Financial Incentives to Fertility: From Short to Long Run

Lidia Cruces*

F. Javier Rodríguez-Román†

December 23, 2024

Click here for the latest version.

Abstract

Are financial incentives effective in increasing fertility rates? Empirical evidence suggests they are, primarily in the short run (around implementation). Can such policies also increase the total number of children in the long run? We address this question by using a structural life-cycle model of fertility and labor supply, calibrated to replicate the short-run effects of a cash transfer paid at childbirth implemented in 2007 in Spain. The model incorporates labor market duality, a defining feature of Spanish labor markets that negatively impacts fertility. Our calibrated model replicates a 6% increase in fertility rates in the short run but only generates a 3% rise in completed fertility over women's lifetimes—the long run. Eliminating labor market duality increases lifetime fertility by 6.62%, but the discrepancy between short- and long-run effects of the incentive persists. These results highlight the limited impact of financial incentives alone to sustain fertility gains.

Keywords: Cash Transfers, Fertility, Female Labor Force Participation, Dual Labor Markets, Life-Cycle.

JEL Codes: J11, J13, J22

*Goethe University Frankfurt. Email: cruces@econ.uni-frankfurt.de.

†Universitat de Barcelona. Email: fjavier.rodriguez@ub.edu.

1 Introduction

To maintain a stable population, the fertility rate must exceed the replacement level. This requires each woman to have, on average, slightly more than two children. However, in a growing list of countries, the total fertility rate (TFR)—the average number of children born to a woman based on current age-specific birth rates—has fallen well below this threshold. In some cases, this low fertility rate has persisted for more than three decades, posing a risk of population decline.[1]

The macroeconomic consequences of low fertility rates are complex and difficult to predict. However, the literature identifies several potential issues: pressures on social security funding (De Nardi et al., 1999), decreased output growth and investment (Aksoy et al., 2019), and theoretically even a situation as dramatic as living standards stagnating as the population vanishes (Jones, 2022).[2]

Many governments have implemented policies to increase fertility rates in response to these concerns. In 1990, only 14% of governments worldwide had pronatalist policies, but by 2019, this figure had doubled to 28% (United Nations, 2021). Financial incentives, such as cash transfers provided at childbirth, represent one common strategy to boost fertility. Empirical evidence indicates that these policies can increase birth rates, although primarily in the short term—within a time window close to their implementation.[3] However, population growth ultimately depends on the total number of children women have at the end of their reproductive years. This measure, known as the completed fertility rate (CFR), can only be observed in the long term.[4]

Evaluating empirically the impact of financial incentives on CFR presents two main challenges: first, tracking treated and control groups of women over their fertile years is often not possible; second, long-term fertility outcomes may be influenced by

[1]For instance, countries like Italy, Japan, and, more recently, China are already experiencing population decline due to persistently low fertility rates. Other countries, such as Germany, Spain and South Korea, are at risk of following suit.

[2]Jones (2022) shows that growth models with endogenous fertility feature two stable steady states. When initial population growth is negative, the economy can move towards the "Expanding Cosmos" steady state, with exponential growth in population and living standards, if it manages to raise fertility rates before a certain level of knowledge per person. However, after that point, the economy gets stuck in the "Empty Planet" steady state, with stagnant living standards for a vanishing population.

[3]For example, studies on the effects of financial incentives in Spain, Canada, Israel, Germany, and South Korea include González (2013), Milligan (2005), Cohen et al. (2013), Riphahn and Wiynck (2017), and Kim (2024), with most employing differences-in-differences and regression discontinuity designs.

[4]See the Online Appendix A for a discussion of the relationship between the TFR and the CFR.

various external factors other than the policy itself, complicating causal identification. Consequently, empirical studies on the long-run effectiveness of financial incentives are scarce.[5] This paper addresses this gap, employing a structural approach.

We first present a simple two-period model of fertility that illustrates why the short- and long-run effects of financial incentives on fertility may be different. The model accounts in a parsimonious way for the trade-offs between early parenthood, late parenthood, and childlessness as a function of the short- and long-term costs of children. Beyond mere changes in the timing of births, it uncovers an additional reason for the effects of the incentives to vary over time: cohorts of different ages at the time the incentive is introduced respond differently over time. Those who are young when the incentive is introduced may have a different fertility behavior later in life than those who are older when the incentive is given for the first time.

Next, we introduce a comprehensive life-cycle model with endogenous labor force participation as well as fertility decisions, and calibrate it to reproduce the short-run effects of a well-studied cash transfer policy implemented in Spain in 2007: the "Cheque Bebé" (baby check). The model features ex-ante heterogeneity in fertility preferences and ex-post heterogeneity in labor market outcomes and attempts to capture the intricate trade-offs between career and family choices that drive decisions such as having an additional child, delaying motherhood, or remaining childless.

Moreover, we model a key characteristic of the Spanish labor market: duality, i.e., the coexistence of temporary and permanent contracts. Young workers typically start their careers in temporary employment and transition to permanent positions after moving through various temporary roles. These are associated with lower earnings and reduced returns to experience, which negatively impact fertility.[6] In the model, some women enter the labor market with a permanent contract, which they retain

[5]To our knowledge, only two empirical studies assess the long-run effects of financial incentives on fertility: Parent and Wang (2007) and Kim (2014). Parent and Wang analyzed the Family Allowance Program in Canada during the mid-1970s, finding a low long-run response compared to the short-run. Kim studies changes in Quebec's Allowance for Newborn Children policy and finds little to no permanent impact on fertility. Both studies indicate no increase in lifetime fertility, only a shift in the timing of childbirth. However, their estimates only have a causal interpretation under very strong assumptions.

[6]Empirical evidence shows that temporary jobs reduce fertility, as demonstrated by studies such as Adserà (2004), De la Rica and Iza (2005), Auer and Danzer (2016), and Lopes (2019). Furthermore, Guner et al. (2024) uses a structural model of endogenous female labor supply and fertility decisions to show that eliminating labor market duality and split-shift work schedules in Spain would significantly increase fertility.

throughout their careers. However, most start with a temporary one and can only transition to permanent contracts by working and accumulating experience.

After showing that our calibrated model performs well in key dimensions relevant to our main question, particularly in reflecting fertility response patterns by age in the short run, we explore which factors drive these patterns. We find that among younger women, those who have lower household incomes and higher desired fertility respond. Among older women, those who respond have relatively high household income and earn a relatively large fraction of it.

Finally, we use our calibrated model as a laboratory for two quantitative experiments. First, we simulate the entire life cycle of a cohort of women who receive a cash transfer incentive from the beginning of their childbearing years to uncover its effect on their completed fertility rate. We find that this long-run effect is about half the size of the (cross-sectional) short-run effect (3% versus 6% increase). This disparity arises because younger women have greater flexibility in timing additional childbirths and prefer to have more additional children earlier and fewer later in life, while older women are more constrained to respond within a few years after the implementation.

In our second experiment, we simulate a counterfactual economy in which there is a single contract instead of both temporary and permanent ones. We find that the completed fertility is 6.62% higher in this scenario and the incentives have a slightly lower effect. Still, the ratio between short- and long-run effects is similar to the one in our benchmark featuring duality.

Previous studies have analyzed the effects of family policies that promote fertility through structural life cycle models with endogenous female labor supply and fertility decisions. Erosa et al. (2010) show that the introduction of mandatory parental leave entitlements in the United States increases fertility rates and employment among mothers with infants. Bick (2016) finds that in West Germany, increasing subsidized child care generates a modest increase in maternal labor force participation and an increase in the intensive margin, but does not boost fertility. Guner et al. (2024) demonstrates that the reduction in childcare costs in Spain increases the total fertility rate and the employment rate of mothers, including those with babies.[7] More closely related, Adda et al. (2017) evaluated the short- and long-run impacts of pronatalist

[7]The effects of childcare costs on female labor supply have been studied, among others, by Attanasio et al. (2008) and Guner et al. (2020) for the United States

(counterfactual) transfer policies in Germany, finding that the long-term effects of subsidy policies are significantly lower than the short-term effects typically reported in the literature. Our paper contributes to this literature in two ways. First, we calibrate our life-cycle model to match empirically identified short-term responses to an actual policy, bridging the gap between empirical and structural approaches to these questions.[8] Second, we incorporate heterogeneity in desired fertility levels based on survey data. This allows us to match the distribution of women by number of children without introducing vast amounts of income heterogeneity, or other forms of ex-ante differences in preferences.

The structure of the paper is as follows. Section 2 introduces the stylized model. Section 3 describes the baby check and presents descriptive evidence regarding the context in which it was implemented. Section 4 outlines the quantitative life-cycle model. Section 5, discusses the parametrization of the model and our choice of parameter values. Section 6 discusses the main quantitative experiments. Finally, Section 7 reviews the key assumptions underlying the model and what they mean for the results, and Section 8 concludes.

2 The Short and Long Run in a Stylized Model

In this section, we present a simple model of fertility decisions to illustrate how a cash transfer upon childbirth can have different short- and long-run impacts on fertility. The model accounts for changes in the timing of births as a potential driver of this difference but shows an additional subtler mechanism: distinct responses by cohorts of women who are given the incentive at different ages.

Consider an economy populated by overlapping generations of households living for four periods: child, young, middle, and old. Households can have children in the young and middle periods. A fraction $1 - \epsilon$ of each cohort has N children in total, evenly distributed between the young and middle periods. From now on, we focus on the remaining ϵ fraction of households. They desire to have one child, and their only decision in each period is whether or not to have it.[9] Childless households experience

[8] Adda et al. (2017) explicitly state that combining clean designs with structural models like theirs can be useful to explore the long-run effects of pronatalist policies.

[9] These assumptions allow us to address the timing and quantity of children in a parsimonious way.

a utility cost of δ and supply 1 unit of labor to the market at an exogenous wage rate ω per period. If a child is born, the labor supply decreases to $1-\xi$ for one period, which means that having a child incurs a one-period time cost of ξ. Furthermore, the wage rate ω increases by a factor of $\underline{X}$ if the household had a child in the previous period and by a factor of $\bar{X}$ otherwise. Thus, having a child results in a permanent reduction in earnings, i.e., a long-run earnings penalty of $\frac{\bar{X}}{\underline{X}}-1$. Finally, there may be a cash transfer upon childbirth ν.

Households maximize their lifetime utility by choosing between early parenthood (having the child when young), late parenthood (having the child in middle age), and childlessness. The respective utilities, disregarding discounting for simplicity, are as follows:

$$U_E = \log\left(\omega\left(1-\xi\right)+\nu\right)+\log\left(\omega\underline{X}\right)+\log\left(\omega\underline{X}\bar{X}\right) \tag{1}$$

$$U_L = \log\left(\omega\right)-\delta+\log\left(\omega\left(1-\xi\right)\bar{X}+\nu\right)+\log\left(\omega\underline{X}\bar{X}\right) \tag{2}$$

$$U_N = \log\left(\omega\right)-\delta+\log\left(\omega\bar{X}\right)-\delta+\log\left(\omega\bar{X}^2\right)-\delta. \tag{3}$$

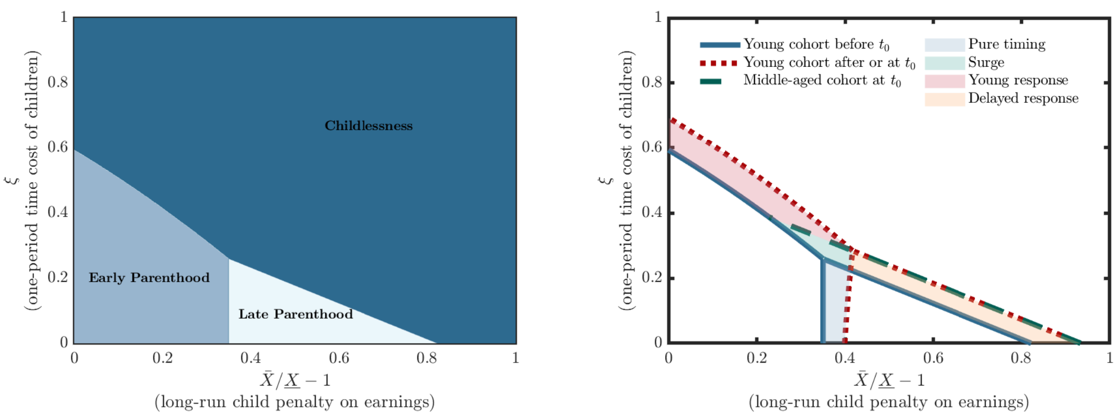
Figure 1a illustrates the fertility decisions of households in terms of the one-period time cost of children, ξ, and the long-run child penalty, $\frac{\bar{X}}{\underline{X}}-1$, in the absence of the cash transfer ($\nu=0$). When both costs are low, and the one-period time cost is relatively small compared to the long-run penalty, households prefer early parenthood. Conversely, if the long-run child penalty is large relative to the one-period time cost, households tend to defer childbirth. Finally, when both costs are high, households choose to remain childless.

Now, suppose $\nu>0$ from t_0 onward. Figure 1b shows the fertility decisions of the relevant cohorts before and after t_0. For young cohorts from t_0 onward, the decision boundaries shift upward and to the right, reducing the region where households opt to remain childless, while expanding those that favor early and late parenthood. For the middle cohort at t_0, the area that favors late parenthood expands. The extent of these shifts depends on the size of the cash transfer.

The shaded regions in Figure 1b represent the combination of parameters in which at least one cohort alters its behavior in response to the cash transfer. We define the short-term response as the percentage change in the birth rate from $t_0 - 1$ to t_0, while the long-term response refers to the percentage change in the completed fertility for cohorts that reach young age at or after t_0, compared to those who reached it before $t_0 - 1$. In the following, we discuss these for two of the shaded regions.[10]

Figure 1. Decision Boundaries: Before and After the Introduction of the Cash Transfer

(a) Before **(b)** Comparison Before and After



Pure Timing. Before t_0, a fraction ϵ of each cohort chooses late parenthood. The cohort who is middle-aged at t_0 maintains this choice, while every subsequent one shifts to early parenthood. Thus, at t_0, for one period only, a fraction ϵ of both the young and middle cohorts has one child. The birth rate increases during this period, even though the completed fertility of each cohort remains unchanged at $(1-\epsilon)N + \epsilon$ children per woman. Thus, the short-run effect results purely from a change in timing.

Surge. Before t_0, a fraction ϵ of each cohort chooses childlessness. The cohort who is middle-aged at t_0 changes to late parenthood, while every subsequent one shifts to early parenthood. Thus, at t_0, for one period only, a fraction ϵ of both the young and middle cohorts has one child. The birth rate increases during this period from $\frac{(1-\epsilon)N}{2}$ to $\frac{(1-\epsilon)N+\epsilon}{2}$, a $\frac{2\epsilon}{(1-\epsilon)N}$ percent increase. The completed fertility rate per cohort increases from $(1-\epsilon)N$ to $(1-\epsilon)N + \epsilon$, a $\frac{\epsilon}{(1-\epsilon)N}$ percent increase. Thus, the increase in births in the short run is greater than in the long run. For one period, there is a surge in births above the new (higher) level.

Notice that the short-run response is an aggregation of the responses by the cross section of women who are young and middle-aged at t_0, while the long-run response

[10]The additional cases are provided in the Online Appendix B.

depends on the life-cycle behavior or those cohorts that are exposed to the incentive from young age, that is, those who reach young age at or after t_0. The discrepancy arises because these cohorts behave differently later in life when they are middle-aged than the cohort that is middle-aged at t_0. This mechanism will be crucial to explain our quantitative results with the more comprehensive model we present in Section 4.

3 Institutional Background

In this section, we describe the Spanish financial incentive to fertility we use to calibrate our model and provide some facts about the economic context in which it was implemented. These facts inform our modeling of fertility preferences and labor markets.

3.1 The Baby Check

Since the early 1990s, Spain has consistently ranked among the countries with the lowest fertility rates.[11] With the explicit goal of encouraging people to have more children, the Spanish government introduced the "Cheque Bebé" (baby check) on July 3, 2007: a universal, lump sum, €2,500 payment to families upon the birth of a child.[12] To give a sense of the magnitude of the incentive, the monthly gross minimum wage in Spain in 2007 was €570 and the median gross monthly earnings of women were €1,190.

The baby check had practically full take-up. In the first full year of the policy (2008), approximately half a million payments were distributed, for a total cost of €1.2 billion, around 0.2% of total public expenditure for that year.[13] However, the policy ultimately would be short-lived. On May 12, 2010, as part of austerity measures implemented in response to the 2008 financial crisis, the baby check was unexpectedly

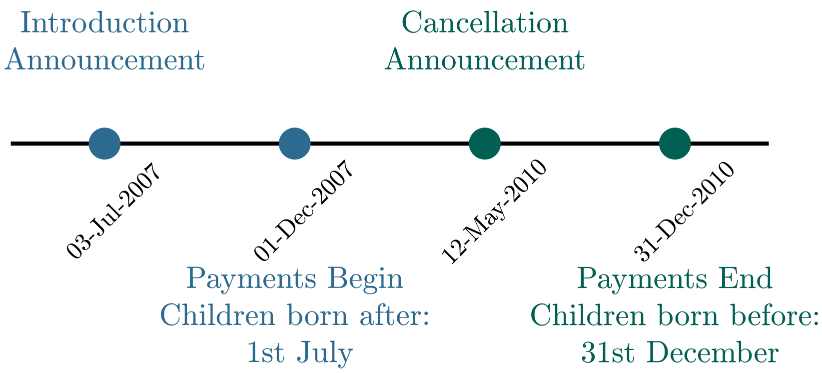
[11] See Section C in the Online Appendix for more on Spain's demographics in context.

[12] The baby check was approved on July 13, and it was declared that parliament would pass a new law in an accelerated procedure, which became *Law 35/2007*, November 15, 2007. The policy had an additional explicit objective of helping families financially with the cost of childbirth. Initially, all births that occurred after the announcement were eligible, but the cut-off date was later moved to the 1st of July for practical reasons. The only requirement stated in the law was two years of prior residency in Spain, but this was relaxed in 2009.

