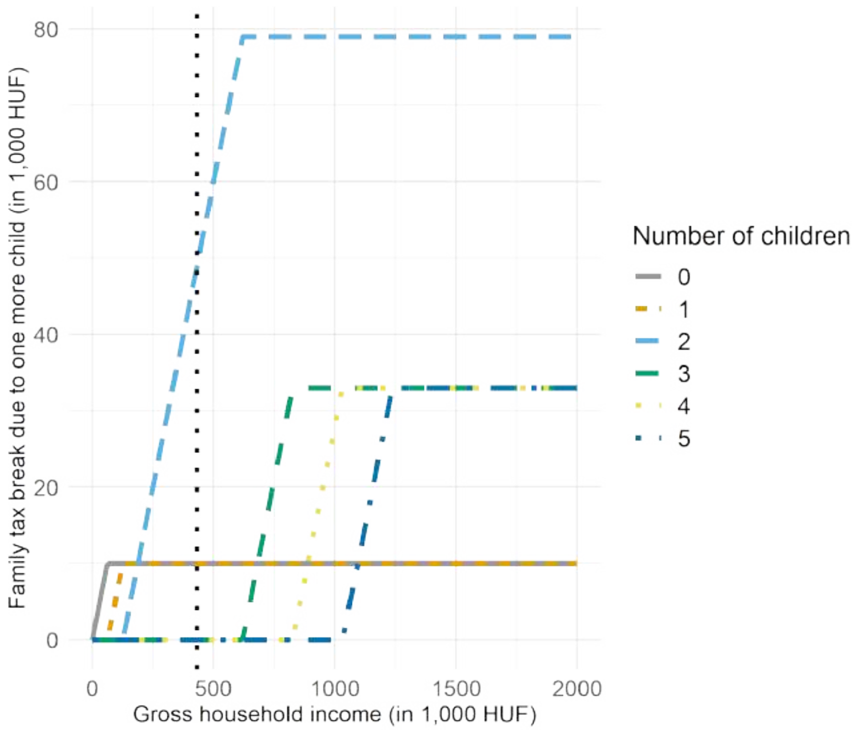
[4]Source: Hungarian Central Statistical Office, I use a 270 HUF/EUR exchange rate which approximates well the

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 would not be able to claim the full amount of the tax break with three children, in fact a family would need 45% more than the average income for that.

Figure 3 displays the additional money a family would receive after having one more child (left panel), and the baseline amount of tax break they would receive for their initial number of children (right panel), against household income by the number of children, with the vertical line indicating 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 an around 80,000 HUF monthly increment if income is high enough, increasing at the slope of 0.16 with income earlier 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.

2011 value

Figure 3: Additional family tax break for one more child




*Note:* the figure shows that amount of family tax break families would get for an 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, and residency status. 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 are required to create a new identifier that we can match the observations on. This is based on the male and female self-identifying as the leader of the household, and spouse of the leader of the household, allowing for both genders to have each role, and allowing for the inclusion of non-married couples. This restriction inherently puts some constraints on the external validity of the results as couple formation cannot be taken into account, however a rich set of control variables are available from the pre-policy period unaffected by the policy from 2011 that might mitigate some of the selection issues.

I excluded from the analysis those individuals who are institutionalized, and more importantly those families where either the female, or the male adult is missing. This means that single-parent households are not considered, as the chance of these households to be found in the 2011 Census would be very small in any case. Furthermore, I restricted the sample to those families where the mother's birth year is 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, while it also includes for additional reference families which 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 am going to 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.[5]

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 a conservative one as it does not utilize stochastic matching algorithms (or more advanced ones such as Abramitzky et al., 2020) or any other possibly more endogenous variables, however other than data errors we can be quite confident that the families found in both Census datasets are indeed the same ones. Table 1 shows the number of family units after the different steps of sample restrictions: we end up with around 54,000 families (representing around 680,000 families) in the 2016 Microcensus that I searched for in the set of around 597,000 families in the 2011 Census.

[5]The fraction of births to mothers forty or more years old is 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 for 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 around 58% could be matched successfully, ranging from 56-59% in different cohorts, so the variance is quite small across cohorts. Table 1 shows the raw number of observations participating in the matching process and that the post-matching data cleaning eliminating potentially erroneous observations and outliers, after which we end up with a final number of observations of around 31,000 families.[6]

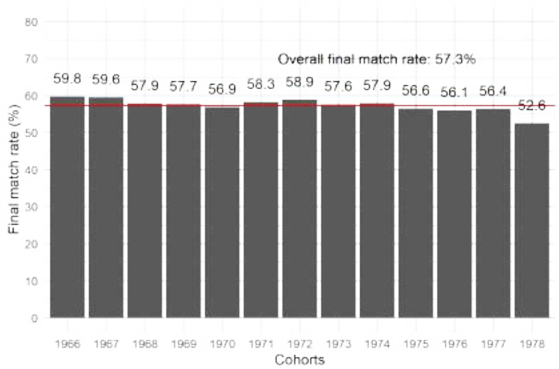
Linking families existing in 2016 to themselves in 2011 requires that the family unit was observed in both censuses. This is not possible if the family unit did not exist earlier, or it was not enumerated because of not being present in the country due to emigration for instance, or they were just not cohabiting officially. This inherently induces a type of sample selection limiting external validity of the findings that can hardly be mitigated. The right panel of Figure 4 helps to give an idea which families were more matched more successfully. In the regression the following predictors were included: mother's birth cohort, number of children, marital status, parental employment status, imputed parental log(salaries), parental education, parental ethnicity, father's birth year, home size. The figure shows the statistically significant predictors at the 5% level, suggesting that the most important factor in finding a family is marital status: married couples given other predictors are around 25% point 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 are all correlated positively with successful matching. This should come as no surprise:

[6]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 and at the individual level, while in my case I match on couples in recent sources which is expected to yield much better results.

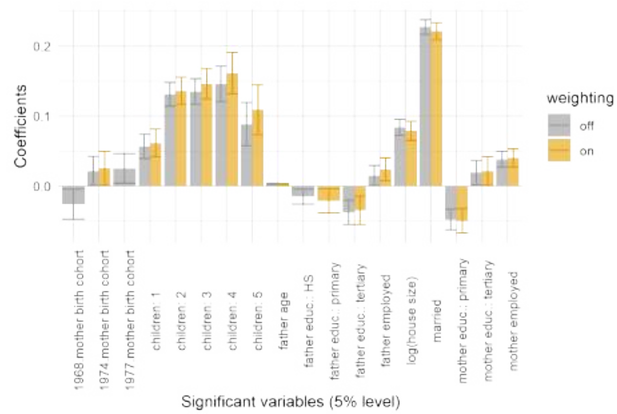
the families described by this constellation are the ones that might be intuitively more stable and longer-lasting. This exercise suggests that when we interpret the findings we should keep in mind that validity might be limited to those couples that are intrinsically more stable.

Figure 4: Matching the 2016 Microcensus and the 2011 Census

(a) Matching success rate by maternal birth cohorts




(b) Predictors of matching success




*Note:* the author's calculations 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 number of children, marital status, parental employment status, imputed parental log(salaries), maternal education, father's birth year. Variables displayed are significant at the 5% level.

## 3.2 Imputation of prospective salaries

To create the treatment variables, I attached to each member of the couple the prospective gross salary estimated using National Wage Survey data of Hungary, which collects salary information on the private and public sector every year since the 1990s, as census data do not contain wage information. However, for most individuals a rich set of work-related observable characteristics is available enabling a fine distribution of imputed salary income at the family level. Prospective income proxies life-cycle earnings of the members of the household, which in the fertility decisions related to the tax break should play a more instrumental role than current earnings, as the tax benefits provides yields for a large portion of the life-cycle, as long as the family receives Family Allowance, 20 years of age the latest (Makay, 2015). Another advantage is that personal unobserved earning capabilities (potentially correlating with fertility decisions as well) within the same subset of people cannot play a role of inducing bias in the estimates, and that currently unemployed individuals can also receive a prospective salary.

