

The fiscal return to childcare policies

David Koll* Dominik Sachs† Fabian Stürmer-Heiber‡ Hélène Turon§

February 15, 2019

Work in Progress

Abstract

We study the long-term fiscal implications of childcare subsidies through their impact on maternal labour supply. Taking human capital accumulation into account, we explicitly capture life-cycle career aspects in a dynamic structural household model of female labour supply and childcare decisions: higher labour supply of mothers today can result in higher future earnings. In our dynamic structural model, we allow households to be heterogeneous in their taste for home produced childcare, their taste for leisure and in their access to informal childcare (e.g. by grandparents). Using German survey data, we provide a structural estimate of the degree to which childcare subsidies are dynamically self-financing through higher labour income tax revenue. Further, we explore how the marginal fiscal returns of childcare subsidies depend on the group of families targeted (e.g. low income, single parents, number of children).

JEL Code: H21, H23, H31, J13, J22, J24, J31

Keywords: Female Labour Supply, Childcare, Learning-by-Doing, Dynamic Scoring, Fiscal Externalities, Germany

*European University Institute, Florence. David.Koll@eui.eu

†LMU Munich. Dominik.Sachs@econ.lmu.de

‡LMU Munich. Fabian.Stuermerheber@econ.lmu.de

§University of Bristol. Helene.Turon-Lacarrieu@bristol.ac.uk

1 Introduction

Kleven, Landais, Posch, et al. 2019 document substantial child penalties in earnings for different countries. The penalty is particularly large for Germany, where female earnings after child birth drop by 80% in the short run and still remain 60% below the pre-birth level in the long run. Besides, there is limited availability of affordable childcare in Germany, and this has been put forward as a reason for the low earnings of mothers. Between 2005 and 2013 several laws have been adopted to expand subsidised childcare availability. One explicit goal of these policies was to promote female labour supply.¹

It has been argued that childcare subsidies may be partly self-financing through higher tax revenue: if childcare subsidies promote female labour supply, this also increases tax revenue.² In this paper we operationalize this idea and propose both a theoretical framework and a structural estimate of this dynamic fiscal effect for Germany.

We show that answers to the following questions are crucial in assessing the fiscal returns to childcare subsidies: (i) which families are marginal in their decision to send their child(ren) to childcare? (ii) Of these families, which are also marginal in their labour supply decision? (iii) How does labour participation affect the future earnings trajectories of these 'marginal' mothers? (iv) How do infra-marginal households in terms of childcare use change the number of hours of childcare in response to a subsidy?

We provide answers to these questions within a structural dynamic framework where households choose maternal labour supply and hours of market childcare. Besides the usual time and budget constraints, households are faced with a 'childcare' constraint, i.e. the age composition of the household's children implies a need for childcare that is quantified in the model. Households are heterogeneous along observable characteristics, i.e. current ages and number of children, father's and mother's (potential) wages, as well as unobserved characteristics. The nature of these is fourfold: households differ in their 'fertility type', defined as the age at first child birth and desired completed family size, in their preferences in terms of home produced childcare and maternal leisure, and in their access to free informal caretakers for their children (e.g. grandparents). We show the importance of both types of heterogeneity for the answers

¹The idea that such policies can indeed achieve this goal is supported by a sizeable empirical literature, see, e.g. Bauernschuster and Schlotter 2015 for the German context, Baker, Gruber, and Milligan 2008 for the Canadian context or Cascio, Haider, and Nielsen 2015 for a summary of different studies for different countries.

²Rainer et al. 2013 provide some back of the envelope calculations for Germany based on static tax revenue effects. We relate to them in more detail below and show what such back of the envelope calculations are missing.

to the above questions and argue that our structural model is well suited to take these into account.

The estimation of the model is carried out with survey data from the German Socio-Economic Panel (GSOEP) and consists of three major steps. First, we estimate wage dynamics and the penalties on wage growth that result from part-time work or non-participation.³ Second, we assign a predicted fertility type to each household since actual types are only observed for households with completed fertility. Finally, we estimate our structural model to obtain the joint distribution of preferences for home produced childcare, leisure and access to informal childcare. Our empirical strategy is to maximise the likelihood of observed household trajectories in terms of female labour supply and used hours of market childcare services. Since tax and childcare price schedules are key to our fiscal calculation, we include a detailed specification of the variations of both of these with family characteristics.

Estimated parameters allow us to simulate our model's response to small perturbations of childcare fees and predict the long-term marginal fiscal returns of childcare subsidies. In particular, we will obtain these predictions for different subgroups, which will inform policy makers about optimal targeting of subsidies. Indeed, since the degree to which subsidies are self-financing is likely to differ non-trivially with observable characteristics as household income, number of children or single versus two-parent households, targeting will bear a crucial impact on the expected costs and benefits of the policy in the long-term.⁴

Related Literature. This paper relates to three strands of literature: reduced form studies of childcare, structural work in labour and some of the public finance literature.