[13] *Presupuestos Generales del Estado, 2008. Ministerio de Economía y Hacienda*, December 26, 2008.

canceled (see Figure 2 for the policy's timeline).[14]

Figure 2. Timeline of the Baby Check



Two key studies provide insights into the fertility and labor supply responses of mothers following the introduction of the baby check. First, González (2013) exploits the unexpectedness of the announcement with regression discontinuity and differences-in-differences designs, and finds a 6% increase in annual births and a 2–4 percentage point drop in maternal labor force participation 12 months postpartum. Second, González and Trommlerová (2023) examine both the introduction and cancellation of the policy, differentiating between the effects due to fewer abortions and more conceptions. They find that a 3.5% rise in births following the policy announcement was driven by fewer abortions, and a further 2.8% increase was attributable to higher conception rates. Conversely, the cancellation triggered a 4.7% short-term rise in births, followed by a 5.7% decline due to increased abortions and reduced conception rates. Importantly, the effects found in these two studies are comparable in magnitude with those of similar policies implemented in Canada (Milligan, 2005) and Israel (Cohen et al., 2013).[15]

Furthermore, González and Trommlerová (2023) argue that the observed effect is not solely a tempo effect—women having children earlier—but also a quantum effect—women having more children overall. They base their claim on two observations. First, the baby check boosted births across all parities. They argue that, for example, it is unlikely that women with two or more children would have additional children if the effect were purely a tempo shift. Second, the policy contributed to higher birth rates

[14] *Royal Decree-Law 8/2010*, May 20, 2010.

[15] González and Trommlerová (2023) convert their findings to be comparable with the other two studies. Milligan (2005) estimates that €1,000 in benefits increases fertility by 2.6 percent, against 1-2 percent in González and Trommlerová (2023). Cohen et al. (2013) find that a 1 percent reduction in the price of children increases fertility by 0.5 percent, while González and Trommlerová (2023) find an effect between 0.2 and 0.4 percent.

among older women. However, the short duration of the policy (3.5 years) and the concurrent economic crisis complicate the empirical evaluation of its long-term effects on completed fertility rates.

In this paper, we evaluate the long-run effects of the baby check by building a life-cycle model that can reproduce these important results and simulating cohorts of women who are given the incentive from the beginning of their reproductive years.

3.2 The Spanish Context

Given that institutional factors shape policy impact, we next provide an overview of childcare arrangements, labor market conditions, and fertility preferences in Spain, which support our model assumptions in Section 4. In this paper, we rely on four primary data sources: the MCVL *Muestra Continua de Vidas Laborales* (Continuous Working Life Sample), EdF *Encuesta de Fecundidad* (Fertility Survey), EPF *Encuesta de Presupuestos Familiares* (Family Budget Survey), and ECV *Encuesta de Condiciones de Vida* (Living Conditions Survey).

3.2.1 Childcare Arrangements

Lack of access to childcare can be an important barrier to the participation of mothers in the labor force, increasing the costs of having children. The situation in Spain in 2007 was quite typical for a developed country in this regard. Between ages 3 and 6, the public education system offered preschool programs universally between 9 am and 5 pm. However, the availability of public childcare for children aged 0–3 was much more limited, with 56.2% of those attending some form of childcare doing so in private centers, which parents have to pay out of pocket. According to the EPF, families with children aged 0-3 using full-time childcare spent an average of €2000 annually in 2007, equivalent to approximately 7% of the average annual wage at the time.

3.2.2 Labor Market Dynamics

Spanish labor markets are characterized by duality. In a dual labor market, older and more experienced workers hold highly protected permanent jobs, while young people experience high turnover rates across precarious temporary ones. According to MCVL, only 43. 7% of women workers aged 25 years had a permanent contract in 2007, this share increasing with experience, but plateauing at 75% at age 45. Having a temporary contract lowers income because the worker experiences periods of unemployment: over the course of a year, workers who have a permanent contract work on average 339 days versus 219 for those who do not have one. Moreover, the returns to experience are lower for people with temporary contracts (Garcia-Louzao et al., 2023). These dynamics, which affect most young women, increase the cost of early career interruptions in these crucial childbearing years. Extensive evidence suggests that temporary contracts lead women to postpone or reduce their fertility. Adserà (2004), De la Rica and Iza (2005), Auer and Danzer (2016), Lopes (2019), and Guner et al. (2024) consistently document these effects. Including duality in our comprehensive life-cycle model seems essential in light of all this evidence.

3.2.3 Fertility Preferences

Across high-income countries, people consistently express a desire for more children than they ultimately have, underscoring the potential influence of policies such as the baby check. Spain exemplifies this trend, as shown by data from the 2018 EdF. Table 1 compares the desired and realized fertility for women aged 40–44 years, whose fertility decisions are likely to be completed. On average, women in this group desired 2.04 children, but had only 1.54, leaving an unmet fertility gap of 0.5 children per woman.

Table 1. Fraction of Women Aged 40–44 by Desired and Realized Number of Children

Number of children	Desired	Realized	Gap
0	7.90%	18.99%	-11.09%
1	15.20%	24.96%	-9.75%
2	49.75%	43.79%	5.95%
3 or more	27.15%	12.26%	14.89%

Source: *Encuesta de Fecundidad 2018, INE.*

The table reveals substantial disparities in childbearing patterns. For example, 18.99% of the women were childless, more than double the 7.90% who explicitly stated they did not want children. Similarly, only 12. 26% of the women had three or more children, much fewer than the 27.15% who desired larger families. These patterns suggest significant barriers to fulfilling fertility preferences.

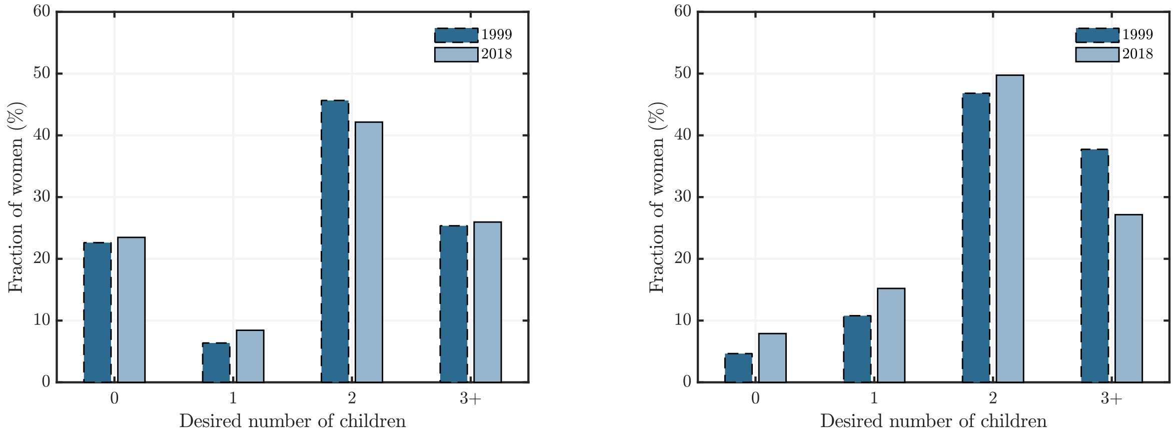
Moreover, new empirical regularities have emerged in the economics of fertility. First, there used to be a negative relationship between fertility and income, both across countries and across families within countries. These relationships have flattened and, in some cases, reversed (Doepke et al., 2022). Spain is not an exception to this trend. In addition, the relationship between *desired* fertility and income is also flat (Figure D.2 in the Online Appendix). Second, there used to be a cross-country negative relationship between women's labor force participation and fertility, which has reversed in high-income countries. Taken these two facts together, compatibility between work and family seems central in explaining systematic fertility patterns, while preferences are central in explaining differences across families. This is quite convenient because we take the latter as exogenous and focus on modeling the former.

This raises the question of whether the baby check influenced fertility preferences. To answer it, an ideal approach would be to use a Regression Discontinuity Design, as in González (2013). Unfortunately, we do not have data on the desired fertility before and after the policy announcement. We have two waves of the survey, allowing us to compare the responses of the same cohort at different ages and cohorts at the same age.

Figure 3 shows the desired number of children by cohort and by survey year. Consider first the cohort aged 21–25 in 1999 and 40-44 in 2018. When the baby check was introduced, these women were 29–33 years old in 2007 and, therefore, were affected by it. Their fertility preferences are represented by the dashed dark blue bar in Figure 3a and the solid light blue bar in Figure 3b, respectively. Notice that the desired fertility experienced some changes between those years. In particular, the fraction of women who declared not wanting children in 1999 (when they were young, e.g., before most of them made any fertility decisions yet) is significantly bigger than the fraction in 2018 (after fertility was completed). Most of the difference is attributed to a higher proportion of women declaring that they want one child. In contrast, the fractions for 2 or 3 children or more are similar. In other words, there is a discrepancy in the extensive margin but not so much in the intensive margin.

Two explanations for this trend arise: (1) the baby check may have influenced fertility preferences for this cohort, or (2) preferences may naturally evolve as individuals age. Evidence from Figures 3a and 3b suggests the latter, since fertility preferences appear stable across cohorts of women aged 21-25 and 40-44 in both 1999 and 2018.

Figure 3. Fraction of Women by the Desired Number of Children for Selected Cohorts



(a) Cohort aged 21–25 **(b)** Cohort aged 40–44

Source: *Encuesta de Fecundidad 1999 and 2018, INE.*

This supports the hypothesis that preferences evolve gradually with age and remain relatively stable across cohorts with different policy environments.

4 The Full Life-Cycle Model

In this section, we present a life-cycle model that is, in essence, similar to the stylized one presented in Section 2. It retains the idea that each household has a desired number of children, that having them is costly, and that households balance these costs against the disutility of not reaching their desired fertility. However, it constitutes a more detailed representation of the economic environment in which the baby check was implemented. The main changes are as follows. First, we jointly consider labor force participation and fertility decisions. Thus, the decline in labor supply after having a child is endogenous and emerges from the trade-off women face between working while raising children or taking a costly career break. Second, we increase the number of periods and shorten their length to one year. Third, we allow women to desire and have multiple children. Fourth, we consider that the biological probability of giving birth decreases with age. Therefore, delaying fertility is risky. Finally, we account for labor market duality by including temporary and permanent contracts.

4.1 Demographics

We model women's life choices between the ages of 25 and 52. A model period corresponds to one year. We denote model periods with j; hence, age 25 corresponds to $j = 1$ and age 52 to $j = J = 28$. Before $j = 1$, each woman is exogenously matched with a spouse/partner two years older than her, and draws a desired number of children $N^* \in \{0, 1, 2, 3\}$. There are no marital transitions (no separations or divorces) and N^* remains constant throughout their life.

From ages 25 to 39, in each period, women must decide whether to try to have an additional child or not. We denote this decision by $b \in \{0, 1\}$. If she decides to do so ($b = 1$), she gives birth to a baby next period with probability α_j, which decreases with age. Therefore, women can give birth between the ages of 26 and 40. Each woman can have a maximum of three children, and we denote the total number of children by $N \in \{0, 1, 2, 3\}$.

After being born, each child transitions stochastically between four stages of life, based on the different kinds of costs that children impose on mothers as they age. The first two stages are newborn (below one year old, thus born in the current period) and baby (between ages 1 and 3). The only difference between them is that newborns can come with a baby check. Children in these two stages incur the same consumption, leisure, and childcare costs. The third stage represents school age (between ages 3 to 12). As children move to this stage, they continue to impose consumption and leisure costs, although quantitatively not identical to those imposed by newborns and babies. However, full-time childcare for school-age children is free for parents (we assume universal public provision). Finally, those in the fourth stage are teenagers / young adults (aged 12 and up), which are not costly at all for mothers in terms of leisure, but represent a larger burden on consumption. The exact mathematical form all these costs take is discussed in the following subsection.

We denote by $\mathbf{n} = [n_0, n_1, n_2, n_3]$ the vector that indicates the number of newborns, babies, school-age children, and teenagers, respectively (hence $N = n_0 + n_1 + n_2 + n_3$). A newborn becomes a baby with probability one after one period. A baby becomes a school-age child, and a school-age child becomes a teen with probabilities $\lambda_1 = 1/2$ and $\lambda_2 = 1/11$, respectively. In expectation, a child spends two years as a baby and eleven years of school age. The teenage/young adult stage is absorbing: once a child reaches this stage, it remains there. We denote this structure by $\mathbf{n}' = \Lambda_j(\mathbf{n}, b)$.[16] The main advantage of modeling children's aging in this way is that we avoid carrying each child's age as a state variable.

4.2 Preferences and Constraints

Women are the sole decision-makers in the household. Alternatively, they hold all the bargaining power, and thus, the household's preferences are perfectly aligned with hers. In each period, apart from their fertility decision, they must also decide their labor force participation $h \in \{0, 1/4, 1/2\}$, where $h = 0$ means no participation, and $h = 1/4$ and $h = 1/2$ stand for part-time and full-time work, respectively.[17]

In each period, utility is derived from consumption (c), leisure (l), and children ($\mathbf{n}$). The functional form for instantaneous utility is given by the sum of CRRA terms for the first two, and an additional term which is a function of the latter:

$$u\left(c, l, \mathbf{n}; N^*\right) = \frac{\left(\frac{c}{\psi(\mathbf{n})}\right)^{1-\gamma_c} - 1}{1-\gamma_c} + \delta_l \frac{l^{1-\gamma_l} - 1}{1-\gamma_l} + \Gamma\left(\mathbf{n}, h; N^*\right), \tag{4}$$

where:

$$\Gamma\left(\mathbf{n}, h; N^*\right) = -\delta_N(\mathbf{n}; N^*) \frac{\exp(j-\gamma_N)}{1+\exp(j-\gamma_N)} |N - N^*| + \mathbb{1}_{\{N>0\}}\zeta + \mathbb{1}_{\{h=1/2\}}\kappa(\mathbf{n}). \tag{5}$$

This general specification for the utility function is extensively discussed in the context of the study of the interaction of income, wages, and fertility by Jones et al. (2010) and was used by Bick (2016) in a life-cycle model like this one. The utility term related to children is less standard and calls for some discussion.

Children affect utility directly and indirectly. The first two terms of the above expression depend purely on the number of children and are therefore the direct effect. First, women experience a utility penalty for the difference between the current (N)

[16] See the Online Appendix E.1.1 for the exact functional form this structure takes.

[17] Out of the 24 hours a day has, we assume 8 are used for sleeping and personal care, leaving 16 hours to be split between work, childcare, and leisure. A full-time job (8 hours per day), therefore, represents $\frac{1}{2}$ of the time endowment, while a part-time one (4 hours per day) represents $\frac{1}{4}$ of it.

and the desired or target number of children (N^*). This can be interpreted as an urge to have children among women who can still have them and lifetime regret among those who cannot. In either way, it is the reason why women have children in the model.

The intensity of the penalty for not having the desired number of children depends on two terms. The first, $\delta_N(\mathbf{n}; N^*)$, introduces a nonlinearity:

$$\delta_N(\mathbf{n}; N^*) = \delta_{N1}\left(1 - \delta_{N2}\mathbb{1}_{\{N=2,N^*=3\}}\right), \tag{6}$$

i.e., we allow for the marginal disutility of not having a third desired child to be different from the marginal disutility of not having the first and the second.

The second introduces age variation. In particular, it is a sigmoid function that increases with age and is asymptotic to 1. That is, younger women experience only a fraction of the utility penalty older women do, but this fraction increases with age. This can be interpreted in various ways. One is that it captures factors not included in the model that cause women to postpone fertility, such as housing and partner disposition. Another one is that it is a child-specific discount rate that decreases with age, as opportunities to have additional children diminish. Finally, there is a fixed (dis)utility of motherhood ζ.

The functional forms described above allow the model to reflect aspects of the quantum and tempo of fertility. Anticipating the calibration, it should be apparent that ζ is crucial for the extensive margin of fertility (i.e., remaining childless or becoming a mother), δ_1 and δ_2 for the intensive margin (having 1, 2, or 3 children), and γ_N for the timing of births.

The third term in $\Gamma(\mathbf{n}, h; N^*)$, $\kappa(\mathbf{n})$ aims to capture the difficulties associated with working full-time while having kids. This includes schedule conflicts between full-time childcare and work, and disutility from not spending time with children.[18]

[18]Childcare usually goes from 9 am to 5 pm, but a full-time job may not correspond to that schedule. The fact that parents with young children must drop off and pick up their children at determined hours themselves or have someone else do it introduces an additional cost to working full-time. Moreover, Guner et al. (2024) report on the unusual organization of the workday in Spain, with long lunch breaks that create split-shift schedules. For example, 50% of workers are still at work at 6 p.m. in Spain, compared to 20% in the UK.

We allow this cost to vary according to the age of the youngest child:

$$\kappa(\mathbf{n}) = \begin{cases} \kappa_1 & \text{if } n_0 + n_1 > 0 \\ \kappa_2 & \text{if } n_0 + n_1 = 0 \text{ and } n_2 > 0 \\ 0 & \text{if } n_0 + n_1 + n_2 = 0, \end{cases} \tag{7}$$

That is, if there is a newborn or baby present in the household, the utility cost of working full time is κ_1, if there are no newborns or babies, but there are school-age children, it is κ_2, and if there are no children other than teenagers, the cost if zero.

Children are costly in terms of consumption (non-childcare), and older children are more so. To reflect this, we use the OECD equivalence scale to adjust per-capita consumption in the household. It assigns a value of 1 to the head of the household, 0.7 to each adult, and 0.5 to each child. Teenagers are counted as adults; hence:

$$\psi(\mathbf{n}) = 1.7 + 0.7 n_3 + 0.5(n_0 + n_1 + n_2). \tag{8}$$

Children also negatively affect their leisure time. We assume that children require a minimum amount of time from the mother, which is a concave function of the number of them to reflect economies of scale in the childcare production function (e.g., it does not take double the amount of time to prepare food for two children than it takes to prepare food for one). Moreover, we allow the minimum time required by children to vary according to the age of the youngest. In particular:

$$\xi(\mathbf{n}) = \begin{cases} \xi_1 \sqrt{n_0 + n_1 + n_2} & \text{if } n_0 + n_1 > 0 \\ \xi_2 \sqrt{n_2} & \text{if } n_0 + n_1 = 0 \text{ and } n_2 > 0 \\ 0 & \text{if } n_0 + n_1 + n_2 = 0, \end{cases} \tag{9}$$

e.g., if there are two children in the household, the cost can be different depending on the ages. If both are of school age, the leisure cost for the mother is $\xi_2\sqrt{2}$, but if one is a baby, the cost is $\xi_1\sqrt{2}$. This provides flexibility for the model to reflect the needs children of different ages may have. Notice once again that teenagers do not impose any leisure costs.