I imputed salaries for 2011 by pooling the 2010 and 2011 Wage Survey data, and similarly for

2016 I pooled data from 2015 and 2016, which I refer to as prospective gross salary as I use the subsample of prime-earning age individuals (30-49 year-olds), so I do not match on age of the individuals. I used three sets of variables to match:

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 the following (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, while gender, region municipality type, and education information is 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. Appendix Figure A1 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, and that these distributions do not differ drastically between the cohorts that are at least 40 years old in 2011 and those that are 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% |

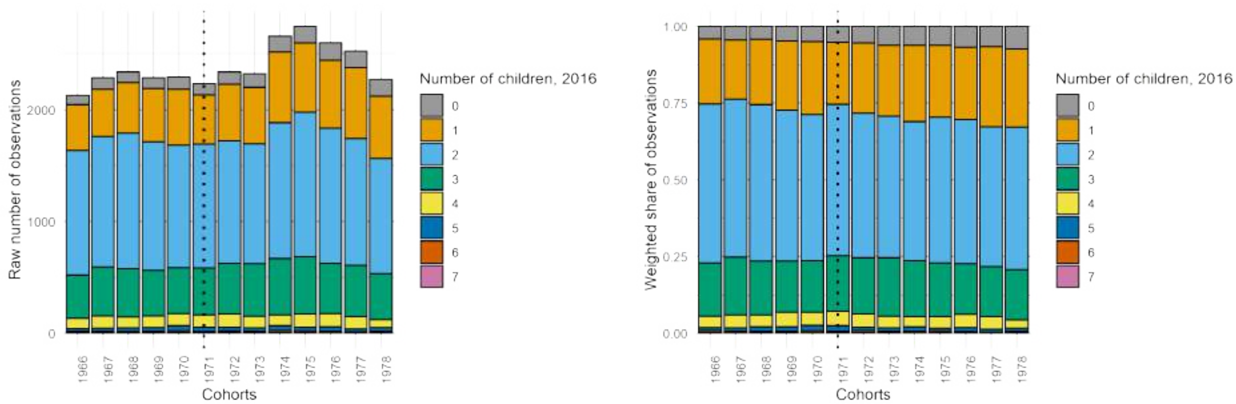
*Note:* The table reports the success rate of joining average salaries along different sets of imputation variables, for the mothers, and for the fathers, for the 2016 Microcensus.

## 3.3 Descriptive statistics

This section gives an overview of the final analysis sample, shows the composition regarding cohorts and the number of children, the imputed prospective household incomes, and the change in value of the outcome variables over time. Figure 5 displays the unweighted and weighted number of observations, and their distribution by the number of children in 2016. For the relevant cohorts, the age of observing them spans from 38 to 50 years, so we can treat all except for the 1977 and 1978 cohorts as around the end of their fertility cycle. We can see that the weights do not alter the composition of the sample substantially. There is also a steady increase in the share of childless couples, which process has continued for younger cohorts not studied in this paper as well, as documented by Spéder (2021). Otherwise for the other birth parities the distribution seems to be similar.

Figure 5: Observations in the analysis sample by number of children in 2016

(a) Unweighted number of observations (b) Shares using the 2016 Microcensus weights



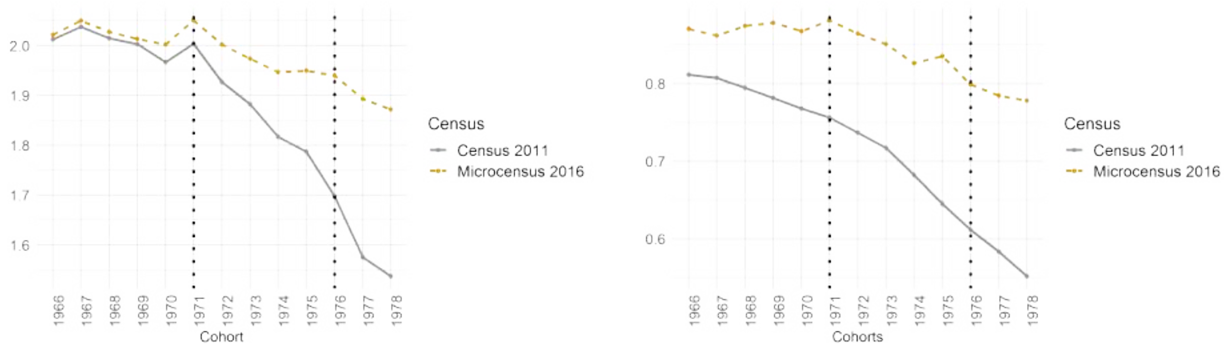

*Note:* the author's calculations of the number of families in the sample, based on the matched dataset of the 2016 Microcensus and the 2011 Census of Hungary.

In this paper I am going to focus on two outcome variables: completed fertility, and to a lesser degree, 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 effects along these aspects as well with additional identifying assumptions. I also inspect the effects on paternal employment and house size as well as validation exercises.

Figure 6 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 was born during these five years, which is around 13% of completed fertility. Shares of employed amongst the matched sample increased for both the fathers and the mothers in these time periods, however, for the latter the relative change is substantially larger. Maternal employment rates also suggest that the 1971 cohort have already transitioned back to the labor market 2016, as their employment rate matches the ones of older cohorts, while in 2011 this is clearly not the case.

Figure 6: Outcome variables in the analysis sample

(a) Number of Children (b) Female Employment




*Note:* the author's calculations based on the matched dataset of the 2016 Microcensus and the 2011 Census of Hungary.

# 4 Empirical strategy

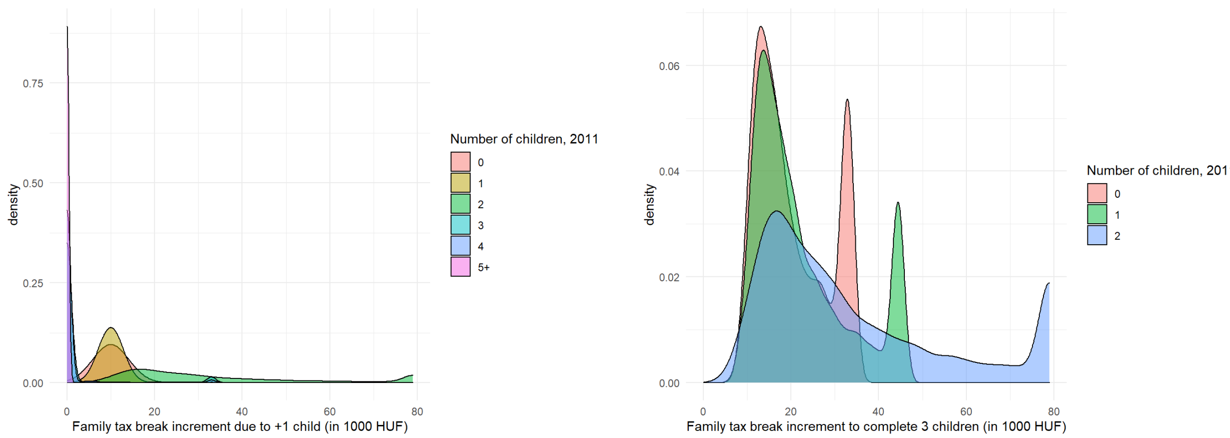
## 4.1 Treatment definitions

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 household gross income. The policy introduces non-linearities as shown earlier (Figure 3) allowing us to examine the treatment effects while being able to control linearly for both income and the number of children, due to additional remaining variation at the interaction as the tax break is non-refundable and eligibility is full. I argue that this remaining variation can be considered quasi-random, as families with similar levels of income but different initial number of children receive differential incentives, and the same occurs to families with the same number of initial children but different levels of income. 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 for families with less than two children to complete three children due to the large jump.

The more straightforward treatment variable we can construct is the family tax break increment for one additional child, as it lowers the relative cost of one additional child. In this setting, most of the variation in the data comes for 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, as ex ante it might occur that a family decides to increase the number of children from zero or one to three instead of two as a response to the policy incentives. Figure 7 displays the estimated kernel densities for these two variables by the number of children. We can see in the left panel that indeed 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. In the right panel we can see that the family tax break increment follows the shape of household income densities by construction, however, due to the non-linearities in the structure of the policy there remains some residual variation even after controlling for the number of children and income.