Reduced form studies that show positive effects of childcare subsidies on labour supply include Bauernschuster and Schlotter 2015 and Gathmann and Sass 2018 for the German context. For other countries, see Baker, Gruber, and Milligan 2008, Bettendorf, Jongen, and Muller 2015, Givord and Marbot 2015, Geyer, Haan, and Wrohlich 2015 and Nollenberger and Rodríguez-Planas 2015.⁵

More structural work includes Bick 2016 and Wang 2019. The former paper studies the effect of childcare policies in Germany and the latter jointly studies the impact of childcare

³The current version only incorporates a calibration, but a careful estimation and presentation of the fit of the model will be added in future versions.

⁴In the current version, we do not incorporate single parents. Doing so will be a straightforward extension that we will add in future versions.

⁵This list is by not complete and will be updated.

and parental leave policies on female employment and fertility. Also related is the paper by Guner, Kaygusuz, and Ventura 2018 who study welfare effect of different child related transfers in the U.S. context. Our approach is different in the sense that we are interested in assessing the true fiscal costs of these childcare subsidies. Our model, in particular the way we model the childcare need and the taste for home produced childcare, borrows heavily from Turon 2018. Lastly, the paper is quite related to Adda, Dustmann, and Stevens 2017 who structurally estimate the career costs of children and Blundell et al. 2016 who structurally estimate the returns to experience for woman in the UK (and thereby also the wage penalties from not working full time).

Many papers in the public finance literature have emphasized that the implied effects on labour supply provide a rationale for subsidizing childcare Bastani, Blomquist, and Micheletto 2017; Domeij and Klein 2013; Ho and Pavoni 2016. The goal of this paper is to make this argument more operational from an applied point of view and ask how large the implied fiscal externality of childcare subsidies really is. The paper is also related to Colas, Findeisen, and Sachs 2018, who study the same question for college education subsidies but where the trade-offs are different.

This paper is organized as follows. In Section 2 we present some policy background and stylized facts which motivate our analysis. Section 3 contains the model setup and the quantification of the model is described in Section 4. In Section 5 we illustrate the fit of the model with data and quasi-experimental studies before we present our policy analysis in Section 6. Section 7 concludes.

2 Background, data, and stylised facts

2.1 Institutional background in Germany

Childcare facilities. Three different types of childcare institutions can be distinguished in Germany: First, children below the age of 3 are taken care of in nurseries (day care centres). Approximately around the time when children turn 3 years old, they enter kindergarten and stay there until they start school. Lastly, during school age children may attend after school care centres in the afternoon.⁶ The distinction between these institutions matters because childcare prices differ by attended institution and, therefore, by child age. This is foremost due

⁶Note that some children might also attend full-day schooling.

to different institutional cost structures and price regulations.⁷ In terms of the opening hours of childcare facilities, there are notable differences between West and East Germany. In West Germany approximately 50 percent of nurseries and kindergartens are open for ten hours or more, while in East Germany the number is 75 percent.⁸

Recent reforms. While all three types of childcare institutions have been continuously present throughout Germany since the early 1990s, their use and prevalence has changed substantially in the past thirty years. During this time, family policy has increasingly focused on childcare as a tool to allow females to both participate in the labour market and have children. A major first step in this policy effort has been the introduction of a legal claim to a kindergarten spot in 1996, introducing legal entitlement to childcare for every child aged 3 until school entry. While this reform led to a substantial increase in the number of kindergarten slots, childcare supply for children below 3 (i.e. nursery slots) remained scarce until 2005. In that year, the German government committed to building up 230,000 additional nursery slots by October 2010, increasing childcare coverage for children aged 0 to 2 from just 5 to 17 slots per 100 children.

Focusing policy efforts further on nurseries, an extension of the legal entitlement to childcare was passed in 2008. From October 2010 on, it introduced a legal claim to a nursery slot for all children below 3 if both parents are working. In August 2013, this legal claim was further extended to cover all children after their first birthday, regardless of the employment status of the parents.⁹

2.2 Data

For our analysis, we focus on German females aged between 20 and 65 that are not currently in education. We track this group over the time span 2000 to 2016 in a representative longitudinal survey data from the core waves of the German Socio-Economic Panel (GSOEP). The GSOEP is an unbalanced household panel that has been running since 1984 in West Germany and includes East Germany since 1991. Its scope is comparable to the US Panel Study of Income

⁷See Appendix C.1.1 for a detailed description of childcare price determinants.

⁸See Authoring Group Educational Reporting 2018, p. 70 and Table C2-13web.

⁹Family policy related law changes: 1996 - Reform of the §SGB VIII, 2005 - Tagesbetreuungsbaugesetz, 2008 - Kinderfoerderungsgesetz

Dynamics, as it provides annual socioeconomic and demographic information on the household and individual level.¹⁰

2.3 Stylised facts

Effect of child birth on maternal labour supply. Persistent effects of parenthood on maternal labour market outcomes – also called child penalties – are a well-established fact in the recent literature for a growing number of countries (e.g. Kleven, Landais, and Søgaaard 2018, Angelov, Johansson, and Lindahl 2016, Kleven, Landais, Posch, et al. 2019).