With this, we can define leisure time for the mother, which is the residual of a time endowment of one unit minus hours at work h minus the time required by the children:

$$l = 1 - h - \xi(\mathbf{n}). \tag{10}$$

Finally, we assume that women in the labor force with newborns and babies must purchase childcare for the same amount of time she works. Resources available for consumption at the household level are therefore the household net income I_{net}^{hh} minus childcare costs:

$$c = I_{net}^{hh} - \lambda 2h\left(n_1 + n_2\right), \tag{11}$$

where λ is the cost of full-time childcare (we assume that part-time childcare costs half as much).[19]

4.3 Income

We seek to model income in a way that captures income risk faced by households, accounts for labor market duality, is capable of reflecting the degree of (after-tax) income inequality (across households and genders, i.e., reflects the gender wage gap) and accounts for the returns of experience at different ages. To this end, we model spouses' (gross) income as depending on a couple of correlated, persistent stochastic shocks, the type of contract available (temporary or permanent), and accumulated experience.

The stochastic shocks for the woman and her partner are given by ϵ^f and ϵ^m, respectively. The household draws a couple of initial shocks $\left(\epsilon_1^f, \epsilon_1^m\right)$ from an exogenous

[19] It is possible that childcare costs do not increase linearly with the number of children (because of sibling discounts, etc.). However, we were not able to find good sources to determine the shape of the function.

joint distribution that subsequently evolves over time following an AR(1) process:

$$
\begin{array}{l} \epsilon^{f\prime} = \phi^f \epsilon^f + \nu^f \\ \epsilon^{m\prime} = \phi^m \epsilon^m + \nu^m \end{array}, \quad \begin{bmatrix} \nu^f \\ \nu^m \end{bmatrix} \sim N \left(\begin{array}{c} \mu^f = 0 \\ \mu^m = 0 \end{array}, \begin{bmatrix} \sigma^2_{\nu^f} & \rho \\ \rho & \sigma^2_{\nu^m} \end{bmatrix} \right). \tag{12}
$$

Experience accumulates over time through working. We assume that the partners always work full time and therefore their experience in period j is $j-1$, $\forall j \in \{1, \ldots, J\}$. Women experience changes from one period to the next according to:

$$
x' = \begin{cases} x & \text{if } h = 0 \\ x \quad \text{w.p. } 1 - \pi_x & \text{if } h = \frac{1}{4} \\ x + 1 \quad \text{w.p. } \pi_x & \text{if } h = \frac{1}{4} \\ x + 1 & \text{if } h = \frac{1}{2}, \end{cases} \tag{13}
$$

that is, experience remains constant if she does not work, increases by one year if she works full time, and increases by one year with probability π_x if she works part-time. Notice that if $\pi_x = \frac{1}{2}$, half a year of experience is accumulated with one year of part-time work. However, $\pi_x \neq \frac{1}{2}$ could possibly be lower, to reflect a penalty of working part-time on experience accumulation. We denote this structure as $x' = \Pi_x(x, h)$.

We model labor market duality only for women. They draw an initial contract type $z_1 \in \{0, 1\}$, where 1 denotes permanent contracts and 0 temporary ones. In subsequent periods, the probability of having a permanent contract depends on experience, age, the type of contract available in the previous period, and whether she worked in the previous period:

$$
z' = \begin{cases} 0 & \text{if } z = 0 \text{ and } h = 0 \\ 0 \text{ w.p. } 1 - \pi_z(x, j) & \text{if } z = 0 \text{ and } h > 0 \\ 1 \text{ w.p. } \pi_z(x, j) & \text{if } z = 0 \text{ and } h > 0 \\ 1 & \text{if } z = 1, \end{cases} \tag{14}
$$

that is, if a woman has a temporary contract and works, she becomes permanent with probability $\pi_z(x, j)$ (which depends on her accumulated experience). If she has a temporary contract and does not work, she will still have only a temporary contract available for her next period. If she already has a permanent contract, she will also have one in the next period (it is an absorbing state). We denote this structure by $z' = \Pi_z(x, j, h, z)$.

Putting all the elements together, the full-time (potential) log income for the woman is:

$$\ln\left(y^f\right) = \eta_0^f + \Delta\eta_0^f \mathbb{1}_{z=1} + \left(\eta_1^f + \Delta\eta_1^f \mathbb{1}_{\{z=1\}}\right) x + \left(\eta_2^f + \Delta\eta_2^f \mathbb{1}_{\{z=1\}}\right) x^2 + \epsilon^f, \tag{15}$$

that is, the type of contract changes the baseline income level and the returns to experience. In particular, if $\Delta\eta_0^f > 0$, $\Delta\eta_1^f > 0$, $\Delta\eta_2^f < 0$ and $\pi_z(x)$ is increasing in x, women have an additional reason to work as much as possible at the beginning of their career, i.e. to increase the likelihood of getting an open-ended contract, under which expected income is higher.

Log income for the husband is given by:

$$\ln(y^m) = \eta_0^m + \eta_1^m (j-1) + \eta_2^m (j-1)^2 + \epsilon^m. \tag{16}$$

The household's net income is the sum of the gross incomes of both partners minus tax liabilities:

$$\begin{aligned} I_{net}^{hh} =& I\left(y^m, y^f, h\right) - T\left(y^m, y^f, h\right) \\ =& y^m + 2hy^f\left(1 - \mathbb{1}_{\{h=\frac{1}{2}\}}\phi\right) - T\left(y^m, y^f, h\right), \end{aligned} \tag{17}$$

where ϕ is an earnings penalty on part-time work, and $T(.)$ is a tax liability function.

4.4 Problem in Recursive Form

Upon entering the economy, women draw a fertility preference N^*, an initial type of contract z_1, and initial income shocks for them and their partners $\left(\epsilon_1^f, \epsilon_1^m\right)$. Then, they enter period 1 with no experience and no children, i.e. $x_1 = 0$ and $\mathbf{n}_1 = [0, 0, 0, 0]$.

From period 1 on, women observe their state vector $\left[\epsilon^f, \epsilon^m, z, x, \mathbf{n}\right]$, and choose their labor supply h and whether or not to try to have an additional child next period $b \in \{0, 1\}$. This continues for every period until they reach 3 children or age 39 ($j = 15$). After this happens, they cannot have any more children, and they choose only labor force participation in each period.

The dynamic problem women solve in period j is given by:

$$
\begin{aligned}
&V_j\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) = \\
&\max_{\substack{h \in \{0, \frac{1}{4}, \frac{1}{2}\} \\ b \in B_j(\mathbf{n})}} u\left(c, l, \mathbf{n}; N^*\right) + \beta \mathbb{E}\left[V_{j+1}\left(\epsilon^{f\prime}, \epsilon^{m\prime}, z', x', \mathbf{n}'; N^*\right) \mid \epsilon^f, \epsilon^m, z, x, \mathbf{n}\right] \\
&\text{s.t.} \\
&l = 1 - h - \xi(\mathbf{n}) \\
&c = I_{net}^{hh} - \lambda 2h\left(n_1 + n_2\right) \\
&I_{net}^{hh} = y^m + 2hy^f\left(1 - \mathbb{1}_{\{h=\frac{1}{2}\}}\phi\right) - T\left(y^m, y^f, h\right) \\
&\ln\left(y^f\right) = \eta_0^f + \Delta\eta_0^f \mathbb{1}_{z=1} + \left(\eta_1^f + \Delta\eta_1^f \mathbb{1}_{\{z=1\}}\right) x + \left(\eta_2^f + \Delta\eta_2^f \mathbb{1}_{\{z=1\}}\right) x^2 + \epsilon^f \\
&\ln\left(y^m\right) = \eta_0^m + \eta_1^m\left(j-1\right) + \eta_2^m\left(j-1\right)^2 + \epsilon^m \\
&\epsilon^{f\prime} = \phi^f \epsilon^f + \nu^f \\
&\epsilon^{m\prime} = \phi^m \epsilon^m + \nu^m \\
&\mathbf{n}' = \Lambda_j(\mathbf{n}, b) \\
&x' = \Pi_x\left(x, h\right) \\
&z' = \Pi_z\left(x, j, h, z\right),
\end{aligned}
\tag{18}
$$

where the choice set for the birth decision is defined as:

$$B_j(\mathbf{n}) = \begin{cases} \{0,1\} & \text{if } N < 3 \text{ and } j < 15 \\ \{0\} & \text{otherwise.} \end{cases} \tag{19}$$

5 The Benchmark Economy

In this section, we explain how we calibrate the model parameters, with most details provided in the Online Appendices E and F. We then evaluate how the calibrated model performs in key areas relevant to our main question. Finally, we briefly explore which households respond to the baby check in our model.

5.1 Calibration

We calibrate our model in two steps. First, we choose or estimate the parameters that can be backed out without solving the full model. This includes assigning a subset of parameters to standard values derived from the existing literature and directly estimating several parameters using Spanish microdata. In the second step, we apply the method of simulation moments (MSM) to calibrate the remaining parameters.

5.1.1 Parameters Chosen Before Solving the Model

Table 2 summarizes the parameter values of those estimated or taken from the data without solving the model. See Appendix F.1 for details on the procedures by which these parameters are estimated.

Figure 4 shows the predicted earnings profiles by years of experience for men, and women in temporary and permanent contracts, based on our estimates from the MCVL. Earnings for women in temporary contracts are lower on average and start flattening out at around 20 years of experience, while those for women in permanent contracts keep growing.

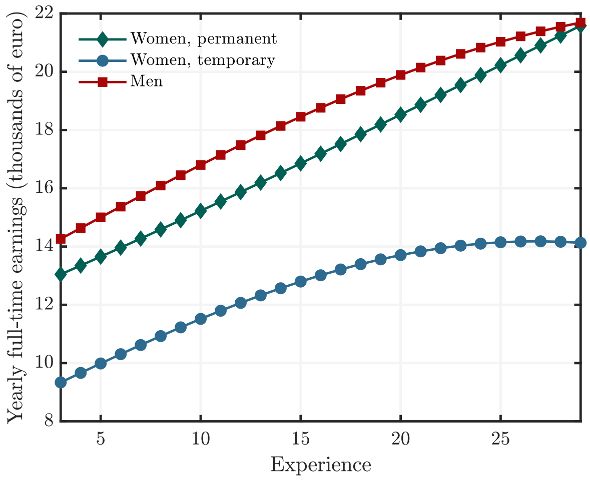
As for the idiosyncratic wage shocks, we observe that men exhibit slightly higher persistence, while women face greater income volatility. Since the MCVL does not include information on spousal wage, we adopt a standard correlation of 0.25 from

Table 2. First Stage Parameter Values

Parameter	Description	Value	Source
Income Process, Contract Type and Taxes			
η_0^f		9.029	
η_1^f	Wage equation for women	0.040	
η_2^f		-0.001	
$\Delta\eta_0^f$		0.376	
$\Delta\eta_1^f$	Wage equation for women	-0.016	
$\Delta\eta_2^f$		0.001	MCVL
η_0^m		9.484	
η_1^m	Wage equation for men	0.028	
η_2^m		0.0004	
ϕ^f	Persistence wage shocks	0.614	
ϕ^m	Persistence wage shocks	0.634	
$\sigma^2_{\nu^f}$	Variance wage shocks	0.370	
$\sigma^2_{\nu^m}$	Variance wage shocks	0.354	
ρ	Income shock correlation	0.25	Hyslop (2001)
$F\left(\epsilon_1^f, \epsilon_1^m\right)$	Distribution initial shocks	See text	ECV
π_x	Experience accumulation	0.5	See text
$\pi_z(x, j)$	Contract transition	See text	MCVL
τ_0		0.8823	
τ_1	Tax function	0.122	García-Miralles et al. (2019)
$\bar{I}$		3069.5	
$\tilde{I}$		1105.02	
Pregnancy Probabilities, Desired Fertility and Childcare Costs			
α_j	Pregnancy probability	See text	Sommer (2016)
$\mathcal{P}(N^* = n)$	Distribution desired fertility	Table 1	EdF
λ	Childcare cost	180.059	EPF
Other			
β	Discount factor	0.96	Kydland and Prescott (1982)

Hyslop (2001) and estimate the initial joint distribution of spousal shocks from the assortative mating observed in the ECV.

Figure 4. Estimated Earnings by Experience



Source: Author's work with data from the MCVL.

For part-time workers, we assume a probability of accumulating a full year of experience of 0.5. Additionally, we estimate the probability of women transitioning from temporary to permanent contracts using the MCVL. Our results show that this probability increases with labor market experience, but decreases with age.

Regarding tax liabilities, we use the function estimated by García-Miralles et al. (2019), which incorporates a nontaxable income threshold, progressivity, and average tax rates, capturing key features of the Spanish tax system.

The probability of success in pregnancy by age, α_j, follows the method in Sommer (2016), starting at 0.85 at 25 years of age and decreasing to 0.48 at 39 years of age. The probability distribution over desired number of children $\mathcal{P}(N^* = n)$ is calculated using the EdF (see Table 1).

Childcare costs are based on the EPF, where the average annual cost for full-time care in 2007 is €2000. We assume that part-time care costs half of this amount. Finally, we set the discount factor to $\beta = 0.96$, following Kydland and Prescott (1982).

5.1.2 Parameters Chosen Via the MSM

The remaining 12 parameters are calibrated by matching 12 moments from the data. The targets can be divided into three groups. The first comprises the labor force participation rates for women with and without children. The second one covers the average quantum and tempo of fertility. The third group involves the response to the cash transfers identified by González (2013). In Appendix F.2, we provide further details on the targeted moments and discuss identification.

Labor Force Participation. We include part- and full-time participation rates for three groups of women (6 targets): childless, mothers whose youngest child is 0–3 (newborn or baby), and mothers whose youngest child is 3–12 (school age). The data moments are taken from the 2004–2007 ECV of the Spanish Statistics Institute. We target the average part-time and full-time yearly rates for childless women between the ages of 25 and 51. For mothers, we target the average rates for all women with children of the respective age.

Number and Timing of Births. To compute the fertility targets, we use the EdF. In particular, we calculated the percentage of women aged 40–44 years in 2018 with 0, 1, 2, and 3 children (see Table 1) and the average age at which this group had their first child. We use this sample because these women were likely very close to their completed fertility, which is the outcome of interest in the long run. These are 4 targets in total.[20]

Baby Check. Finally, González (2013) states in the conclusions of her paper that she finds that after one year of the introduction of the baby check, the number of births increased by 6 percent, and mothers were 2–4 percentage points less likely to work after the introduction of the baby check in Spain. We take these numbers as her preferred estimates and target a 6 percent increase in births and a 3 percentage points reduction in participation. These are the last two targets.

[20]One of the fractions of women by the number of children is residual, and therefore, there are only three targets for quantum and one for the tempo of fertility.

5.2 Model's Performance

5.2.1 Targeted Moments

Table 3 shows the calibrated parameters and describes their roles in the utility function and the budget constraint. The utility of consumption is nearly logarithmic, with γ_c very close to 1. The penalty for not having a third child is 70% lower than the penalty for not having the first two. In addition, there is an extra utility from the first child, as indicated by the positive value of the fixed utility of motherhood, ζ. This result suggests that the utility cost of remaining childless exceeds the cost of having fewer children than desired, which seems reasonable.

The time cost associated with younger children is higher, while the additional cost of full-time work is consistent between the two groups of children. Finally, the model shows that working part-time results in a 37% reduction in hourly earnings compared to working full-time.

Table 3. Second Stage Parameter Values

Parameter	Description	Value
γ_c	Curvature of consumption	0.972
γ_l	Curvature of leisure	0.113
γ_N	Age-varying weight on fertility gap	26.250
δ_l	Weight on leisure	1.026
δ_{N1}	Base weight on fertility gap	0.421
δ_{N2}	Weight on fertility gap, 3rd child	0.295
ζ	Fixed utility of motherhood	0.160
ξ_1	Time cost, youngest child 0–3	0.349
ξ_2	Time cost, youngest child 3–12	0.314
κ_1	Cost of full-time work, youngest child 0–3	-0.030
κ_2	Cost of full-time work, youngest child 3–12	-0.030
ϕ	Part-time earnings penalty	0.370

Table 4 shows the model's outcomes versus the data targets. In general, we achieve our objective of closely reproducing the data targets with the calibrated model. In particular, it is important that we were able to get the correct magnitude of the effect of the baby checks on fertility. We fall slightly short of the magnitude of the effect on LFP, but we are not too far from the lower range of a drop of 2 percentage points in González (2013). The bottom line is that the effect is small and the model is qualitatively close.