Figure 7: Kernel densities of the family tax break variables




*Note:* the author's calculations 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 a 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 to number of children $N_i$ can be defined 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 which is limited in size by the tax base of the family, tax base being the sum of the gross salaries of the parents (who are not required to be married).[7] Based on this setup, I define the two types of continuous treatment variables targeting the increment effects, measuring how much additional money the family tax break policy would imply given the prospective household salaries. These are the following:

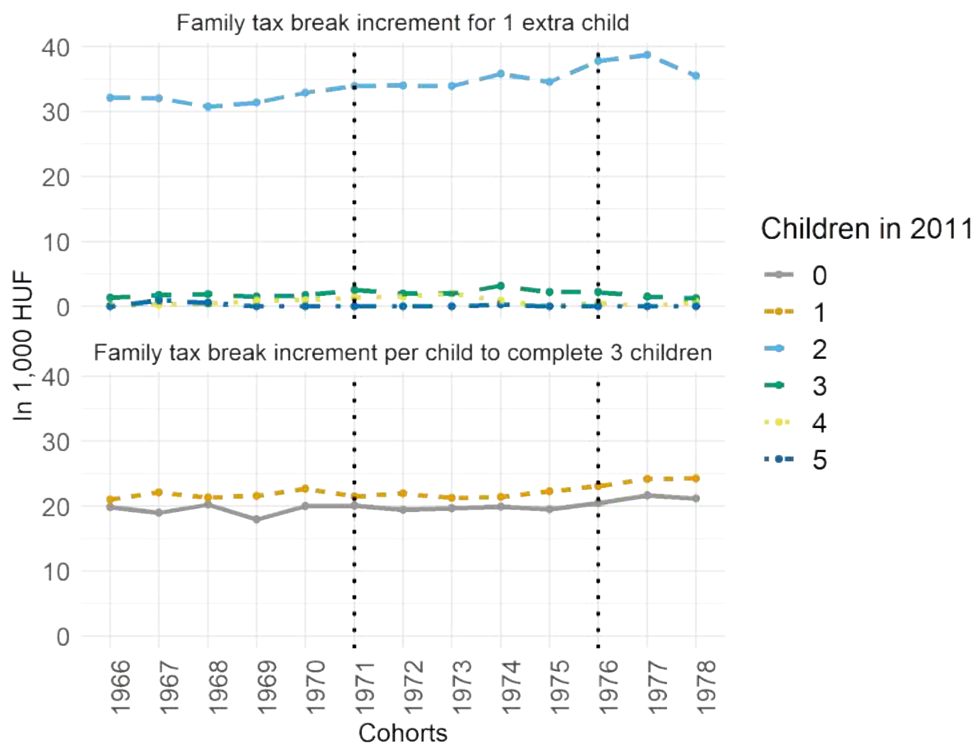
1. $T_{1,i} = M(N_{i,2011} + 1, S_{i,2011}) - M(N_{i,2011}, S_{i,2011})$: Additional money in 10,000 Hungarian Forints (HUF) per month for one additional child (examine separately also for families with at least two children, as it varies very little due to the low amount of deduction for zero and one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

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

Figure 8 shows the average values of the family tax break variables defined earlier, along with the imputed prospective household income, by cohorts and number of children. By definition the number of children and income jointly determines the amount of tax break an individual family would receive. This figure demonstrates that indeed there is no large variation in the treatments families received across cohorts, and that meaningful variation is expected to happen for the third and fourth birth parity.

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

Figure 8: Monthly average family tax break measures by cohorts and the number of children




*Note:* The figures shows 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 some external threshold, we can also introduce a discrete version of the treatment where we deem the treatment value 'large enough'. Poverty line estimates by the Hungarian Statistical Office indicates 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 children exceeds half of this level then the family is considered treated, which setup requires less stringent identification assumptions in a difference-in-difference estimation compared to the continuous version (Callaway et al., 2021). The discrete treatment results are used as supplementary robustness checks to the main ones, but we can conceptualize the treatment as the state financing half of the basic needs of an additional child.

### 4.2 Identification

The most important obstacle in identifying the treatment effect of such a policy is that families that receive more treatment might end 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. Fertility decisions should most reasonably be assumed to be connected to other life-cycle decisions such as educational, occupational, and residential choices, meaning that within a single cohort it is unlikely that there are no systematic differences in fertility between those that receive more tax break, compared to those that receive less (for instance, due to 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 a more adequate control group, which needs to be unable to react to the policy, while on average 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. I argue that a valid counterfactual can be constructed by using cohorts that are past fertile age, rendering them unable to react to the policy. To those cohorts the received family tax break environment acts as placebo regarding fertility: even though a similar treatment is received, the researcher knows ex ante that no effect can be expected. As mentioned, in my case a randomized experiment is not available so we cannot be certain whether there is unobserved, confounding heterogeneity biasing the within-cohort results. However, as 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 similar on average to the younger cohorts (conditional on several observable characteristics). Evidence indicates that fertility preferences remained stable throughout the years earlier to 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 should have to be the same 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 $T \in \Omega_T$ and controls $X \in \Omega_X$. Then the required assumption is the following:

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

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

Now suppose that the following linear equation governs the outcome, and let us add additively separable unobserved heterogeneity denoted by $\eta_{i,c}$. Then for cohort $c$ we get that

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

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 either be treated or untreated by the tax break, and unobserved heterogeneity $\eta$ represents the extra births from families with preferences for many children. If on expectation the extra births from such families is different in the treated group than it 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 for identification. Due to biological reasons we know that $\theta_{1966}$ (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 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 parallel 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 have a strong effect on the estimates. As even the 1971-72 cohorts should not respond strongly to the policy, instead of the 1966 cohort I select the 1970-1972 cohorts as the reference category for the regressions and use the other cohorts as pre-trend, while also grouping yearly cohorts to biannual ones to increase statistical power.

# 5 Results

I examine completed fertility (measured by the logarithm of the value + 1)[8] and maternal employment as relevant outcomes of the policy.[9] 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 also employ a rich set of control variables denoted by $X_i$ from the 2011 Census about the education, employment, health, and living conditions of the family, to eliminate any potential selection bias or composition change along these variables.[10] There are two treatment variables: one is the additional money received from one more child, which has meaningful variance for families with at least two children, and the additional 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 purpose was to include those 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.

[8]The results are robust to count-based regression specifications, as shown later.

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

[10]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.

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. The point of this exercise is 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 can we 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.

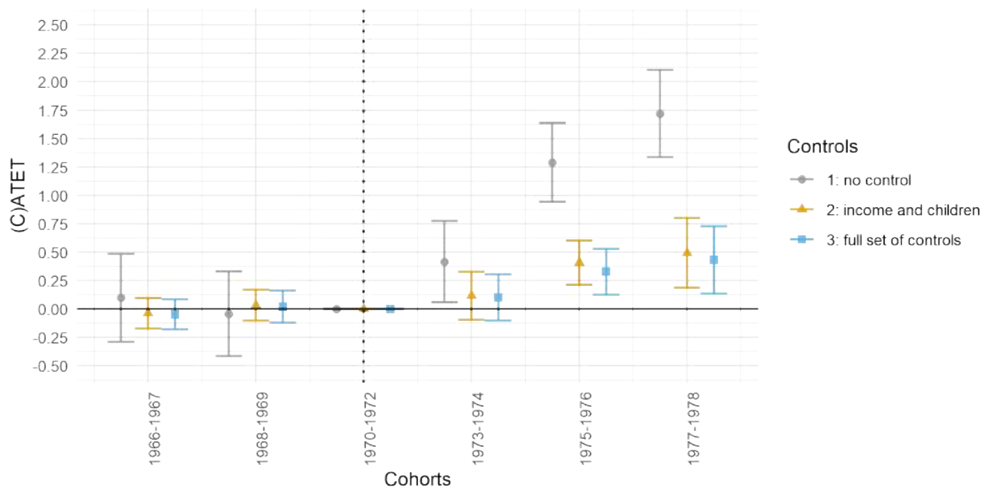
I present the estimates the following way. For each cohort, I display the point estimates of the interaction term with the 95% confidence interval bands calculated using cluster-robust standard errors on the sub-county administrative level.[11] For each cohort groups we can see three estimates, each of them corresponds 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 meaningfully react to the incentives and adjust their completed fertility.

### 5.1 Completed fertility

Figure 9 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 around 0.33 after including income and children. At the mean treatment size of around 2.29, the policy is implied to have an 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 even without controls for the older cohorts indicate that the assumption regarding unconfoundedness is supported, and we could indeed interpret the estimates for the fertile cohorts causally.