Table 4. Targeted Moments

Moment	Model	Data
Labor force participation:		
Childless women:		
Part-time rate	0.121	0.122
Full-time rate	0.671	0.694
Mothers, youngest child 0–3:		
Part-time rate	0.177	0.212
Full-time rate	0.511	0.507
Mothers, youngest child 3–12:		
Part-time rate	0.175	0.198
Full-time rate	0.521	0.516
Fertility:		
Share of women with:		
0 children	0.181	0.190
1 child	0.259	0.250
2 children	0.436	0.438
3 children	0.124	0.123
Average age at first birth	31.654	29.593
Effects of cash transfers on:		
Annual number of births	6.296	6.000
Mother's LFP over the first year	-1.076	-3.000

5.2.2 Non-Targeted Moments

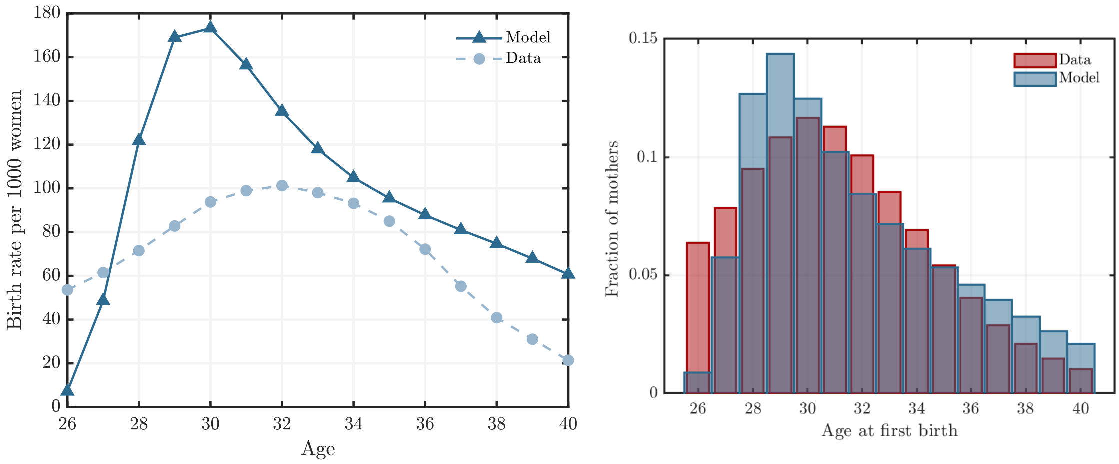
Below we discuss the fit of the model for a set of non-targeted moments. We discuss a few additional ones can be found in Online Appendix F.

Timing of Births. The overall magnitude of the response of fertility to the baby check depends on age-specific responses weighted by baseline rates. Figure 5 shows that the model delivers reasonable birth rates by age, as well as distribution of age at first birth. Thus, we can be assured that the overall effect of the baby check is not driven just by women within certain ages, but by the combined responses of women of all ages.

Effect of the Baby Check by Age in the Short Run. We extend the work in González (2013) by estimating the effect of the baby check by the mother's age. We estimate eight specifications based on RDD and DiD designs, as in the original paper. Figure 6 shows the results obtained for the specification in which the overall effect is

Figure 5. Timing of Births, Model and Data

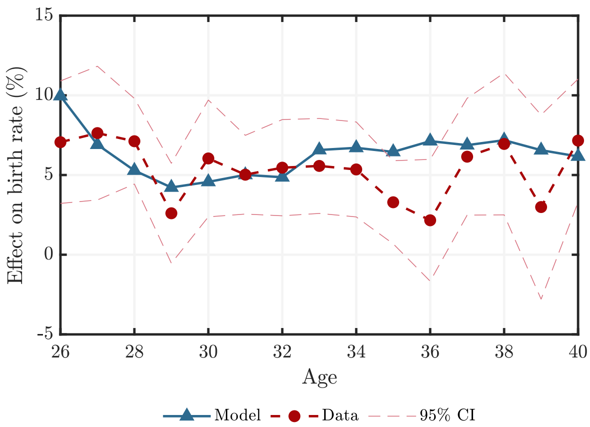
(a) Age-Specific Birth Rate **(b)** Age at First Birth



Source: For age-specific birth rates we use data from the Spanish Statistical Office in 2007. For the age at first birth we use cohorts aged 40–44 in the EdF in 2018 (aged 29-33 in 2007).

closest to our target 6%, as well as the effects predicted by our model. The details of the procedure we followed, as well as the results for the other seven specifications, can be found in online Appendix G. The main takeaway is that the short-run effects of the baby check on fertility are positive for women of all ages. The model-generated response falls within the 95% confidence interval of the data estimates for almost all ages. Thus, our model generates plausible short-term fertility responses not just in aggregate but also conditional on age.

Figure 6. Effect of the Baby Check by Age in the Short Run, Model and Data



Source: Author's work with vital statistics data by the Spanish Statistical Office.

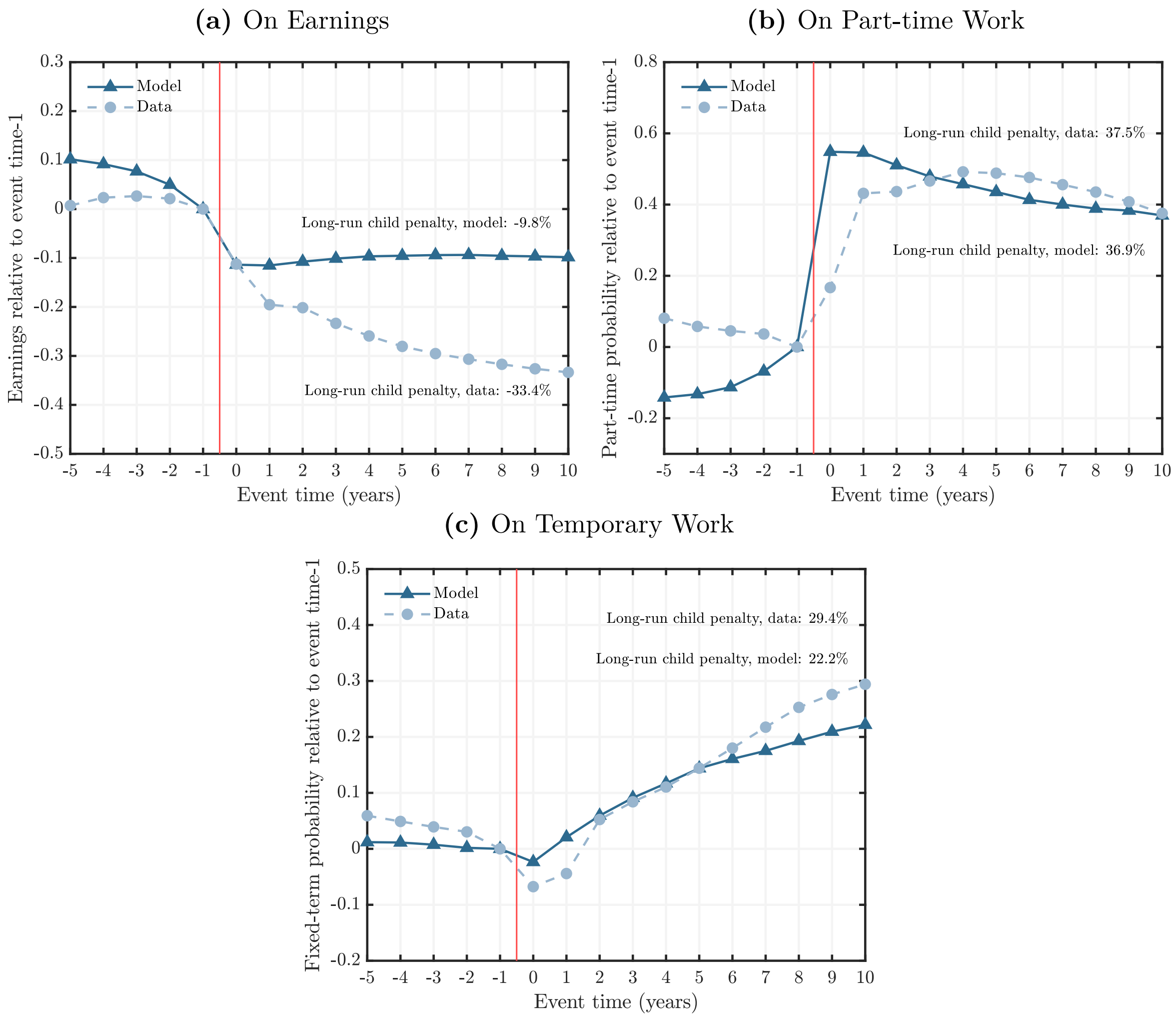
Note: The effects from the data correspond to the DiD 7-year specification.

Labor Market Child Penalties. Kleven et al. (2019) propose an event study methodology to estimate the effect children have on parents' labor market outcomes. de Quinto et al. (2021) applied this methodology to the same Spanish Social Security data that we use to estimate the parameters of our income process. Figure 7 shows their estimates of earnings, part-time and temporary probability, as well as the results obtained from the estimation of the event study with our simulated data. Qualitatively, the model displays adequate behavior in all three outcomes: earnings and part-time penalties increase rapidly in the first two years after giving birth to the first child, while the temporary penalty is negative at the beginning and then increases gradually. Moreover, the three outcomes show a long-term penalty, that is, the effects persist even 10 years after the first childbirth. The reason why earnings and part-time probability increase immediately after the first birth is straightforward: after having their first child, many women switch to part-time employment or drop out of the labor force, which means they earn less or nothing at all (the earnings child penalty is estimated unconditionally on employment status). The reason behind the shape of the penalty on the probability of being on a temporary contract is more subtle: women on permanent contracts are more likely to have children. However, over time, those who have children but are on temporary contracts convert them at a much lower rate than those who don't because many of the former take career breaks or switch to part-time. Finally, the long-run child penalty on earnings in the model is smaller than the one estimated in the data. The model captures the initial drop in earnings, but in the data, they decrease further over time. We think this is because there are factors that the model does not feature that affect earnings among mothers over time, such as loss of skills, in the spirit of (Adda et al., 2017).

Effect of Cancellation and the Role of Expectations A crucial determinant of people's behavior in response to incentives is whether they expect them to be permanent or not. In our case, if households were expecting the baby check to be phased out, it is likely that some that were planning to have children in the future anyway will have them earlier, thus affecting the fertility rate in the short run but leaving their total number unaffected.

We take advantage of how the policy was canceled to rule out such behavior. González and Trommlerová (2023) exploit the fact that the *announcement* that the

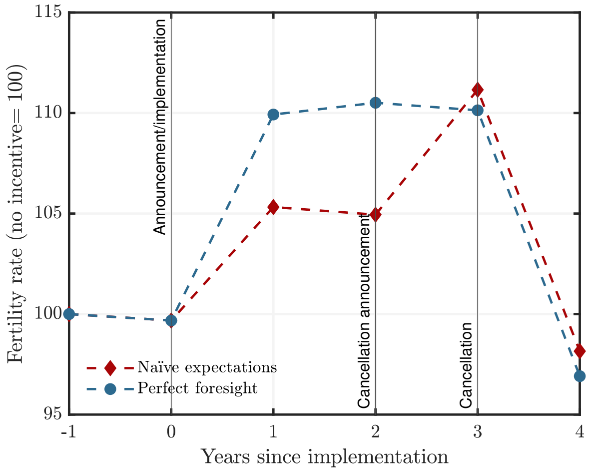
Figure 7. Child Penalties



Source: Data estimates from de Quinto et al. (2021) using the MCVL.

policy was going to be canceled occurred before the *actual* cancellation and estimate that there was a 4.7% increase in births following the announcement. We run two simulations: one in which people are taken by surprise by both the introduction and cancelation of the baby check (which corresponds to what we assume in the benchmark case and which we call "naive expectations") and one in which they are certain that the policy will be canceled in its fourth year as happened in reality (called "perfect foresight"). Figure 8 shows that, if people had perfect foresight, the effect of the incentive would have been greater, but crucially, the cancellation announcement would have had zero effect on births. Furthermore, the increase in the birth rate after the cancellation announcement under naïve expectations leads to a 5.9% increase in births, roughly consistent with the effect observed in reality. We conclude that our assumption of "naive expectations" is appropriate.

Figure 8. Introduction and Cancellation of the Baby Check



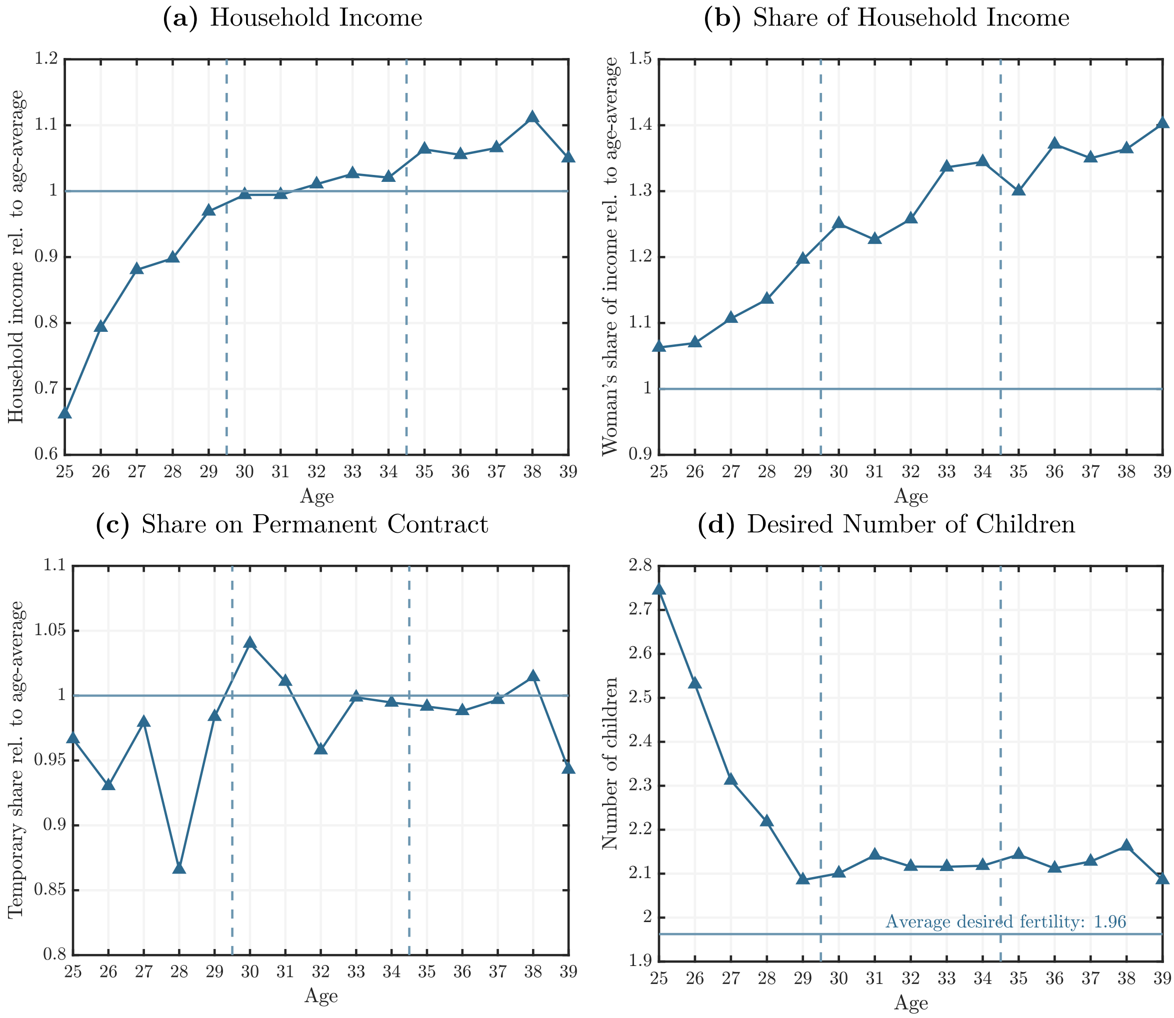
5.3 Who Responds to the Baby Check in the Short Run?

To shed more light on what drives the fertility response in the short run across women of different ages in our model, we characterize those who respond to the introduction of the baby check by having a child in the period following its introduction. Figure 9 shows four dimensions by the age of the mother: household income, women's share in household income, share in temporary work (these first three measured during the period when the baby check is announced), and desired number of children.

The main takeaways are as follows. First, women who respond to the policy contribute a relatively high share of their household income. Second, among younger women (ages 25 to 29), those with lower household incomes and very high desired fertility respond to the baby check. This is in line with classic Beckerian theory. Third, older women (aged 34 to 39) who respond to the policy tend to have above-average household incomes and contribute a relatively high share of them. The same happens when we look at lifetime income: women who respond in this age range contribute a relatively high share of their household's lifetime income. This result can be linked to the driver of procyclical fertility in Coskun and Dalgic (2024): when women's share of household income increases, fertility falls. Thus, women who are in a position of earning a relatively high share of household income are less likely to be close to their desired fertility level and, thus, more likely to respond to the baby check.

Another aspect of interest is whether it is childless or women who already have children who respond to the baby check. We find that 63.11% of the newborns that result from the policy are the first child, 23.25% are the second child, and 13.63% are

Figure 9. Characteristics of Women That Respond to the Baby Check in the Short Run by Age When it is Introduced



the third child. That is, fertility responds both in the extensive margin (being childless versus being a mother) and in the intensive margin (having additional children), consistently with González and Trommlerová (2023), who find that the baby check increased births at every parity.

6 Quantitative Experiments

In this section, we use our calibrated life-cycle model to perform a series of experiments that address our main research question: How effective are financial incentives in increasing completed fertility rates? In addition, we analyze the role of labor market duality in shaping fertility outcomes and its interaction with these incentives.

6.1 The Long-Run Impact of the Baby Check on Fertility

In our benchmark model, introducing a financial incentive at time t_0 that emulates the Spanish baby check leads to a 6% increase in the fertility rate one year later, at t_1. To understand its longer-term impact, we simulate the entire life cycle of women aged 25 to 39 years at t_0 over the next 15 years. We assume that the incentive remains in place indefinitely, and that all women expect it to.

We focus on this 15-year window because it ensures that the youngest cohort at t_0, aged 25, reaches the end of their fertility cycle at age 39. This allows us to measure the completed fertility rate for this cohort, defined as the total number of children per woman by age 40. This measure represents the long-run effect of the incentive. In our model, the completed fertility rate increases from 1.502 children per woman in the baseline scenario without the incentive to 1.546, an increase of 2.96%, which accounts for approximately 47% of the initial effect on the fertility rate at t_1. Why does this discrepancy arise?

Figure 10a illustrates the impact of the financial incentive on age-specific fertility rates. It compares two groups: the cross section of women who can have children at t_0 and the entire life cycle of women aged 25 at t_0. Notice that the two lines coincide at age 25 and remain close until around age 30, when they begin to diverge. This divergence occurs because women who had access to the incentive in earlier periods are more likely to have had additional children by that age, bringing them closer to their desired fertility and reducing the likelihood of having more children later.

To illustrate this further, consider the effect of the incentive on a 39-year-old woman at t_0 (blue) versus a 25-year-old woman at t_0 when she reaches age 39 (red). The 39-year-old woman has had only one period to respond to the policy, while the 25-year-old woman will have had 14 years of exposure. As a result, the 39-year-old is more likely to respond by having an additional child. Interestingly, the woman who had access to the incentive throughout her fertility cycle is around 3% less likely to give birth at ages 38 and 39 compared to the benchmark without the incentive. This indicates that part of the effect of the policy is due just to changes in the timing of births.

The effect on the fertility rate one year after the policy implementation is given by the weighted average of the cross-sectional, age-specific effects, while the effect on completed fertility is derived from the weighted average of life-cycle age-specific effects. In both cases, the weights are determined by the benchmark age-specific fertility rate.

Figure 10. Short and Long-Run Effect of the Financial Incentive to Fertility

(a) On Age-Specific Fertility Rate **(b)** On Yearly Fertility Rate

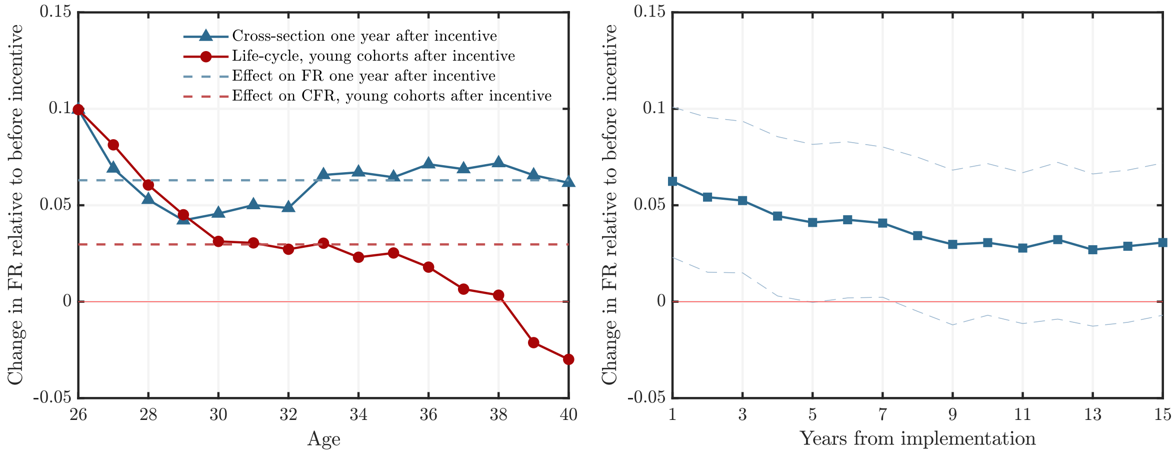


Figure 10b shows that the yearly fertility rate increases by 6.30% in the first year, that is, in t_1 compared to t_0, but then gradually declines to a level 2.96% above the pre-incentive fertility rate. This decline occurs because younger cohorts gradually replace older ones. From year 15 onward, the yearly fertility rate aligns with the completed fertility rate. This happens because, in our model, every cohort that reaches age 25 after t_0 faces the same environment and thus has children at the same rates at each age.

Finally, even though only a small fraction of women actually change the number of children they have, all receive the baby check conditional on having children, so it is possible that there may be effects on labor force participation. However, this is not the case. Among women with children aged 0–3, who may actually receive the baby check, labor force participation falls by less than 0.67 percentage points, mostly due to a reduction in full-time work. Among women with children 3–12, the reduction is less than half of that, only 0.31 percentage points. This is consistent with González (2013) finding that the baby check has no effect on labor force participation two years after birth.

6.2 The Impact of Labor Market Duality on Fertility

Temporary contracts are associated with delayed and depressed fertility. Although most of the literature exploring the effect of this type of working conditions on fertility centers on uncertainty and stability, we highlight another mechanism: returns to experience at critical ages. Most women in our benchmark scenario start their careers under a temporary contract and anticipate that their participation in the labor market may be reduced after having their first child. Therefore, they have an additional incentive to work full-time, accumulate experience, and postpone childbirth, that is, obtaining a permanent contract first.

In our model, the difference between temporary and permanent contracts is reflected in the parameters of the income process. In particular, the constant term is larger for the permanent one, reflecting that temporary workers experience unemployment spells throughout the year. In addition, the returns to experience are greater for permanent workers, possibly reflecting firm- or tenure-specific skill accumulation that temporary workers do not accumulate.

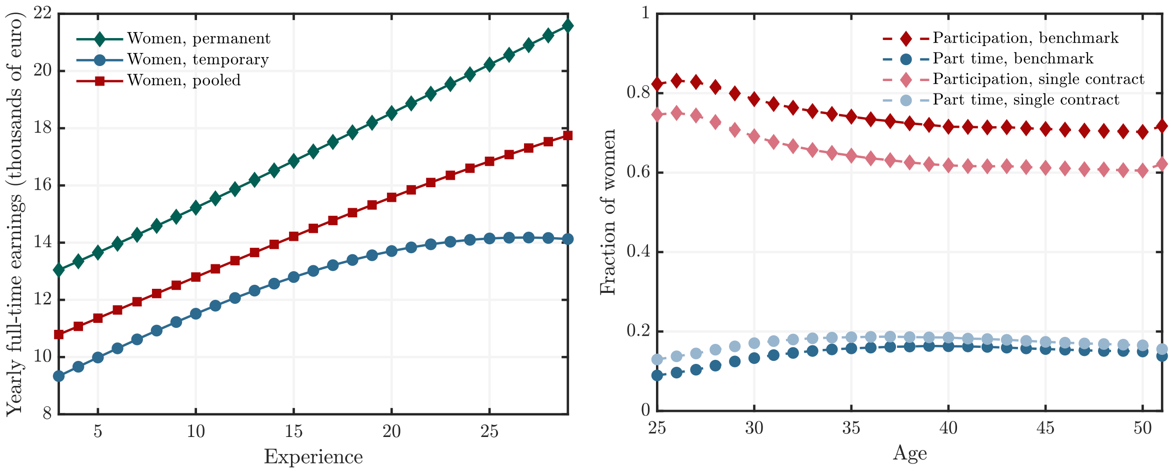
The counterfactual we propose is a single contract represented by a unique earnings equation with parameters estimated from pooling all female workers (see the Online Appendix H for more details). Among younger workers, this contract offers full-time earnings that are lower than the permanent contract, but higher than the temporary one in the benchmark scenario. Therefore, the returns to experience are not as large at the beginning of women's careers. Figure 11a shows the earnings profile by experience level under the two types of contracts present in the benchmark economy, and under the single-contract counterfactual used here.

We recompute the model solution under the new parametrization for the earnings equations, simulate the life cycle for cohorts of women that live under these labor market conditions, and compare it to the benchmark (with dual labor markets and no cash transfers). Our main result is a 6.6% increase in completed fertility rates, which is more than twice as much as the long-run increase we found results from cash transfers. Moreover, the first age at birth is anticipated by about half a year.

To better understand what drives these effects, Figure 11b shows overall and part-time labor force participation rates in the benchmark and under a single contract. Overall participation is lower throughout the life cycle, but in child-bearing years, part-time participation is higher under single contracts. It is less costly for women to reduce their market hours in the latter case, because they do not face the additional career cost of being stuck in a temporary contract.

Figure 11. The Effect of Duality on Labor Markets

(a) Earnings Profiles by Experience **(b)** Labor Force Participation



Source: Author's work using the MCVL for earnings profile by experience.

Finally, the last experiment we carry out is implementing the financial incentive on top of the single contract and repeating the simulations. In this counterfactual, the effect of the cash transfer in the period after implementation and the effect on the completed fertility rate of the 25-year-old cohort in t_0 become 5.54% and 2.26%, respectively. The ratio between short- and long-run effects is 0.41, compared to 0.47 in our benchmark. These results imply that the findings of our first experiment, i.e., the effects on fertility of giving cash transfers upon birth are smaller in the long run than in the short run and modest in magnitude, are not heavily dependent on the labor market duality featured in our benchmark model and the Spanish economy.

7 Discussion

7.1 Cost Effectiveness

Using our benchmark 6.3% increase in fertility in the short run, a simple calculation reveals that the cost per additional birth in the first year after implementation of the baby check was €42,182.[21] González and Trommlerová (2023) find that, over the course of the 37 months that the policy was in place, the cost per additional birth was at least €63,400. They explicitly recognize that this number includes mere changes in timing. Using our model, we are in a position to provide an estimated net of any timing effects: The result of our first experiment implies that the cost per additional birth, in the long run, increases to €86,959.

A natural follow-up question is how much it would cost to close the gap between the desired and the realized fertility. We compute the minimum yearly transfer that accomplishes this in the calibrated model and find it to be 2.89 times the value of the baby check, or €7,326 per child, per year, up to age 20. A back-of-the-envelope calculation estimates that the cost of this program would amount to around 8% of the Spanish GDP.[22]

7.2 Model Assumptions

In this section, we discuss how to interpret our results in light of the assumptions we make. First, we assume that the desired fertility is a deep parameter that does not change in response to policy. We think that this assumption is not unreasonable when studying the baby check, since it was a relatively small transfer in terms of household lifetime income. In Figure 3, we show that the desired fertility in similar age groups did not change dramatically between 1999 and 2018. This is suggestive evidence that the baby check did not change fertility preferences. This is less likely to be true in our calculation of the recurrent transfer that closes the gap between the desired and realized fertility presented in the previous subsection. In any case, our results in terms of the effectiveness of the baby check in the long run should be seen as a lower bound, as there can be additional effects through desired number of children.

Second, a related concern is that we take the distribution of women by their desired number of children from survey data. There are a few reasons why responses may not reflect deep preferences. First, they may be conditional on economic conditions. Second, they may be affected by choice-supportive bias, in the sense that women who

[21] If B_0 is the baseline total number of births, the cost per additional birth is given by the total cost, $2,500 \times (B_0 \times 1.063)$, divided by the additional number of births, $B_0 \times 0.063$.

[22] González and Trommlerová (2023) reports that the Spanish government paid 491,577 baby checks in 2008. We multiply this by 1.25, which is the increase in fertility, and then by the amount of the transfer, and divide by the GDP of Spain in 2008.

really wanted fewer children than they ended up having may rationalize the choice taken ex post by saying they wanted a higher number all along. Third, they may be affected by rosy retrospection bias, in the sense that women may say that they would have liked to have more children because they do not fully recall the difficulties of raising the ones they had. However, we believe that it is still a powerful fact that a large share of women declare wanting more children at the end of their fertile years than they had, indicating there is scope for pro-nationalist policies to work.

Third, another important assumption we make is that women are the sole decision makers in the household. Introducing intra-household bargaining would be challenging in our rich setting. The state-of-the-art research on bargaining over fertility is Doepke and Kindermann (2019). They find that births are likely only if both partners desire an additional child and that in low-fertility countries, women are more likely to be opposed to having one. Moreover, in the Online Appendix D, we show that men's and women's preferences are similar, and that the gap between them decreases with age.[23] In light of this, our assumption of women being the decision makers does not seem inadequate.

Finally, households in our model do not have the ability to save or borrow. Adda et al. (2017) do find that, in their model, household saving rates increase in anticipation of a birth. This would provide an additional channel through which the baby check may have an effect. However, there are a few reasons to believe that savings and assets are not crucial to understanding the effect of the baby check. First, the transfer is relatively small, so it has a negligible impact on lifetime income. Second, if households are credit-constrained, the behavior of younger ones should be close to hand-to-mouth if earnings profiles are increasing with age. In fact, the fact that the baby check had any effect on fertility provides indicative evidence that households are credit-constrained. For older households, the behavior in the absence of savings may differ. However, our model delivers effects by age that are not too far away from the data for this age group in the short run. Thus, the role of savings in the effect of the baby check is likely small. For policies involving larger incentives, considering savings is probably necessary.

Overall, we believe that, while including some of the features mentioned above would provide opportunities to study interesting interactions, the main results of our

[23] For individuals aged 40-44, the gaps between men and women are 2.84, -0.26, 3.25, and -5.05 percentage points (pp) for wanting 0, 1, 2, or 3+ children, respectively.

paper would not change substantially.

8 Conclusions

This paper studies the effects of financial incentives on fertility, differentiating between short and long run. Short-run effects arise from the responses of a cross section of women of varying ages when incentives are introduced. There is empirical evidence of their magnitude. The long-run effects arise from life-cycle decisions of women who receive the incentive from the beginning to the end of their childbearing years. Identifying them empirically is challenging.

To bridge this gap, we develop a life-cycle model of labor force participation and fertility decisions, calibrated to replicate the short-run effects of Spain's 2007 baby check, a cash transfer paid at childbirth. Our findings reveal that the long run impact of the incentive is about half the short run effect, with completed fertility increasing by 3% compared to an initial 6% rise in annual fertility. This occurs because women who are young when the incentive is introduced do not respond in the same way later in life as those who are older when the incentive is given for the first time. The former respond by having additional births earlier in life and fewer later, while the latter, constrained by fewer reproductive years left, have additional births later in life.

Labor market conditions also play a crucial role in shaping fertility. Spain's labor market duality has been shown to negatively affect fertility. Simulating an alternative single-contract labor market scenario reveals a significant 6.6% increase in completed fertility, more than twice the long-run effect of cash transfers. Addressing structural labor market challenges, such as the prevalence of temporary employment, could substantially increase fertility.

Even under improved labor market conditions, the long-run effects of the baby check remain modest and smaller than the short-run effects. This suggests that financial incentives alone are insufficient to induce large and sustained fertility increases. Other policies aiming to improve family-work balance, reduce job instability, and provide consistent support throughout women's life cycles should be carefully studied. We leave that for future research.

Acknowledgments

We thank Andrés Erosa, Luisa Fuster, Nezih Guner, Timothy Kehoe, John Knowles, Zöe Kuehn, Matthias Kredler, Hamish Low, Alessio Moro, Rachel Ngai, Virginia Sánchez-Marcos, and participants at seminars at the University of Cagliari, Universidad Carlos III de Madrid, Goethe University Frankfurt, and Universitat de Barcelona, as well as attendees at the Annual ESPE Conference, the Brazilian Meeting in Family and Gender Economics, the CRENoS Workshop, the Symposium of the Spanish Economic Association, the Frankfurt–Mannheim Macro Workshop, the SOLE Meeting, the SED Meeting, the Spanish Labour Economics Meeting, the Lisbon Macro Workshop and the Princeton Faculty Macro Lunch for their valuable feedback. Javier thanks the Centre for Macroeconomics at LSE and the International Economics Section at Princeton University, and Lidia the Department of Economics at the University of Minnesota for their hospitality during academic visits while working on this project.

Funding

This work was financially supported by the Spanish Ministry of Science, Innovation and Universities, Spain [grant numbers PID2020-114040RB-I00 and PID2019-107614GB-I00]; the Catalan Research Agency [grant number 2021SGR00678]; and the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement number 101095175 (SUSTAINWELL project). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the UK can be held responsible for them.

References

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

Adserà, A. (2004). Changing fertility rates in developed countries. the impact of labor market institutions. *Journal of Population Economics*, 17(1):17–43.

Aksoy, Y., Basso, H. S., Smith, R. P., and Grasl, T. (2019). Demographic structure and macroeconomic trends. *American Economic Journal: Macroeconomics*, 11(1):193–222.

Attanasio, O., Low, H., and Sánchez-Marcos, V. (2008). Explaining Changes in Female Labor Supply in a Life-Cycle Model. *American Economic Review*, 98(4):1517–52.

Auer, W. and Danzer, N. (2016). Fixed-Term Employment and Fertility: Evidence from German Micro Data. *CESifo Economic Studies*, 62(4):595–623.

Bick, A. (2016). The quantitative role of child care for female labor force participation and fertility. *Journal of the European Economic Association*, 14(3):639–668.

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

Coskun, S. and Dalgic, H. C. (2024). The emergence of procyclical fertility: The role of breadwinner women. *Journal of Monetary Economics*, 142:103523.

De la Rica, S. and Iza, A. (2005). Career Planning in Spain: Do Fixed-term Contracts Delay Marriage and Parenthood? *Review of Economics of the Household*, 3(1):49–73.

De Nardi, M., İmrohoroğlu, S., and Sargent, T. J. (1999). Projected u.s. demographics and social security. *Review of Economic Dynamics*, 2(3):575–615.

de Quinto, A., Hospido, L., and Sanz, C. (2021). The child penalty: evidence from Spain. *SERIEs: Journal of the Spanish Economic Association*, 12(4):585–606.

Doepke, M., Hannusch, A., Kindermann, F., and Tertilt, M. (2022). The economics of fertility: A new era. Working Paper 29948, National Bureau of Economic Research.

Doepke, M. and Kindermann, F. (2019). Bargaining over babies: Theory, evidence, and policy implications. *American Economic Review*, 109(9):3264–3306.

Erosa, A., Fuster, L., and Restuccia, D. (2010). A General Equilibrium Analysis of Parental Leave Policies. *Review of Economic Dynamics*, 13(4):742–758.

Garcia-Louzao, J., Hospido, L., and Ruggieri, A. (2023). Dual returns to experience. *Labour Economics*, 80:102290.

García-Miralles, E., Guner, N., and Ramos, R. (2019). The Spanish personal income tax: facts and parametric estimates. *SERIEs: Journal of the Spanish Economic Association*, 10(3):439–477.

González, L. and Trommlerová, S. K. (2023). Cash transfers and fertility: How the introduction and cancellation of a child benefit affected births and abortions. *Journal of Human Resources*, 58(3):783–818.

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.

Guner, N., Kaya, E., and Sánchez-Marcos, V. (2024). Labor market institutions and fertility. *International Economic Review*.

Guner, N., Kaygusuz, R., and Ventura, G. (2020). Child-Related Transfers, Household Labour Supply, and Welfare. *The Review of Economic Studies*, 87(5):2290–2321.

Guvenen, F., Pistaferri, L., and Violante, G. L. (2022). Global trends in income inequality and income dynamics: New insights from grid. *Quantitative Economics*, 13(4):1321–1360.

Hyslop, D. R. (2001). Rising u.s. earnings inequality and family labor supply: The covariance structure of intrafamily earnings. *American Economic Review*, 91(4):755–777.

Iskhakov, F., Jørgensen, T. H., Rust, J., and Schjerning, B. (2017). The endogenous grid method for discrete-continuous dynamic choice models with (or without) taste shocks. *Quantitative Economics*, 8(2):317–365.