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

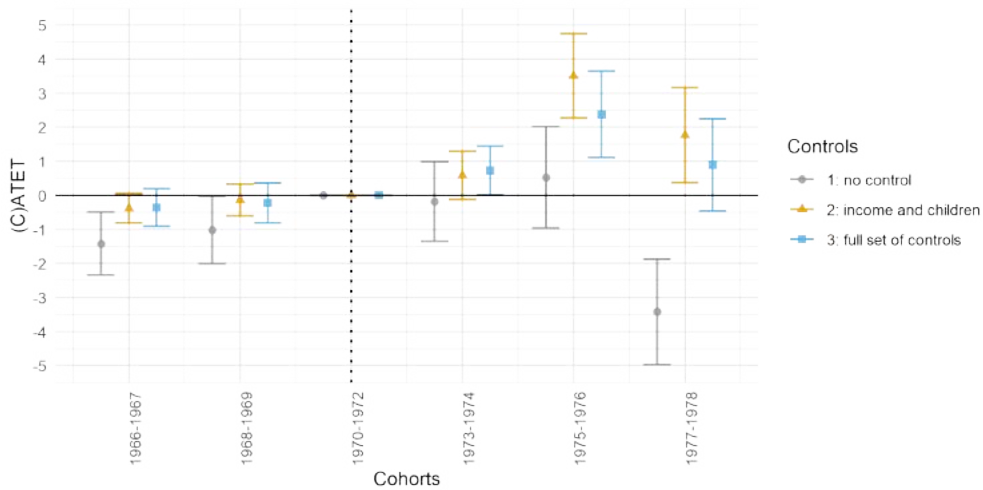
Figure 9: Effect of 10,000 HUF family tax break on completed fertility (in %), families with at least 2 children




*Note:* the author's calculations based on regressing the log(1+number of children) in 2016 on cohort groups interacted with treatments, with baseline of 1970-1972. Full set of controls include 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. Sample includes families with at least 2 children by the mother's age of 35, with sample size of 20,721. Mean treatment size is 2.29.

The results for the families with less than two children are reported by Figure 10. While we see no significant effects without controls along with some evidence for pre-trends at the older cohorts, after adding income and children as controls pre-trends turn insignificant while for the 1975-1976 cohort statistically significant effects are revealed. With a full set of controls 2.38% per 10,000 HUF is estimated, which at the mean treatment of 2.37 points to an around 5.64% effect at the lower birth parities.

Figure 10: Effect of 10,000 HUF family tax break on completed fertility (in %), families with less than 2 children




*Note:* the author's calculations based on regressing the log(1+number of children) in 2016 on cohort groups interacted with treatments, with baseline of 1970-1972. Full set of controls include 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. Sample includes families with less than 2 children by the mother's age of 35, with sample size of 10,243. Mean treatment size is 2.37.

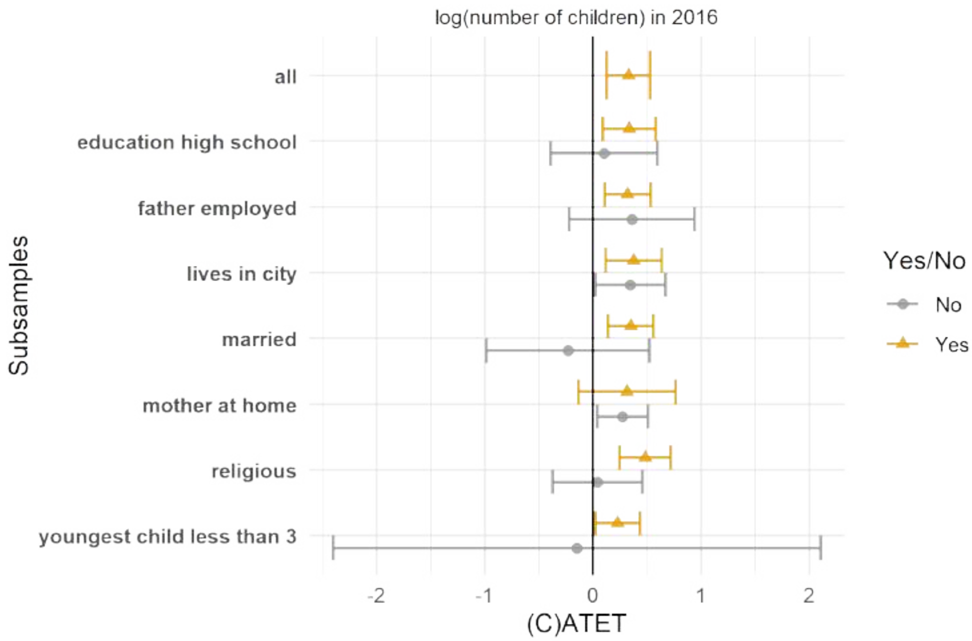
The heterogeneity of the results could also be of interest as we could learn which segments of the population are more responsive to financial incentives regarding their completed fertility. To address this question, I run the same regressions for several relevant subgroups[12], and I display the estimates for the 1976 cohort in Figure 11 and Figure 12. For families with at least two children, there are three aspects along which some differences can be seen. Although in all cases the confidence intervals overlap, the confidence bounds for religious, married couples, and where the youngest child is less than three years old are above the point estimate of their counterpart. One-sided Z-tests[13] would suggest that religiousness is significant at the 5% level, while marital status is so at the 10% level. For families with less than two children the picture is somewhat different (Figure 12). While the youngest child being less than three years old again seems to

[12] 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 employed status of the male parent the week previous to the Census survey with the ideal date of 01/10/2011. 'Lives in 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.

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

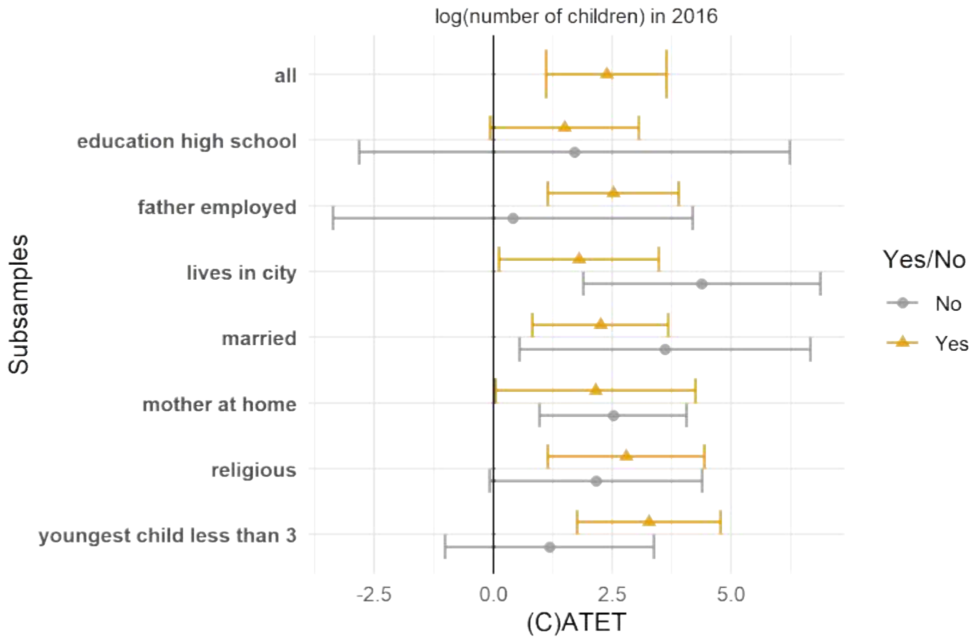
matter (one-sided Z-test indicating significance at the 10% level), no other variable plays a role creating heterogeneity in the effect of the policy.

Figure 11: Effect of 10,000 HUF family tax break on completed fertility for the 1975-1976 Cohorts (in %), families with at least 2 children




*Note:* the author's calculations based on regressing the log(1+number of children) in 2016 on cohort groups interacted with treatments for the 1975-1976 cohort, with baseline of 1970-1972 for different subsamples. Full set of controls include 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. Sample includes families with at least 2 children by the mother's age of 35. Sample and treatment sizes vary between subsamples.

Figure 12: Effect of 10,000 HUF family tax break on completed fertility for the 1975-1976 Cohorts (in %), families with at least 2 children



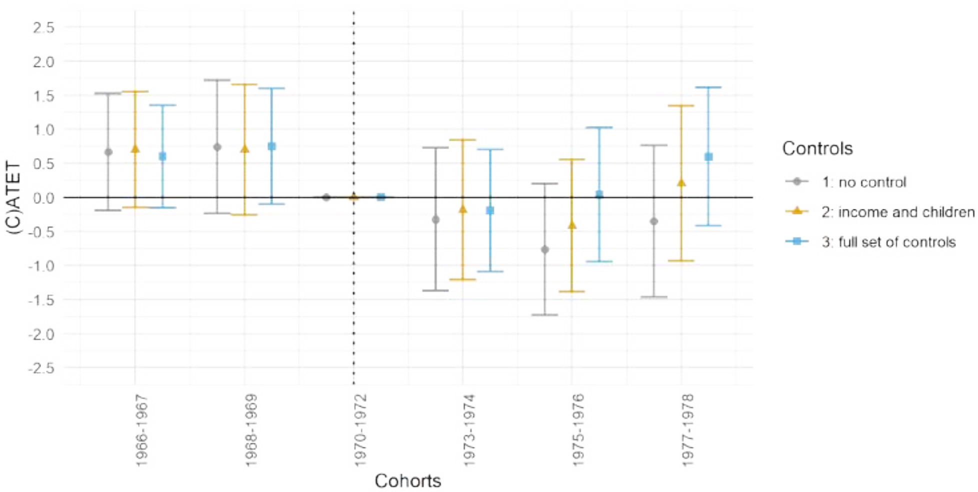

*Note:* the author's calculations based on regressing the log(1+number of children) in 2016 on cohort groups interacted with treatments for the 1975-1976 cohort, with baseline of 1970-1972 for different subsamples. Full set of controls include the following from 2011 besides household income and number 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. Sample includes families with less than 2 children by the mother's age of 35. Sample and treatment sizes vary between subsamples.

## 5.2 Mother's employment

As a second outcome we examine the employment status of the mother in the household. These results reflect a short-run labor market change that should be understood in the context of our previous findings on fertility. There are economic models describing how short-run and long-run labor supply decisions are intrinsically related to fertility choices of women due to heavy 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 we can use this as a validation exercise.

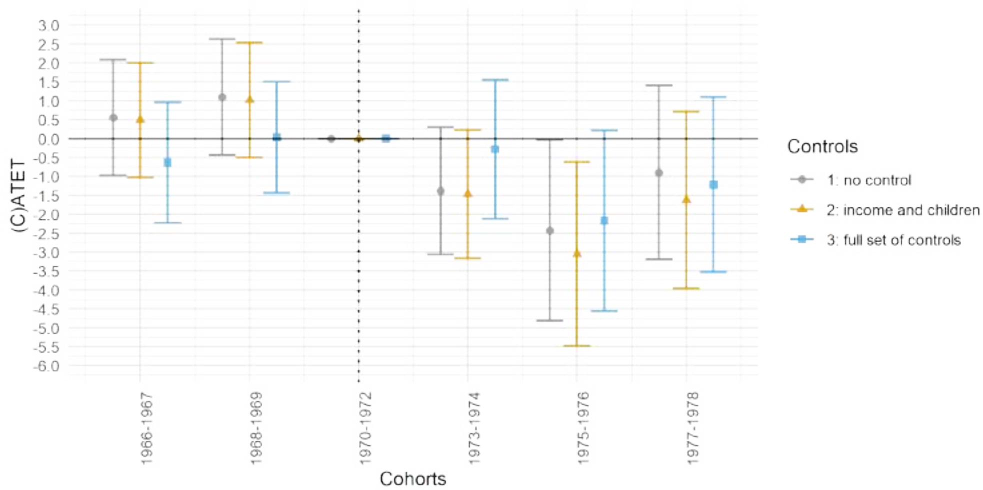
Figures 13 and 14 present the estimates for families with at least two, and less than two children respectively. We can see that no robust effects can be detected in either case, with some significant negative point estimates for families with less than two children if we do not include all controls. Reassuringly, the estimates for older cohorts are not statistically significant either, supporting the previous identifying assumptions.

Figure 13: Effect of 10,000 HUF family tax break on maternal employment (in percentage points), families with at least 2 children




*Note:* the author's calculations based on regressing the indicator for maternal employment in 2016 on cohort groups interacted with treatments, with baseline of 1970-1972. Full set of controls include 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. Sample includes families with at least 2 children by the mother's age of 35, with sample size of 20,721. Mean treatment size is 2.29.

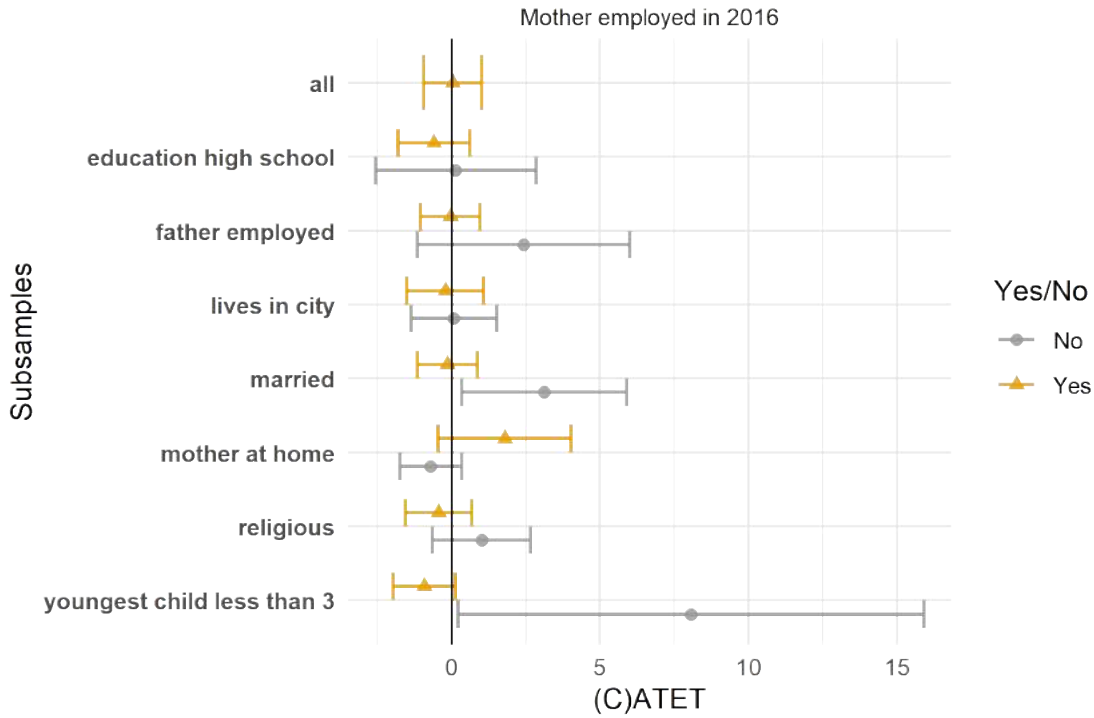
Figure 14: Effect of 10,000 HUF family tax break on maternal employment (in percentage points), families with less than 2 children




*Note:* the author's calculations based on regressing the indicator for maternal employment in 2016 on cohort groups interacted with treatments, with baseline of 1970-1972. Full set of controls include 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. Sample includes families with less than 2 children by the mother's age of 35, with sample size of 10,243. Mean treatment size is 2.37.

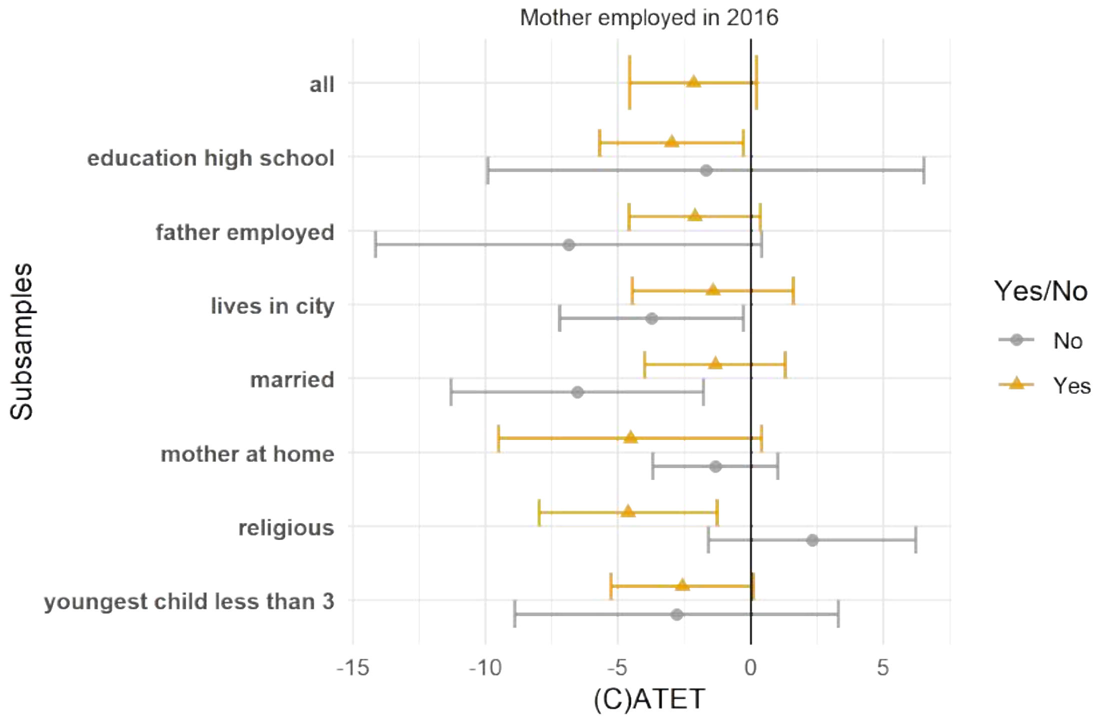
Finally, we can turn to the heterogeneity of the employment effects, displayed by Figures 15 and Figures 16, again using the 1970-1972 cohorts as baseline. For families with at least two children (Figure 15) we can detect lower employment exactly where higher completed fertility is visible (religion, 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 less than three years old. However, we can see some statistically significant positive employment effects for mothers who are at home in 2011. For families with less than two children Figure 16 suggests significant differences only along religiosity at the 5% level: maternal employment for more religious couples seem to be lower than their counterparts.