Jones, C. I. (2022). The End of Economic Growth? Unintended Consequences of a Declining Population. *American Economic Review*, 112(11):3489–3527.

Jones, L. E., Schoonbroodt, A., and Tertilt, M. (2010). Fertility theories: Can they explain the negative fertility- income relationship? In Shoven, J. B., editor, *Demography and the Economy*, pages 43–106. University of Chicago Press, Chicago.

Kim, W. (2024). Baby bonus, fertility, and missing women. *Journal of Human Resources*.

Kim, Y.-I. A. (2014). Lifetime impact of cash transfer on fertility. *Canadian Studies in Population [ARCHIVES]*, 41(1-2):97–110.

Kleven, H., Landais, C., and Søgaard, J. E. (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.

Kohler, H.-P., Billari, F. C., and Ortega, J. A. (2002). The emergence of lowest-low fertility in europe during the 1990s. *Population and Development Review*, 28(4):641–680.

Kydland, F. E. and Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370.

Lopes, M. C. (2019). Job security and fertility decisions. mimeo, Universidad Carlos III de Madrid.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*. Academic Press, New York.

Milligan, K. (2005). Subsidizing the stork: New evidence on tax incentives and fertility. *The Review of Economics and Statistics*, 87(3):539–555.

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.

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

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

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*, 20(2):177–181.

Tauchen, G. and Hussey, R. (1991). Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica*, 59(2):371–396.

Trussell, J. and Wilson, C. (1985). Sterility in a population with natural fertility. *Population Studies*, 39(2):269 – 286. Cited by: 60.

United Nations (2021). World population policies 2021: Policies related to fertility. Technical report, Department of Economic and Social Affairs, Population Division.

United Nations (2022). World Population Prospects. Department of Economic and Social Affairs, Population Division.

Financial Incentives to Fertility: From Short to Long Run

Lidia Cruces and F. Javier Rodríguez-Román

December 23, 2024

Online Appendix

A Relationship Between the TFR and the CFR

The **total fertility rate (TFR)** refers to the average number of children a woman would have over her lifetime, based on current age-specific birth rates. In contrast, the **completed fertility rate (CFR)** measures the total number of children a woman has at the end of her reproductive years.

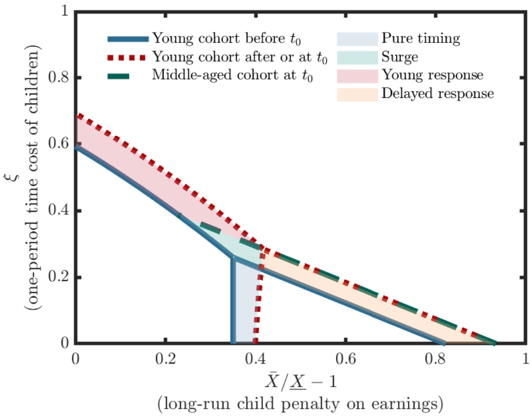
The TFR is reported more frequently because it can be calculated annually using data from women of reproductive age, whereas the CFR becomes observable only after a cohort has completed childbearing. When the TFR remains stable over several years, it often serves as a reasonable approximation of the CFR. However, sustained TFR levels below 2 typically indicate CFRs below 2 for the corresponding cohorts. It is important to note that the TFR can vary significantly from year to year without a corresponding change in cohort CFRs, as such fluctuations often reflect shifts in the timing of fertility rather than changes in lifetime fertility.

B Stylized Model: Further Discussion

In this section, we complement the discussion of the short- and long-run effects of introducing a cash transfer in the stylized model presented in Section 2. In particular, we show two additional cases in Figure B.1, the "Delayed Response" and the "Young Response" areas.

Let P_t^y and P_t^m be the measures of the young and old cohort at time t. If the fraction ϵ of households choose childlessness before t_0, the number of children born per

Figure B.1. Decision Boundaries: Before and After the Introduction of the Cash Transfer



time period is $P_t^y(1-\epsilon)N/2 + P_t^m(1-\epsilon)N/2$, and thus the fertility rate is $(1-\epsilon)N/2$. After the introduction of the financial incentive at t_0, it could be that the number of children increases by $P_{t_0}^y\epsilon$, $P_{t_0}^m\epsilon$ or $(P_{t_0}^y + P_{t_0}^m)\epsilon$, depending on whether only the young, only the middle-aged, or both cohorts respond at t_0. The corresponding increases in fertility rates are given by $\frac{P_{t_0}^y}{P_{t_0}^y+P_{t_0}^m}\epsilon$, $\frac{P_{t_0}^m}{P_{t_0}^y+P_{t_0}^m}\epsilon$ and ϵ. Letting $S_{t_0}^y = \frac{P_{t_0}^y}{P_{t_0}^y+P_{t_0}^m}$ and $S_{t_0}^m = \frac{P_{t_0}^m}{P_{t_0}^y+P_{t_0}^m}$, we can write the percent increase in the period fertility rates as $\frac{2S_{t_0}^y\epsilon}{(1-\epsilon)N}$, $\frac{2S_{t_0}^m\epsilon}{(1-\epsilon)N}$ or $\frac{2\epsilon}{(1-\epsilon)N}$.

On the other hand, the percent change in completed fertility rates for any given cohort is necessarily zero or $\frac{\epsilon}{(1-\epsilon)N}$, the latter being the case in which the fraction ϵ of households change their decision from childlessness to early or late parenthood after the incentive to fertility is introduced.

In the following, we discuss the two additional shaded regions in Figure B.1.

Delayed Response. Before t_0, the optimal choice was childlessness. However, at t_0, both young and middle-aged households would optimally choose late parenthood. As a result, the completed fertility rate increases from $(1-\epsilon)N$ to $(1-\epsilon)N+\epsilon$ children per woman, representing a percentage increase of $\frac{\epsilon}{N(1-\epsilon)}$. The percentage increase in short-run fertility is $\frac{2S_{t_0}^m\epsilon}{(1-\epsilon)N}$, where $S_{t_0}^m$ is the share of the middle cohort in the fertile population at t_0 (young and middle). Therefore, if the middle cohort is smaller (larger) than the young cohort, the percentage increase in fertility in the short run will be smaller (larger) than in the long run.

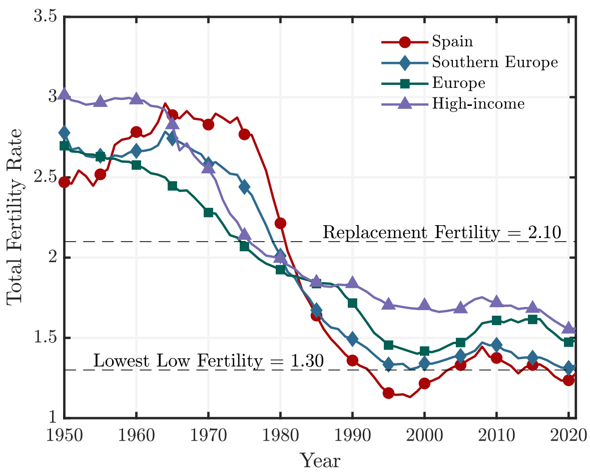
Young Response. The young cohort moves from childlessness to early parenthood, while the middle-aged one decides to remain childless. Similarly to the delayed response, the long-run response is $\frac{\epsilon}{N(1-\epsilon)}$, while the short-run response is given by $\frac{2S^y_{t_0}\epsilon}{(1-\epsilon)N}$, where $S^y_{t_0}$ is the share of the young cohort in the fertile population at t_0 (young and middle). Thus, if the size of the young cohort is smaller (larger) than that of the middle-aged one, the percentage increase in fertility in the short run is smaller (larger) than in the long run.

C Institutional Background

C.1 Spain's Demographic Context

In the early 1990s, Spain became one of the first countries in the world to attain what demographers call lowest-low fertility, that is, a total fertility rate below 1.3 (Kohler et al., 2002). A group of countries, mostly in Southern, Central, and Eastern Europe, followed. Figure C.1 shows how Spain experienced a relatively late baby boom in the 1960s and 1970s, during which its TFR increased above that of others in its geographic proximity and that of the rest of high-income countries[24]. However, it fell rapidly in the 1980s, dipping below the lowest-low fertility threshold for three decades between 1990 and 2021, and below that of its peers for the entirety of that period.

Figure C.1. Total Fertility Rate, Selected Countries (1950-2021)

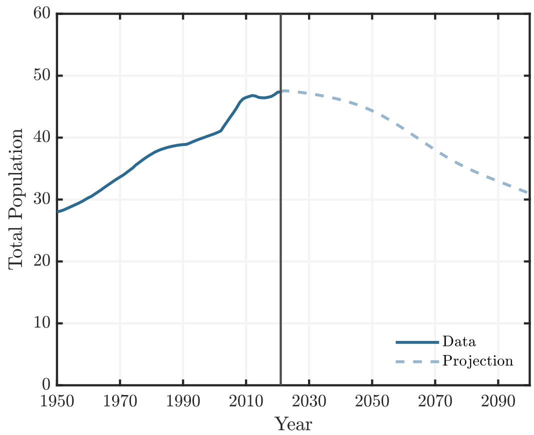


Source: United Nations (2022).

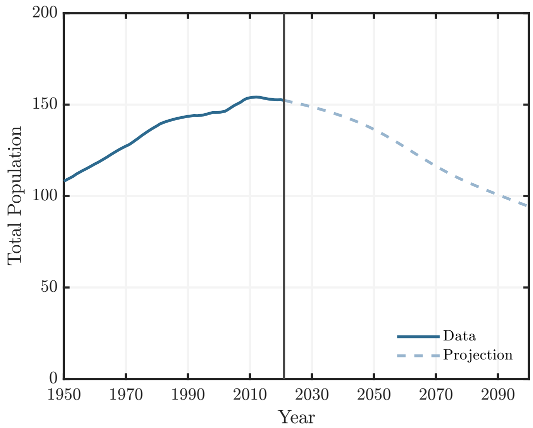
[24]This Spanish baby-boom occurred later than that of other high-income countries, like the United States, where the baby-boom took place right after WWII and lasted until the mid-1960s.

Moreover, since the early 1980s, most of the developed world has had a TFR below the replacement rate of 2.1. This means that, in the absence of immigration, that population will eventually decline. Figure C.2 shows the estimated and projected population in the countries for the same group of countries. In all of them, the population peak is likely very close or has already passed. In the case of Spain, it is apparent that the population would have already peaked if it were not for the large influx of immigrants (mainly from Latin America) it received in the 2000s. In fact, considering the very low fertility levels, in the absence of large-scale immigration, it is all but certain that most of the countries considered here will see population decline soon.

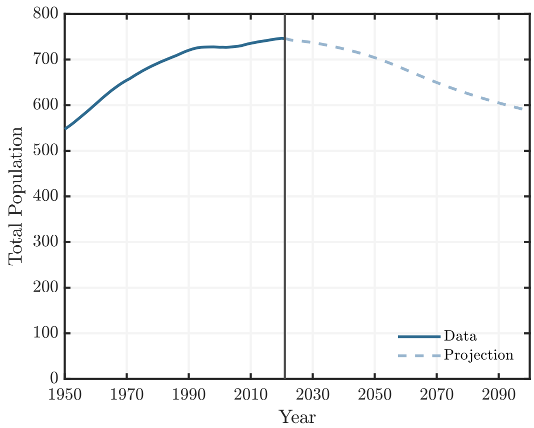
Figure C.2. Total Population, Selected Countries (1950-2022)



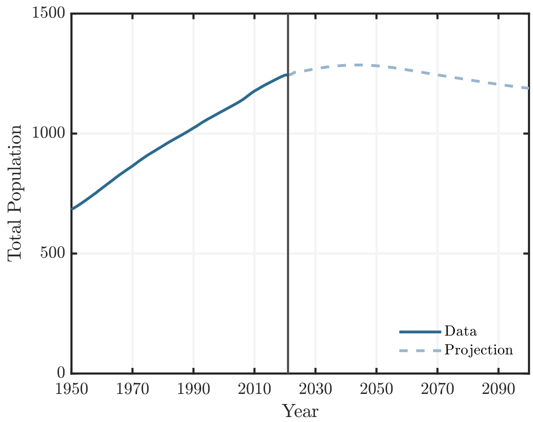
(a) Spain



(b) Southern Europe



(c) Europe



(d) High Income

Source: United Nations (2022).

Note: Projections account for fertility, mortality, and migration, and we use the medium scenario. The total population is measured in thousands.

All said, Spain's demographic situation, while a bit more extreme than that of other similar countries, is still comparable in broad terms: a TFR below replacement, in all likelihood, will lead to population aging and decline.

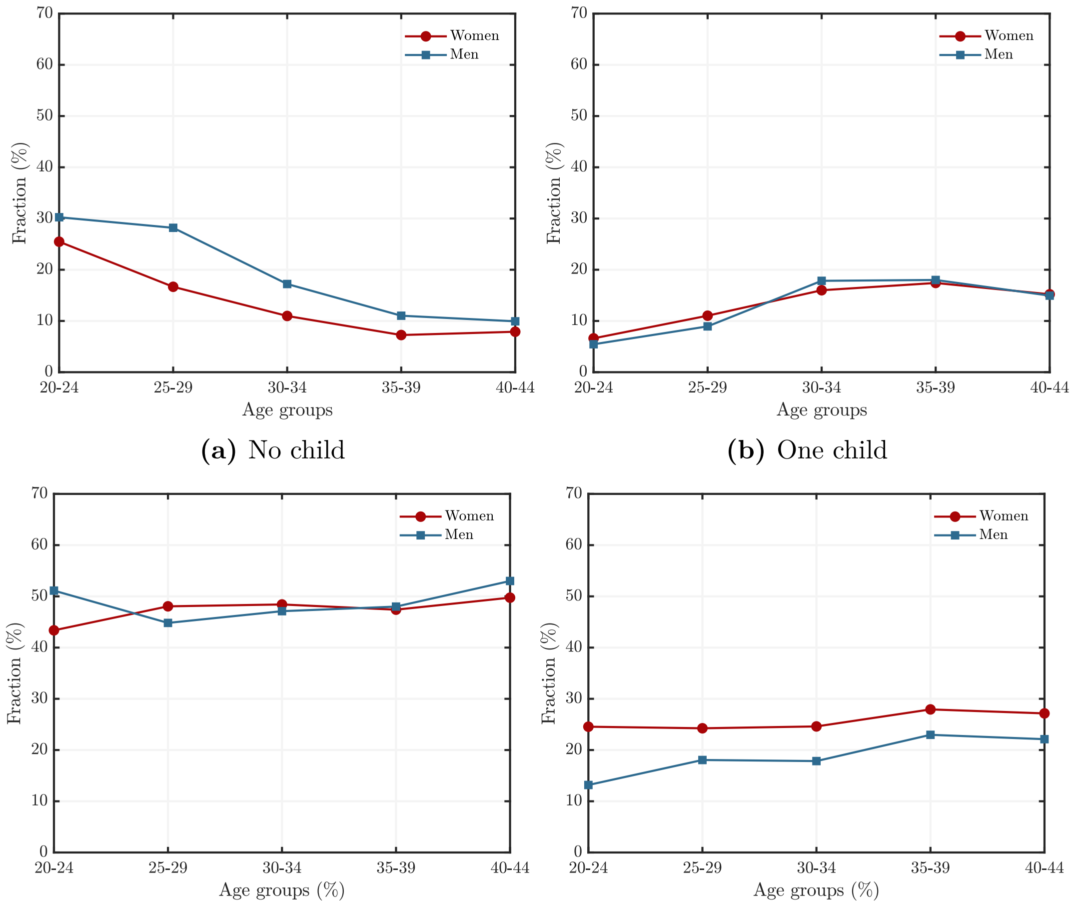
D *Encuesta de Fecundidad* (Fertility Survey, EdF)

We use the 2018 Fertility Survey (*Encuesta de Fecundidad*) to compute the gap between desired and realized fertility. The 2018 survey is a longitudinal study that includes both men and women aged 18 to 55, living in family households across Spain. It follows the 1977, 1985, and 1999 surveys, which initially focused only on women, but the 2018 wave also expanded to include men. The survey asks individuals about their ideal number of children and whether their current family size aligns with their desired family size. If there is a discrepancy between realized and desired fertility, respondents are asked to provide reasons for having fewer children than they initially wanted. These reasons include factors such as age, relationship status, health, economic conditions, work-life balance, and other personal factors. For the main analysis, we select women aged 40-44, as this group is very likely to have completed their fertility.

To assess the robustness of our choice, we compare in Figure D.1 the desired level of fertility between men and women for different age groups in 2018. The most notable discrepancies are observed in the proportions of individuals wanting no children (higher for men) and those wanting three or more children (higher for women than for men). However, in aggregate terms, the disagreement is minor and decreases with age. For individuals aged 40-44, the gaps between men and women are 2.84, -0.26, 3.25, and -5.05 percentage points (pp) for wanting 0, 1, 2, or 3+ children, respectively.

Figure D.2 shows the gap between the desired and realized fertility in 2018 by household monthly income. As can be seen, the gap remains relatively constant across the income distribution.

Figure D.1. Desired Fertility by Gender and Age in Spain



(a) No child **(b)** One child

(c) Two children **(d)** Three or more children

Source: *Encuesta de Fecundidad 2018, INE.*

E Full Life-Cycle Model

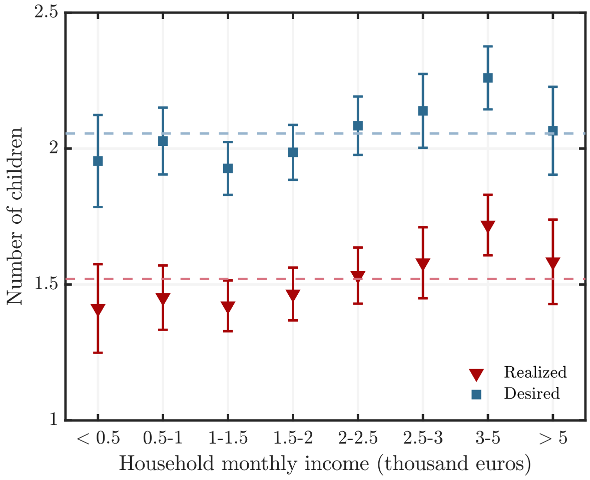
E.1 Solving the Model Numerically

E.1.1 Stochastic Structure for Children's Aging

In this subsection, we show how to retrieve the stochastic structure that governs the transition probabilities for the vector state of the number of children of different ages, which we denote by $\mathbf{n}' = \Lambda_j(\mathbf{n}, b)$.