Figure 15: Effect of 10,000 HUF family tax break on maternal employment (in percentage points) for 1975-1976 Cohorts, families with at least 2 children




*Note:* the author's calculations based on regressing the indicator for maternal employment in 2016 on cohort groups interacted with treatments for the 1975-1976 cohort, with baseline of 1970-1972 for different subsamples. Full set of controls include 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. Sample includes families with at least 2 children by the mother's age of 35. Sample and treatment sizes vary between subsamples.

Figure 16: Effect of 10,000 HUF family tax break on maternal employment (in percentage points) for 1975-1976 Cohorts, families with less than 2 children




*Note:* the author's calculations based on regressing the indicator for maternal employment in 2016 on cohort groups interacted with treatments for the 1975-1976 cohort, with baseline of 1970-1972 for different subsamples. Full set of controls include 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. Sample includes families with less than 2 children by the mother's age of 35. Sample and treatment sizes vary between subsamples.

## 5.3 Robustness of the results

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 as well as 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 number of children. In the fourth column I present results from an alternative specification where the base family tax break 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 around 0.023. With the mean outcome at around 2.49 an around 0.9% effect size could be calculated, close to the main specification. The alternative, simpler specification in the fourth column suggests quantitatively very similar point estimates, but the sample selection induces a somewhat higher mean treatment size of 2.41. Finally, the discrete treatment specification yields an around 1.7% higher completed fertility if the treatment is above 12,500 HUF, which is the case for around 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 2 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 |
| controls | full | full | full | full | full |
| mean of outcome | 1.2282 | 1.2342 | 2.4888 | 1.2245 | 1.2282 |
| sd of outcome | .1966 | .2011 | .8013 | .1946 | .1966 |
| mean of Family Tax Break increment | 2.2892 | 2.0766 | 2.2892 | 2.4147 | .3681 |
| sd 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. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 4 reports the robustness checks for families with less than two children. While the estimates for the unweighted regression differs slightly in this case, the differences are within one standard error of the coefficients. In the regression with fertility measured in number of children the estimate suggests an around 0.07 child increase in family due to 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. This is quite large compared to the main specification. Finally using a discrete treatment definition I find an around 10% point higher completed fertility for a tax break higher than 12,500 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 for the 1966-1967 cohorts the estimate 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, and we should be somewhat careful about the exact magnitude of these policy effects.

Table 4: Robustness checks: completed fertility effect estimates for families with less than 2 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 |
| controls | full | full | full | full | full |
| mean of outcome | .6713 | 0.6708 | 1.0648 | .6349 | .6713 |
| sd of outcome | .3397 | .3332 | .6501 | .325 | .3397 |
| mean of Family Tax Break increment | 2.3723 | 2.2497 | 2.3723 | 2.1829 | .2958 |
| sd 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 Mechanism and alternative policy scenarios

In this section I construct a very simple model describing partial equilibrium household behavior, focusing on the key variables of fertility and maternal employment. 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 on female labor supply. The original family tax break policy (abbreviated as FTB in the figures) connects child support with employment status and the level of household income, 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 up to the level where the entire amount of the tax break can be used. Besides comparing alternative policies, I use this model to validate the previous empirical results, while also shedding light on the potential female labor supply effects of the policy that were not possible to be identified.

## 6.1 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 a 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 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 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, $\nu$ 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[14], labor supply, and the number of children. I abstract away from the male labor supply decision as men in Hungary are predominantly working (around 90% employed in this sample).

The household derives utility from consumption and the quality of children, while derives disutility from not having her ideal number of children. Consumption is taken into account 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 in accordance with the calculations of the Hungarian Statistical Office (HCSO, 2015). This abstracts away from the decision of splitting household consumption between the parents and the children, which would entail having separate utility from parental consumption and children's. 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 itself rather than labor market activities. Note that I assume that employment is purely the choice of the household and labor demand plays no part.

### 6.2 Model fit

I study the behavior of the model 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 $(\alpha, \beta)$ to match the empirical completed fertility and maternal employment of the untreated, 1966-1971 birth cohorts of the matched sample. 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.[15]

Table 5 reports the targeted and the fitted simulated values of the moments. We can see that

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

[15] 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.

the model can reproduce aggregate empirical moments reasonably well, while Figure A5 shows that the found parameter values of $\alpha = 1.8838$ (SE: 0.0087), $\beta = 0.9087$ (SE: 0.0186) indeed minimize the loss function.

Table 5: 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 predictive power of the model, Table 6 compares non-targeted empirical and simulated moments. We can see that despite the estimation not directly matching on the fertility rates conditional on income, the model can reproduce these moments for low and middle income families quite accurately. For high income families the model overpredicts completed fertility, which can be the result of the independence assumption between income and 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 exercise are fairly close.

Table 6: 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%.

## 6.3 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[16]

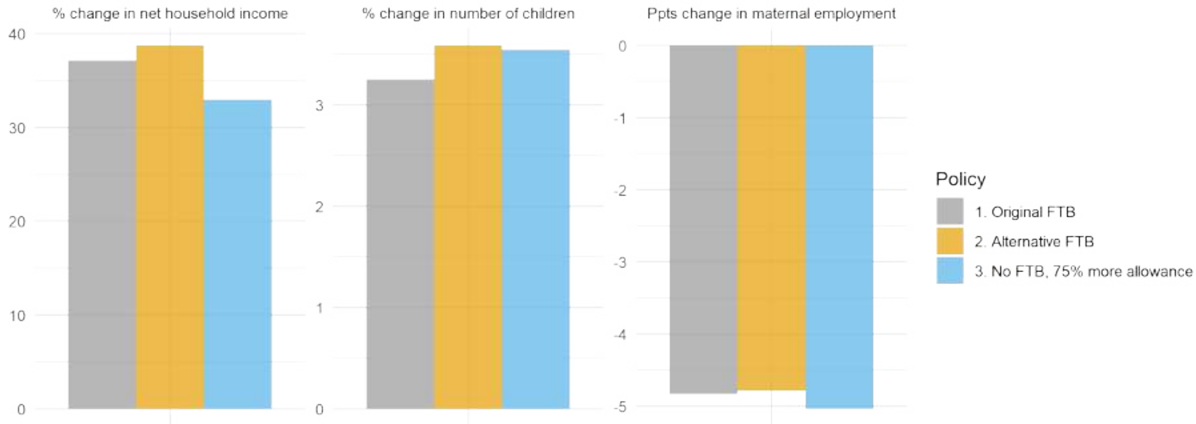
- No Family Tax Break, but a 75% increase of the Family Allowance

The point of the exercise is to see whether the model can recreate the estimated policy effects, which groups react to the policy according to the model, and how alternative simpler policies would fare. While the Alternative Family Tax Break simulates a simpler version keeping the incentives to work, the increase of the Family Allowance simply 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 17 shows the aggregate changes due to different policies compared to the baseline in 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 higher, 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 close 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 a slightly higher fertility (3.59%), with 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.

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

Figure 17: Policy scenarios: aggregate changes in outcomes compared to the no-policy baseline

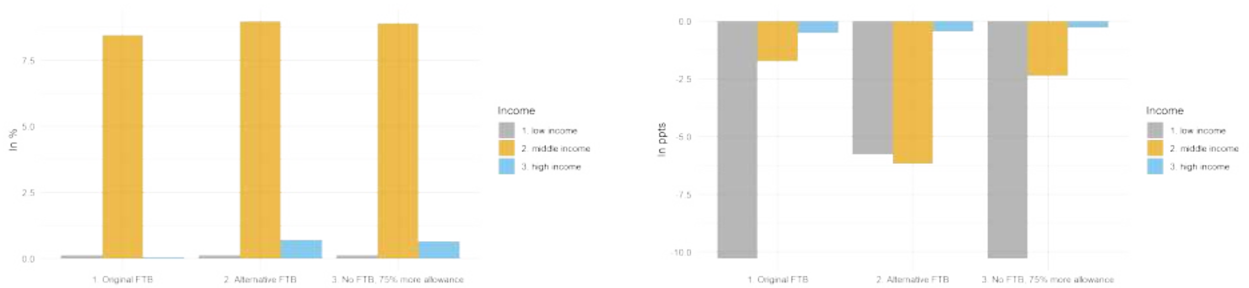



*Note:* the figure reports the model simulations' aggregate outcomes under different policies of the original Family Tax Break (FTB), the alternative Family Tax Break, and the no tax break scenario with elevated Family Allowance.

Figure 18 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 of either policy is generated by the middle income households, with around 8% increase in the number of children, while the effect on the two other groups is minimal. Both alternative policies result in higher fertility for the middle and high income families leaving the low income families unaffected. The impact on maternal employment however, is different. While the original family tax break and the elevated family allowance decreases 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 of female employment for both low and middle income families in the long-run, while still keeping high income families untouched.

Figure 18: Policy effects compared to the baseline by income groups

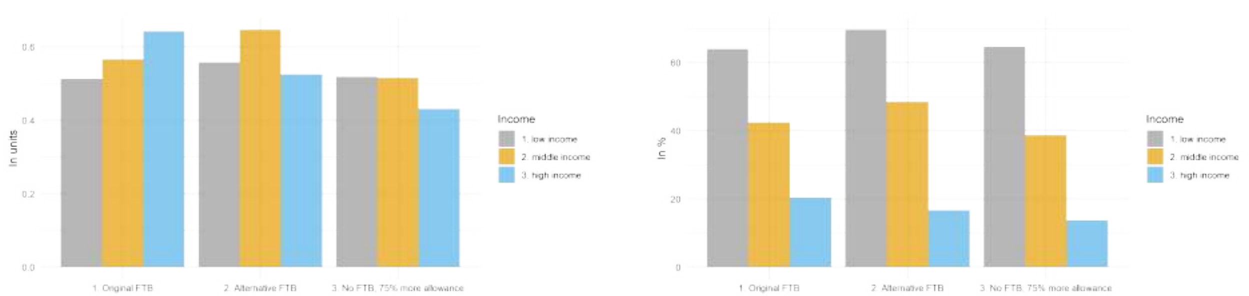
(a) Number of children (b) Maternal employment




*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 19 reports the change in net income, on the left hand side in absolute terms (units of poverty line consumption), while on left hand side it is expressed in percentage terms. 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 end up enriching the middle income households the most, while providing the least to the high income households. In percentage terms all of the examined policies provide more to the poorer households, as they start with a lower basis.

Figure 19: 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%.

# 7 Discussion

In this section I put the findings into perspective: discuss the internal and external validity, and place them into international context based on the comparable 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 relationship existed both in 2011 and 2016. That inherently puts some constraints on the external validity of the results, as the sample necessarily includes only more stable relationships. The family tax break policy affects families over a long portion of the life-cycle, so more stable relationships could expect more benefits from it in the long-run. This might lead to on overestimation of the policy effect for the entire population. However, the rich set of controls I include in the regressions might mitigate this bias, as I find that controls slightly decrease the size of the point estimates.

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 strongest case in terms 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 number of children pre-trends disappear, the indicated treatment effects might deviate from the main estimates might with different regression specification, which suggests that we should be more careful with the interpretation of the results.

With that in mind, we can summarise the findings of this paper 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 induced an around 0.76% increase in completed fertility, and for families with less than two children a less robust, around 5.64% effect is identified. Regarding maternal employment, there is no evidence for 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 yields 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 7 puts the findings of this paper into context by collecting the findings of some relevant papers in the literature. The present results are lower than the short-run estimates of literature, while somewhat higher than the previous completed fertility estimates suggest. One possible explanation is that this paper aims to measure only the quantum effects, not incorporating changes in fertility timing due to the policy, while examining a country where additional financial resources might indeed make a difference in the life of a family, while the size of the treatment and the type of the policy being a recurring tax break might provide higher incentives for families compared to a one-time transfer per child. It is important to note that we cannot yet observe the effect for families that adjust their entire life-cycle in light of the changed tax incentives, and the comparative statics exercise presented here suggests a somewhat larger effect on completed fertility. As a last point I mention that the magnitudes presented here are aligned with previous aggregate or quasi-panel results for Hungary (Gábos et al., 2009; Bördős and Szabó-Morvai, 2021).

Table 7: 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<br>PV: 13,000 EUR<br>(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.

# 8 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 induced quasi-experimental variation along household income and the initial number of children due to its non-linear jump at the third child. I take advantage of the structure of the policy by defining treatment variables as the additional net income that the household receives for additional children, based on a unique, family-level linked dataset I create of the 2016 Microcensus and 2011 Census of Hungary augmented with prospective parental salaries from the National Wage Surveys.

I find that the policy led to around 0.76% higher completed fertility for families with at least two children, and to around 5.64% for those with less than two children. 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, however, there seem to be some negative employment effects for those segments where completed fertility increased. I also fit a very 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 an around 3.23% higher completed fertility, and around 4.6 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 the reported preferences of women in Hungary, indicating that there might be even more important obstacles than financial constraints in realizing family size goals. These might include: the availability of nurseries, the difficulty to return to the labor market, the participation of fathers, 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.

Billari, F. C., Philipov, D., and Testa, M. R. (2009). Attitudes, Norms and Perceived Behavioural Control: Explaining Fertility Intentions in Bulgaria. *European Journal of Population / Revue européenne de Démographie*, 25(4):439–465.

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.

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

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.

Frejka, T. and Sardon, J.-P. (2004). *Childbearing Trends and Prospects in Low-Fertility Countries: A Cohort Analysis*, volume 13 of *European Studies of Population*. Springer Netherlands, Dordrecht.

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.

Goldstein, J., Sobotka, T., and Jasilioniene, A. (2009). The End of "Lowest-Low" Fertility. *Population and Development Review*, 35(4):663–699.

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.

Greenwood, J., Guner, N., and Vandenbroucke, G. (2017). Family Economics Writ Large. *Journal of Economic Literature*, 55(4):1346–1434.

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.

Seltzer, J. A. (2019). Family Change and Changing Family Demography. *Demography*, 56(2):405–426.

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 sontelles 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.

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

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.

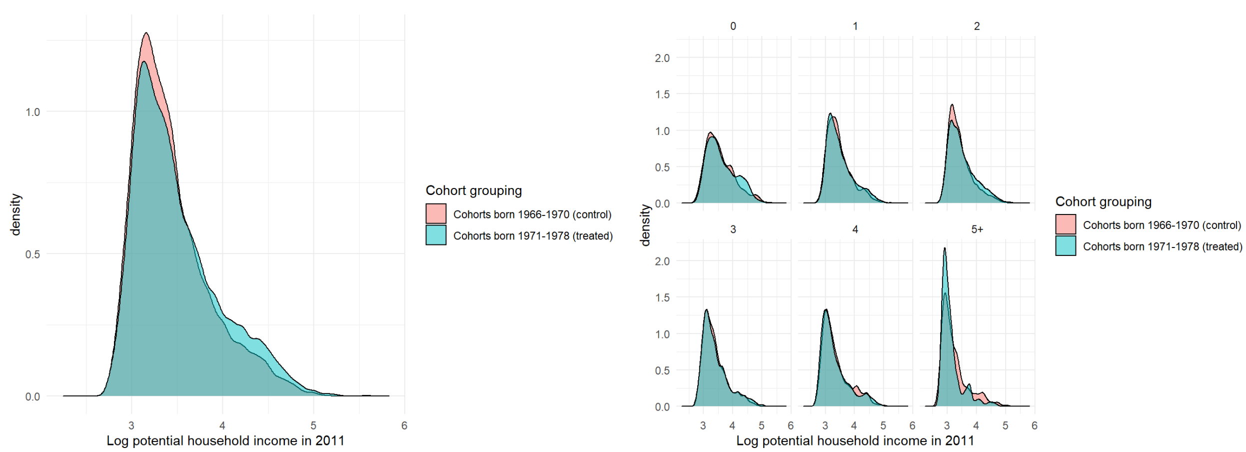
# A Appendix

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<br>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 |