We denote by λ_1 and λ_2 the probabilities that an individual baby becomes a school-age child, and that an individual school-age child becomes a teen in a given period, respectively. Moreover, we assume that the aging event is independent across

Figure D.2. Women's Desired and Realized Fertility by Monthly Household Income



Source: *Encuesta de Fecundidad 2018, INE.*

Note: We restrict the sample to women aged 40-44. Spikes represent 95% confidence intervals. Dashed lines represent the average realized and desired fertility (1.54 and 2.04, respectively).

children. Notice that $b\alpha_j$ is the probability that there is a newborn in the next period. Denote by $P_i(x \mid n_i)$ the probability that x children in stage $i \in \{1, 2\}$ (babies, school age) move on to the next stage the next period (school age, teenager), conditional on there being n_i children in that stage in the current period. Table E.1 shows these for babies and school-age children.

Table E.1. Probabilities of aging by number of children, $P_i(x \mid n_i)$

n_i	Number of children aging			
	0	1	2	3
0	1	0	0	0
1	$1-\lambda_i$	λ_i	0	0
2	$(1-\lambda_i)^2$	$\lambda_i(1-\lambda_i)$	λ_i^2	0
3	λ_i^3	$\lambda_i(1-\lambda_i)^2$	$\lambda_i^2(1-\lambda_i)$	λ_i^3

To compute the whole set of probabilities of transition from state $\mathbf{n} = [n_0, n_1, n_2, n_3]$, we follow the algorithm:

E.1.2 Computation With Taste Shocks

All the decisions women make in the model are discrete. Unfortunately, such models often produce abrupt aggregate responses to parameter changes as individuals tend to

for $x_1 \in \{0, 1, 2, 3\}$ **do**
  for $x_2 \in \{0, 1, 2, 3\}$ **do**
    if $n_0 = 1$ **then**
      $\mathbf{n}' = [0, n_1 - x + 1, n_2 + x - y, n_3 + y]$ w.p. $(1 - b\alpha_j)P_1(x_1 \mid n_1)P_2(x_2 \mid n_2)$ or $\mathbf{n}' = [1, n_1 - x + 1, n_2 + x - y, n_3 + y]$ w.p. $b\alpha_j P_1(x_1 \mid n_1)P_2(x_2 \mid n_2)$;
    else
      $\mathbf{n}' = [0, n_1 - x, n_2 + x - y, n_3 + y]$ w.p. $(1 - b\alpha_j)P_1(x_1 \mid n_1)P_2(x_2 \mid n_2)$ or $\mathbf{n}' = [1, n_1 - x, n_2 + x - y, n_3 + y]$ w.p. $b\alpha_j P_1(x_1 \mid n_1)P_2(x_2 \mid n_2)$;
    end
  end
end

adjust their decisions simultaneously unless there is substantial heterogeneity.

To facilitate the numerical solution of the model, we include a taste shock to women's utility in every period. This helps smooth out labor force participation and fertility decisions. The shocks can be interpreted as unobserved state variables that add noise to the women's decisions. Moreover, the calibration and results are robust to their inclusion. For an in-depth discussion of this computational method, see Iskhakov et al. (2017).

Thus, we assume that in all periods, women receive a vector of additive-separable taste shocks μ. In periods when they can still have children and need to choose pregnancy $b \in \{0, 1\}$ in addition to participation in the labor force $h \in \{0, \frac{1}{4}, \frac{1}{2}\}$, they receive a vector of six shocks, one for each element in $\{0, \frac{1}{4}, \frac{1}{2}\} \times \{0, 1\}$. In periods when they cannot have any more children and need only to choose labor force participation, they receive a vector of three shocks, one for every element in $\{0, \frac{1}{4}, \frac{1}{2}\}$:

$$\mu = \begin{cases} \left(\mu_{0,0}, \mu_{\frac{1}{4},0}, \mu_{\frac{1}{2},0}, \mu_{0,1}, \mu_{\frac{1}{4},1}, \mu_{\frac{1}{2},1}\right) & \text{if } j < 15 \text{ and } N(\mathbf{n}) < 3 \\ \left(\mu_0, \mu_{\frac{1}{4}}, \mu_{\frac{1}{2}}\right) & \text{otherwise} \end{cases} \tag{1}$$

All of these shocks are i.i.d., drawn from an Extreme Value Type I distribution with scale parameter σ_μ. The modified value function in states $\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)$ is:

$$W_j\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}, \mu; N^*\right) = \begin{cases} \max\{W_j^{h,b}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) + \sigma_\mu \mu_{h,b}\}_{h\in\{0,\frac{1}{4},\frac{1}{2}\}, b\in\{0,1\}} & \text{if } j < 15 \text{ and } N(\mathbf{n}) < 3 \\ \max\{W_j^h\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) + \mu_h\}_{h\in\{0,\frac{1}{4},\frac{1}{2}\}} & \text{otherwise,} \end{cases} \tag{2}$$

where $W_j^{h,b}$ and W_j^h represent the value, ex-taste shock, of choosing labor force participation h and pregnancy status b for a woman in period j, or just labor force participation h, in states $\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)$:

$$\begin{aligned} W_j^{h,b}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) &= u^{h,b}\left(c, l, \mathbf{n}; N^*\right) + \beta \mathbb{E}^{\sigma_\mu}\left[W_{j+1}\left(\epsilon^{f\prime}, \epsilon^{m\prime}, z', x', \mathbf{n}', \mu; N^*\right)\right] \\ W_j\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) &= u^h\left(c, l, \mathbf{n}; N^*\right) + \beta \mathbb{E}^{\sigma_\mu}\left[W_{j+1}\left(\epsilon^{f\prime}, \epsilon^{m\prime}, z', x', \mathbf{n}', \mu; N^*\right)\right]. \end{aligned} \tag{3}$$

$\mathbb{E}^{\sigma_\mu}$ denotes the expectations over future taste shocks, and in both cases, choice and state variables need to be retrieved from the constraints and laws of motion:

$$\begin{aligned} & l = 1 - h - \xi(\mathbf{n}) \\ & c = I - \lambda 2h\left(n_1 + n_2\right) \\ & I = y^m + 2hy^f\left(1 - \mathbb{1}_{\{h=\frac{1}{2}\}}\phi\right) - T\left(y^m, y^f, h\right) \\ & \ln\left(y^f\right) = \eta_0^f + \Delta\eta_0^f \mathbb{1}_{z=1} + \left(\eta_1^f + \Delta\eta_1^f \mathbb{1}_{\{z=1\}}\right)x + \left(\eta_2^f + \Delta\eta_2^f \mathbb{1}_{\{z=1\}}\right)x^2 + \epsilon^f \\ & \ln\left(y^m\right) = \eta_0^m + \eta_1^m\left(j-1\right) + \eta_2^m\left(j-1\right)^2 + \epsilon^m \\ & \epsilon^{f\prime} = \phi^f \epsilon^f + \nu^f \\ & \epsilon^{m\prime} = \phi^m \epsilon^m + \nu^m \\ & \mathbf{n}' = \Lambda_j(\mathbf{n}, b) \\ & x' = \Pi_x\left(x, h\right) \\ & z' = \Pi_z\left(x, h, z\right). \end{aligned} \tag{4}$$

The main consequence of introducing the taste shocks is that the policy function becomes probabilistic. Given the distribution assumed for them, the probability that a woman chooses pregnancy decision b and labor force participation h in states

$\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)$ when $j < 15$ and $N(\mathbf{n}) < 3$ is the logit probability:

$$P_j\left(h, b \mid \epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) = \frac{\exp\left(\frac{W_j^{h,b}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)}{\sigma_\mu}\right)}{\sum_{i\in\{0,1,2,3\}} \sum_{k\in\{0,1\}} \exp\left(\frac{W_j^{i,k}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)}{\sigma_\mu}\right)}. \quad (5)$$

Otherwise, the probability that a woman chooses labor force participation h in states $\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)$ is the logit probability:

$$P_j\left(h \mid \epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) = \frac{\exp\left(\frac{W_j^{h}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)}{\sigma_\mu}\right)}{\sum_{i\in\{0,1,2,3\}} \exp\left(\frac{W_j^{i}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)}{\sigma_\mu}\right)}. \quad (6)$$

An additional benefit of using Extreme Value Type I shocks is that the expected value function is given by the tractable log-sum formula from (McFadden, 1973):

$$\mathbb{E}^{\sigma_\mu}\left[W_{j+1}\left(\epsilon^{f\prime}, \epsilon^{m\prime}, z', x', \mathbf{n}', \mu; N^*\right)\right] =$$
$$\begin{cases} \sigma_\mu \log\left(\sum_{i\in\{0,1,2,3\}} \sum_{k\in\{0,1\}} \exp\left(\frac{W_j^{i,k}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)}{\sigma_\mu}\right)\right) & \text{if } j < 15 \text{ and } N(\mathbf{n}) < 3 \\ \sigma_\mu \log\left(\sum_{i\in\{0,1,2,3\}} \exp\left(\frac{W_j^{i}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)}{\sigma_\mu}\right)\right) & \text{otherwise.} \end{cases} \quad (7)$$

Using backward induction starting in period J, one can easily retrieve the expected value functions and the probabilistic policy functions.

F Calibration

F.1 First Step: Directly Estimated Parameters

F.1.1 Income Process

Data and Sample. To estimate the parameters that govern the income process, we use Spanish administrative data, specifically the *Muestra Continua de Vidas Laborales* (MCVL, Continuous Working Life Sample). This dataset is a 4% random sample of individuals registered with the Spanish Social Security system, covering the years from 2004 to 2019. The unit of observation in the MCVL is any change in an individual's employment status or a variation in job characteristics, such as changes in occupation or contractual conditions within the same firm.

The MCVL provides comprehensive labor market histories, including both employment and unemployment spells, starting from the first recorded employment date (or 1980 for older cohorts) and continuing to the present. For each labor market spell, the data includes information on part-time or full-time status, contract type (temporary or permanent), sector, industry, occupational skill category, working hours, and monthly labor earnings. The data has two types of earnings: non-censored earnings derived from tax records (excluding two regions: Navarra and País Vasco) and censored earnings from Social Security records. When both are available, we prioritize earnings from tax records and use the Social Security earnings as a substitute when tax data is unavailable (i.e., for the other two regions). In addition, individual-level characteristics such as age, gender, country of birth, nationality, and educational attainment are available in the dataset.

We construct a yearly panel dataset by assigning a single "main job" to each individual each year, following the approach described in Guner et al. (2024). For individuals with a unique spell during the year, this assignment is straightforward, comprising 82.14% of the sample. For individuals with multiple jobs within a year, we apply the following rules: (i) select the job with the longest duration within the year (based on days worked); (ii) if multiple jobs have the same duration, choose the job with the longest total duration across all years. Job characteristics for the year are aggregated using the modal values, while earnings represent total yearly earnings, and days worked also represent the total days worked.

Our analysis focuses on the period from 1998 to 2019, as reliable reporting of contract types (open-ended versus fixed-term) became available only after 1998. The sample consists of native individuals born between 1964 and 1985, who were between 22 and 45 years old during the implementation of the baby check in Spain. The analysis is further limited to individuals under 55 years old who are employed and earn at least three months of minimum wage annually (this wage restriction is the same as in Guvenen et al. (2022)). All earnings are adjusted in real terms using the 2009 consumer price index.

Earnings Equation Estimation. For men, we estimate the following earnings equation:

$$\ln(w) = \eta_0^m + \eta_1^m \text{exp} + \eta_2^m \text{exp}^2 + \gamma \text{CType} + \delta X + \varepsilon \tag{8}$$

where $\ln(w)$ represents the log of yearly wages, exp is the years of experience in the labor market, CType is a dummy variable for the contract being permanent, and X is a vector of control variables, including education, occupation, sector, province, and a dummy for part-time employment. The experience term is specified both linearly and quadratically to account for potential nonlinear effects of experience on wages.

For women, we modify the specification to include interaction terms between experience and contract type. This allows us to capture the differential effects of the type of contract on wages. Thus, we estimate the following regression:

$$\begin{aligned} \ln(w) = \eta_0^f + \eta_1^f \text{exp} + \eta_2^f \text{exp}^2 + \Delta\eta_0^f \text{CType} + \Delta\eta_1^f \left(\text{exp} \times \text{CType}\right) \\ + \Delta\eta_2^f \left(\text{exp}^2 \times \text{CType}\right) + \delta X + \epsilon^f \end{aligned} \tag{9}$$

where the interaction terms of experience and experience squared with the contract-type dummy are included to capture the distinct impact of contract type on women's earnings. Table F.1 reports the estimated equation coefficients for men and women.

The Shock in Log Wages. We then take the residuals obtained from these estimations and regress them on their time lags at the individual level to obtain the persistence parameters of the AR(1) process for the stochastic shocks and the variance of the innovations. Table F.2 shows the estimation results for both men and women. For the correlation coefficient between spousal shocks, we follow Hyslop (2001) and set a value of 0.25. Note that in the MCVL, it is not possible to identify spouses.

Discretization and Implementation. For the numerical solution, we approximate the auto-regressive vector of stochastic shocks for the woman and her partner with a discrete-valued Markov chain using the method proposed by Tauchen (1986) and Tauchen and Hussey (1991). We use a 5 by 5 grid for the values of the shocks, where

Table F.1. Coefficients from Wage Estimation Regressions (8 and 9) for Men and Women

	Male	Female
Experience	0.0284***	0.0396***
	(111.40)	(81.45)
Experience^2	-0.000383***	-0.000740***
	(-48.46)	(-36.32)
Permanent contract	0.337***	0.307***
	(415.17)	(99.85)
Permanent contract × Experience		-0.0157***
		(-27.89)
Permanent contract × Experience^2		0.000598***
		(26.49)
Constant	9.616***	9.322***
	(2194.32)	(1700.21)
Observations	2178651	2063360
Province	Yes	Yes
Education	Yes	Yes
Occupation	Yes	Yes
Sector	Yes	Yes
Part-time contract	Yes	Yes

t statistics in parentheses

Source: MCVL 1998-2019

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table F.2. Estimated Parameters For The Wage Shock Process for Men and Women

	Men	Women
Persistence	0.634	0.614
Variance	.125	.137

Source: MCVL 1998-2019

each point is calculated so that the income of the n-th point is the average income of the n-th quintile for men and women at age 26.

We adjust the intercept coefficients so that we match average earnings for men and women at 5 and 3 years of experience—the average experience for men and women at 27 and 25 years, which are our starting ages. These are the coefficients reported in 2.

F.1.2 Estimating Temporary to Permanent Job Transitions

To estimate the transition probability from temporary to permanent contracts, we use the MCVL dataset, building on the yearly panel data constructed to estimate the income process, but with additional restrictions. First, we restrict the sample to women older than 25. Next, we identify individuals whose first observed contract was temporary to ensure that we capture their initial employment conditions. A key part of our process involves tracking the timing and occurrence of permanent contracts and determining how many years individuals spend under temporary contracts both before and after obtaining a permanent position. We achieve this by generating cumulative sums of contract types over time and removing all temporary spells that occur after obtaining a permanent contract. Thus, the sample includes all the spells of individuals initially on temporary contracts who eventually transition to a permanent contract.

We then estimate the transition probability using a probit regression to model the likelihood of transitioning from a temporary to a permanent contract based on experience and age:

$$\Pr(\text{transition\_prob} = 1) = \Phi(\beta_0 + \beta_1 \text{exp} + \beta_2 \text{age} + \epsilon) \tag{10}$$

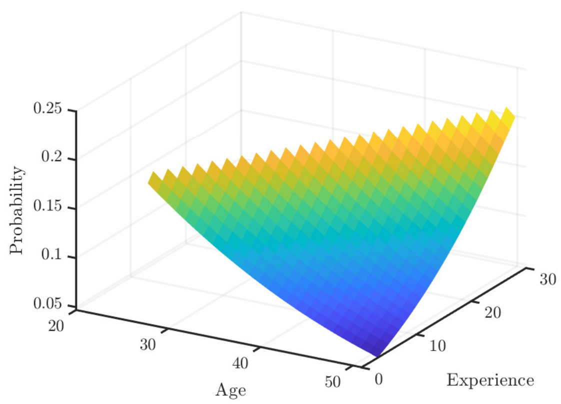
Table F.3 shows the results of this estimation. To introduce them practically in the code, we compute marginal effects for various experience levels, conditional on age, to quantify the likelihood of transitioning changes with career progression. By looping over age groups from 25 to 51, we store the marginal effects for different experience levels into a matrix, which we use in the model. Figure F.1 shows the graphical representation of the predicted probability of having a permanent contract, conditional on age and experience.

Table F.3. Coefficients From Regression 10

	Pr(transition_prob = 1)
Experience	0.00745***
	(53.37)
Age	-0.00642***
	(-70.22)
Constant	0.321***
	(119.18)
Observations	533912

t statistics in parentheses
Source: MCVL 1998-2019
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure F.1. Estimated Transition Probabilities by Experience and Age



Source: Author's work with data from the MCVL.

F.1.3 Initial Distribution of Shocks.

Initial Distribution of Shocks. To capture to some extent the degree of assortative mating present in the data, we calculate a 5x5 marital sorting matrix by income quintile. Since we do not observe partners or spouses in the MCVL, we rely on the *Encuesta de Condiciones de Vida* (Living Conditions Survey, hereafter ECV) for this. This database is provided by the Spanish Statistics Institute. The ECV is a rotating panel survey in which households are interviewed for four consecutive years. The sample includes 16,000 households, with 4,000 households leaving the panel and 4,000 new households entering it each year.