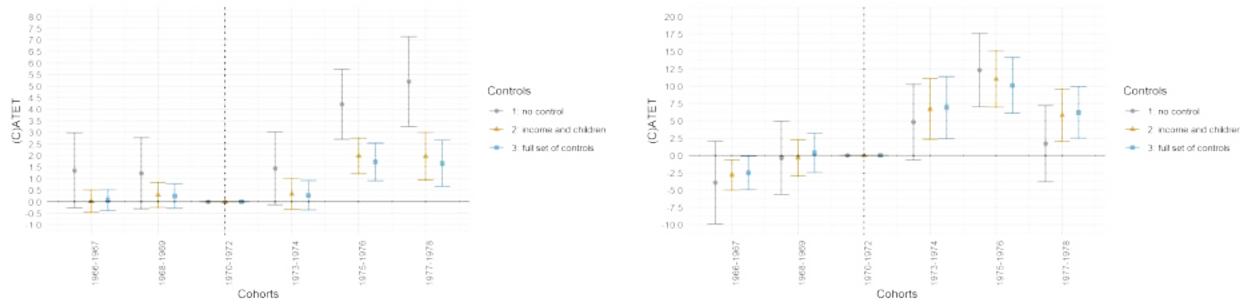
*Note:* the table reports the most important family policies in Hungary apart from the family tax break, based on Makay (2015). CSED and GYED are subject to personal income tax, GYES is subject to a 10% pension contribution.

Figure A1: Kernel density of imputed prospective household income for groups by cohorts (left), and by number of children in 2011 (right)



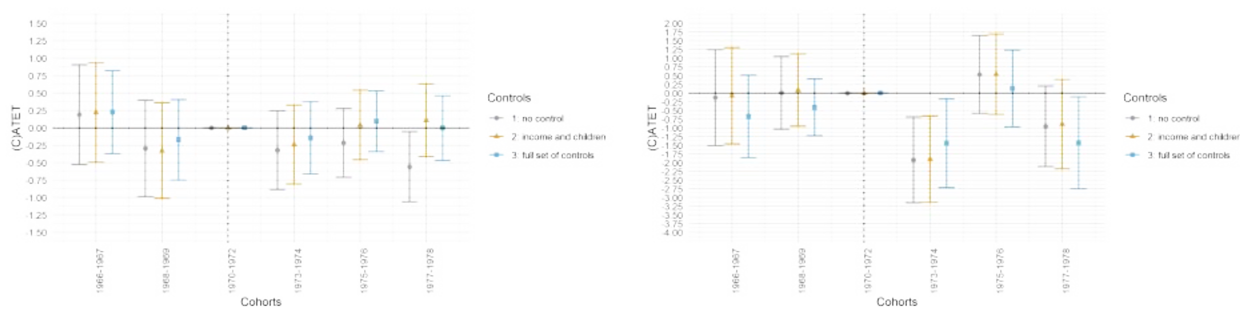

*Note:* the author's calculations based on the matched dataset of the 2016 Microcensus and the 2011 Census of Hungary.

Figure A2: Effect (discrete) on completed fertility for families with at least 2 children



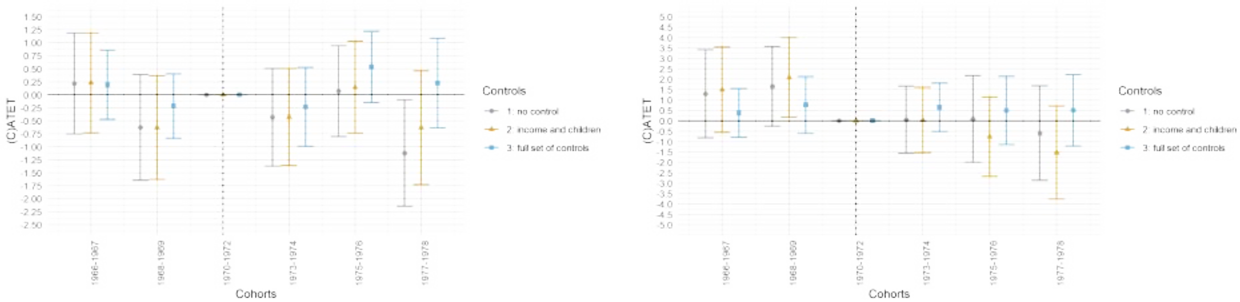

*Note:* the author's calculations based on regressing the log(1+number of children) in 2016 on cohorts interacted with treatments, with baseline cohort being 1971. X-controls include the following from 2011: gender composition of children; parental education, employment, occupation, industry, log of 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 subcounty ('járás') level. Sample includes families with at least 2 children before mother's age of 35, with sample size of 20,377. Mean treatment size: 0.3681 (increment), 0.0385 (base). Adjusted R-squared of the regressions: 0.192, 0.389, 0.901.

Figure A3: Increment effect on father's employment for families with at least 2 children



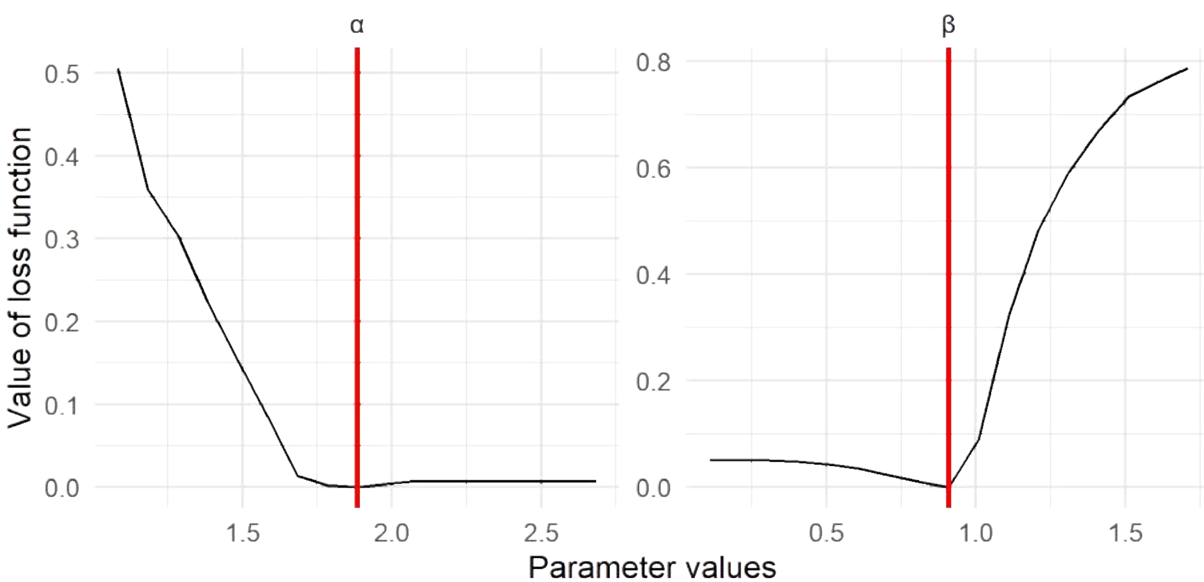

*Note:* the author's calculations based on regressing the father's employment in 2016 on cohorts interacted with treatments, with baseline cohort being 1966. X-controls include the following from 2011: gender composition of children; parental education, employment, occupation, industry, log of prospective 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 subcounty ('járás') level. Sample includes families with at least 2 children before mother's age of 35, with sample size of 20,721. Mean treatment size: 2.29 (increment), 2.90 (base).

Figure A4: Effect on home size for families with at least 2 children




*Note:* the author's calculations based on regressing the log(father's salary) in 2016 on cohorts interacted with treatments, with baseline cohort being 1966. X-controls include the following from 2011: 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 subcounty ('járás') level. Sample includes families with at least 2 children before mother's age of 35, with sample size of 20,721. Mean treatment size: 2.29 (increment), 2.90 (base).

Figure A5: 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.