To estimate the initial distribution of shocks, we proceed as follows: first, we restrict the sample to married women between 25 and 30 years old with nonmissing earnings, both their own and their spouse's. Then, we compute the earnings quintile cutoffs for women and spouses separately, and assign each couple to a combination of wife and husband earnings quintile. Finally, we calculate the share of marriages in each category. Table F.4 shows the results. We use the probabilities implied by this matrix to draw the initial joint income shocks for households at 25 years of age.

Table F.4. Marital Sorting

Wife	Husband				
	Q1	Q2	Q3	Q4	Q5
Q1	0.083	0.051	0.037	0.020	0.016
Q2	0.053	0.045	0.045	0.044	0.017
Q3	0.024	0.045	0.047	0.034	0.038
Q4	0.027	0.035	0.056	0.043	0.039
Q5	0.013	0.023	0.028	0.046	0.090

Source: Author's work with data from the ECV.

Initial share of women with temporary contracts. To estimate the initial share of women with temporary contracts, we use the MCVL dataset and restrict the sample to women with 3 years of experience, corresponding to the average experience of women at age 25. Based on this restricted sample, we calculate the proportion of women working under temporary contracts, which is 37.44%.

F.1.4 Taxes

We model the total tax liabilities of a household as follows:

$$T\left(y^m, y^f, h\right) = \tau \cdot I\left(y^m, y^f, h\right), \tag{11}$$

where τ represents the average tax rate, and $I\left(y^m, y^f, h\right)$ denotes the gross income of the household. Gross income is determined by the income of the male household member (y^m), the income of the female household member (y^f) and other characteristics of the household (h).

Following García-Miralles et al. (2019), we specify the average tax rate, τ, as follows:

$$\tau = \begin{cases} 0 & \text{if } I\left(y^m, y^f, h\right) < \tilde{I}, \\ \max\left\{1 - \tau_0 \left(\frac{I\left(y^m, y^f, h\right)}{\bar{I}}\right)^{-\tau_1}, 0\right\} & \text{if } I\left(y^m, y^f, h\right) \geq \tilde{I}. \end{cases} \tag{12}$$

Under this specification, households with gross income below the threshold $\tilde{I}$ are exempt from taxes ($\tau = 0$). For households with gross income above $\tilde{I}$, the average tax rate increases with the ratio of household income to the average income $\bar{I}$. This progressive structure ensures that higher-income households are subject to higher tax rates. Moreover, the parameter τ_0 determines the overall scale of taxation, while τ_1 governs the progressivity of the tax system. This formulation captures both a tax-free region for low-income households and a progressively increasing tax rate for higher-income households.

The parameters τ_0 and τ_1 are borrowed from the estimates provided in García-Miralles et al. (2019). Specifically, $\tau_0 = 0.8823$ and $\tau_1 = 0.1224$ reflect the overall scale of taxation and the degree of progressivity, respectively. The remaining parameters are calibrated to fit the specific year of interest: the income threshold is set to $\tilde{I} = 1404$, and the average household income is set to $\bar{I} = 3900$. The yearly average gross household income for women aged 25 to 50 is calculated using the ECV data for the year 2007 and amounts to 36,834.14 euros.

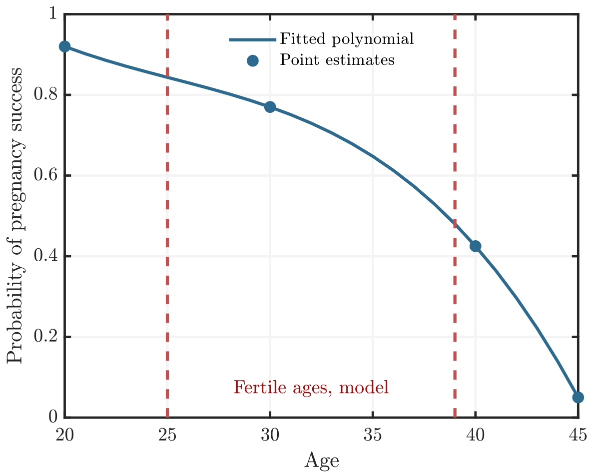
F.1.5 Pregnancy Success Probabilities

We follow Sommer (2016) in estimating the probability of pregnancy success by age α_j. We use the following point estimates of natural infertility from Trussell and Wilson (1985):

$$\text{Infertility probability} = \begin{cases} 0.080 & \text{at age } 20 \\ 0.230 & \text{at age } 30 \\ 0.575 & \text{at age } 40 \\ 0.950 & \text{at age } 45, \end{cases} \tag{13}$$

Then we fit a polynomial through (the inverse of) them, as shown in Figure F.2. Notice that we only have women having children in the model between ages 25 and 39.

Figure F.2. Pregnancy Success Probability Conditional on Age



Source: Author's work, point estimates by Trussell and Wilson (1985).

F.2 Second Step: Identification

F.2.1 Labor Force Participation

To estimate the part-time and full-time participation rates for three groups of women (six targets)——childless women, mothers whose youngest child is 0–3 years old (newborn or baby), and mothers whose youngest child is 3–12 years old (school-age)—we use the 2004–2007 ECV.

For each woman, we have information on participation by month. We compute an average annual participation rate, counting each month worked full-time as 1, each month worked part-time as $\frac{1}{2}$, and dividing by 12. Following Bick (2016), we create a yearly participation status variable that falls into one of our three categories: If the average participation rate is above 0.75, we count the woman as having worked full time that year, if it is between 0.75 and 0.25 we count her as having worked part time, and if it falls below 0.25 we count them as out of the labor force.

F.2.2 Identification

Although all parameters affect all model moments once we solve the model, some are more important than others for specific targets. The first two parameters, γ_c and γ_l, govern how quickly the marginal utility of consumption and leisure decreases, respectively. These parameters are critical in determining women's willingness to substitute between consumption, leisure, and children. They play a significant role in decisions about the intensive margin of participation in the labor force (LFP) (part-time versus full-time) and influence the responses of fertility and LFP to baby checks.

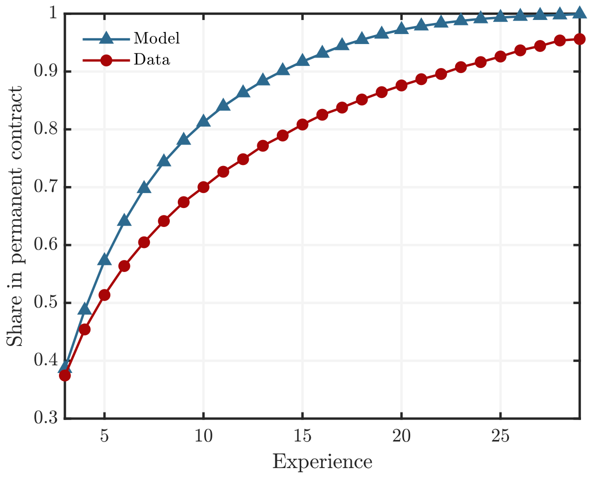
The parameter on the age-varying weight of the fertility gap determines how early the craving for children begins in the women's life cycle, making it essential for the average age at first birth. The weight on leisure, δ_l, is crucial for the extensive margin of LFP decisions. For the distribution of women by the number of children, δ_{N1}, δ_{N2}, and ζ are key. The first parameter, δ_{N1}, increases the penalty for not achieving the desired number of children. As it rises, more women will likely achieve or approach their desired number of children. The second parameter, δ_{N2}, reduces the marginal penalty for not having a third child, influencing the fraction of women who ultimately have three children. The final parameter, ζ, represents an additional utility of motherhood and is vital for the extensive margin of fertility—specifically the decision to remain childless or have children.

The next four parameters, ξ_1, ξ_2, κ_1, and κ_2, are relevant to the decisions about LFP of mothers with children of different ages. The first two parameters, ξ_1 and ξ_2, affect the extensive margin, while κ_1 and κ_2 influence the intensive margin by increasing the cost of full-time work. Finally, the part-time earnings penalty, ϕ, directly impacts the likelihood of part-time work.

F.3 Model Fit: Untargeted Moments

In this section, we show the model fit for a set of additional moments. Figure F.3 compares the share of women with permanent contracts in the model and the data by experience levels. Figure F.4 illustrates the gender wage gap by age in both the model and the data. Figure F.5 shows the gap in labor force participation between mothers and childless women by age.

Figure F.3. Share of Women in Permanent Contracts by Experience, Model and Data



Source: Author's work with data from the MCVL.

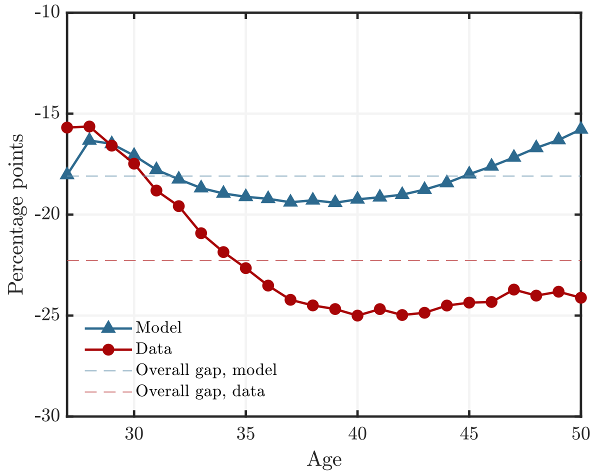
G Baby Check Effect by Age in the Short Run

We build on the analysis in González (2013) by independently estimating the following two equations for each age group. The first specification is the following.

$$\log(C_m) = \alpha + \gamma_1 m + \gamma_2 (m \cdot post) + \beta \cdot post + \lambda X_m + \epsilon_m \tag{14}$$

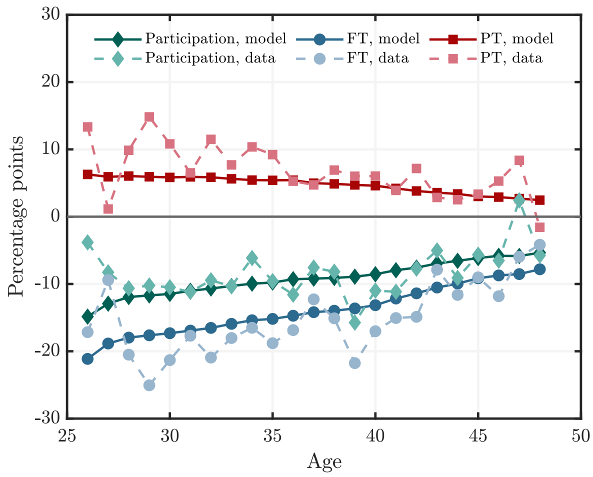
where C_m represents the natural logarithm of the estimated number of conceptions in month m, *post* is a binary variable equal to 1 for all months starting in July 2007, and X_m is the number of days in month m. The coefficient β captures the discrete jump in

Figure F.4. Gender Wage Gap: Model and Data



Source: Author's work with data from the MCVL.

Figure F.5. Labor Force Participation Gap Between Mothers and Childless Women by Age: Model and Data



Source: Author's work with data from the ECV.

the monthly number of conceptions in July 2007. A positive value of β would indicate that the policy successfully encouraged fertility.

The second specification extends the model by including calendar month-of-conception dummies:

$$\log(C_m) = \alpha + \gamma_1 m + \gamma_2(m \cdot post) + \beta \cdot post + \lambda X_m + \sum_{c=2}^{12} \mu_c month_m + \epsilon_m. \tag{15}$$

In this equation, the coefficient of interest, β, continues to measure the discrete jump in the number of conceptions between June and July 2007, controlling for the average difference in conceptions between these months in other years.

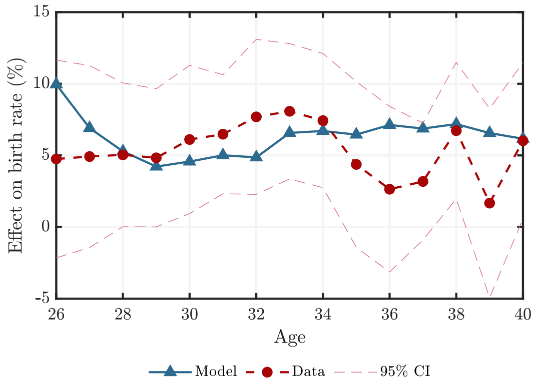
Figure G.1 presents the results of the estimation of equations 14 and 15 across the eight different specifications, separated by age group.

H Quantitative Experiment: Labor Market Duality

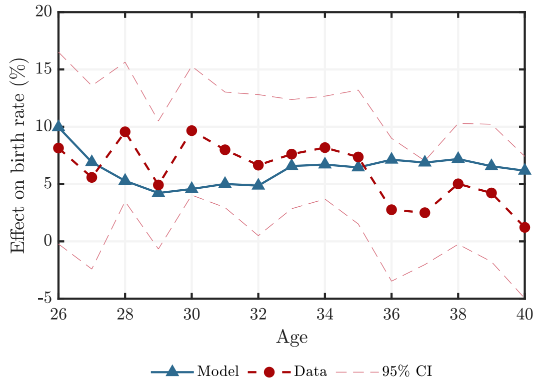
For this counterfactual analysis, we run the same regression as in 8 for women using the same sample as in the original estimation of income for women. We show the estimation results in Table H.1, where we include men for comparison.

Figure G.1. Effect of the baby check by age in the short run, model and data

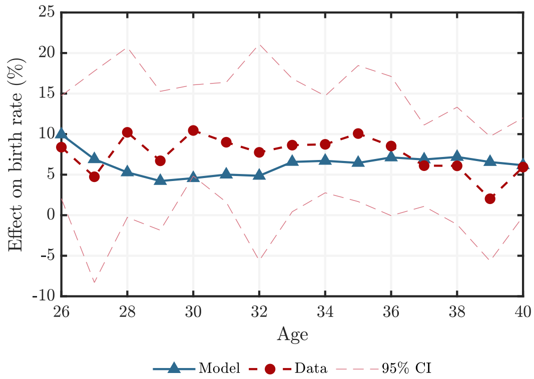
(a) RDD 10 years



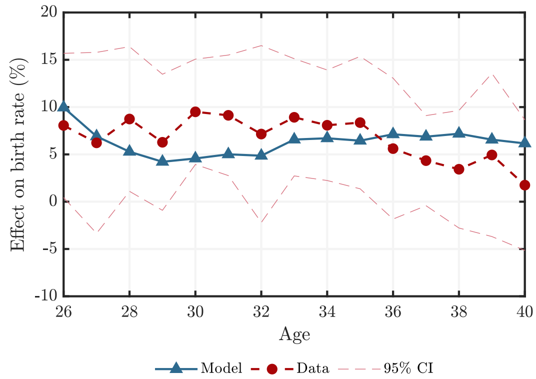
(b) RDD 5 years



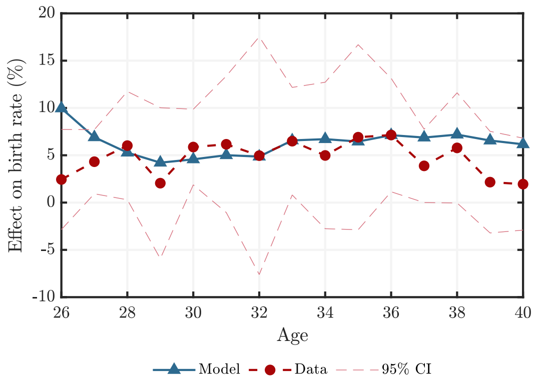
(c) RDD 12 months



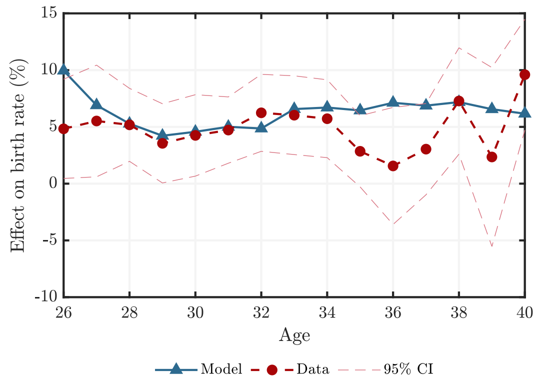
(d) RDD 9 months



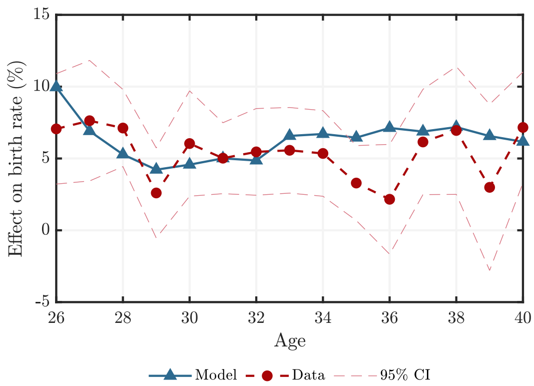
(e) RDD 3 months



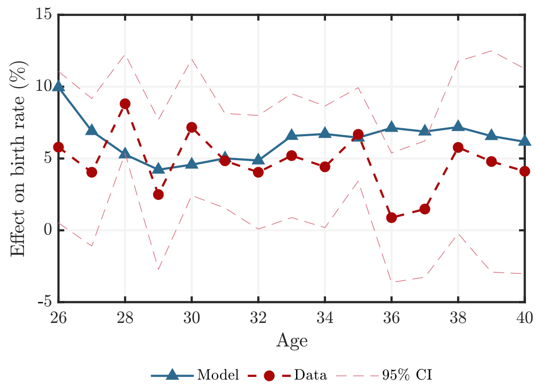
(f) DiD 10 years



(g) DiD 7 years



(h) DiD 5 years



Source: Author's work using vital statistics by the Spanish Statistical Office.

Table H.1. Coefficients from Wage Estimation Regressions (8) for Men and Women

	Male	Female
Experience	0.0284^{***}	0.0280^{***}
	(111.40)	(118.76)
Experience^2	-0.000383^{***}	-0.000277^{***}
	(-48.46)	(-33.15)
Constant	9.616^{***}	9.376^{***}
	(2194.32)	(1827.44)
Observations	2178651	2063360
Province	Yes	Yes
Education	Yes	Yes
Occupation	Yes	Yes
Sector	Yes	Yes
Part-time contract	Yes	Yes
Permanent contract	Yes	Yes

t statistics in parentheses

Source: MCVL 1998-2019

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$