This also holds for Germany where childbirth has a substantial and sustained effect on maternal employment status.¹¹ Figure 1 illustrates this effect in the GSOEP data. It compares maternal pre-birth employment rates to their post-birth developments with age of the youngest child. The average pre-birth employment rate of future mothers is almost 90 percent and most of these females work full-time, namely more than 75 percent. In the first year after childbirth the total employment rate drops by more than 70 percentage points to below 16 percent and only five percent of mothers with newborns work full-time. It is important to note that during this time and until the third birthday of the child mothers may take maternity leave. During such a maternity related temporary 'non-employment'-spell mothers partially receive financial compensation and have the right to return to their pre-birth job afterwards.

As children age, the share of mothers working full-time rises gradually. While it reaches 18 percent by the end of statutory maternity leave at age 3, from then on it only increases slowly to just above 20 percent when the youngest child is 10 years old. Afterwards the full-time share gradually rises up to 40 percent by the time the youngest child reaches adulthood. However, this level is still 30 percentage points lower than the pre-birth full-time share, suggesting that a substantial fraction of mothers never return to full-time work.

Part-time contracts illustrate another side of the employment pattern of mothers: Before the first birth, only 14 percent of future mothers work part-time. This share then increases sharply following the first year after childbirth on to reach 54 percent by age 5. At this point, total employment reaches almost 75 percent. This is in line with many mothers opting for part-time when their children are young, although they worked full-time before parenthood.

¹⁰A detailed description of the GSOEP can be found in Goebel et al. 2018.

¹¹Kleven, Landais, Posch, et al. 2019 use the event study approach by Kleven, Landais, and Søgaaard 2018 and data from 6 countries (US, UK, Denmark, Sweden, Austria and Germany) to show that long-run earnings penalties range from 21% in Denmark to 61% in Germany.

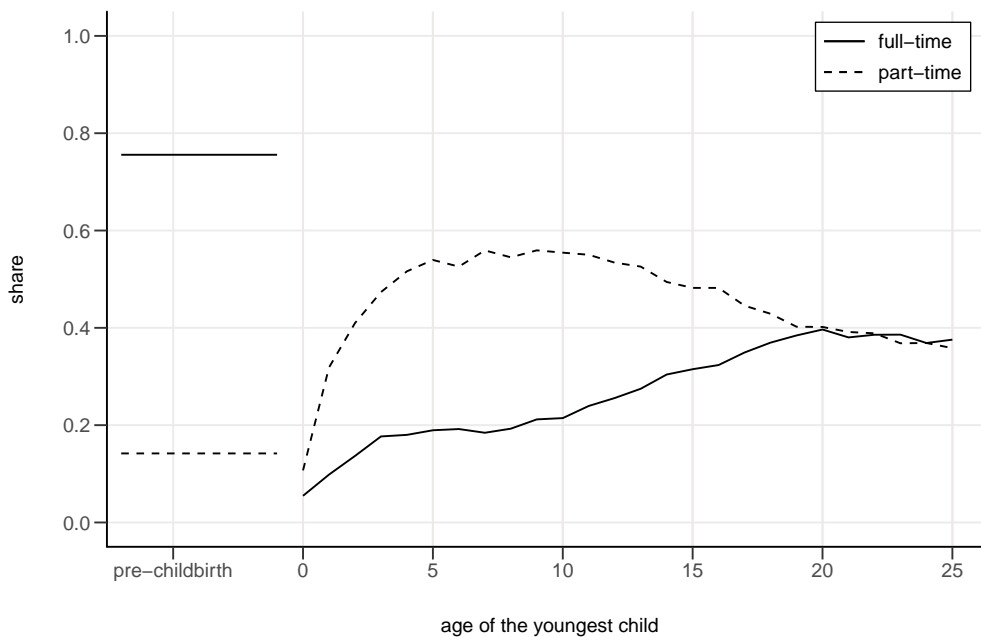


Figure 1: Employment status by age of the youngest child

Notes: Maternity leave is treated as non-employment. Sample definition: females aged 20 to 65 that are not in education, conditional on having at least one child. Source: 2000 to 2016 core GSOEP samples.

After the youngest child has turned 10, the part-time share gradually declines to 40 percent. Together with the simultaneously increasing full-time share, this trend suggests an intensive margin decision, i.e. mothers returning from part-time to full-time. Summing up, participation remains almost flat around 75 percent from age 5 on.

Taken together, Figure 1 shows that childbirth substantially reduces the hours that women work, but that part-time arrangements do play a large role for mothers' labour market attachment. These part-time jobs yield lower wage growth rates, which we will document in Section 4.3. This part-time penalty additionally decreases mothers earnings over the life cycle, amplifying the effect of working fewer hours after childbirth.

Childcare enrollment. One of the explicit goals of the recent family policy reforms in Germany has been to increase maternal labour market participation, both, in terms of part-time and full-time work. As laid out in Section 2.1, the expansion of subsidised childcare, especially for children below the age of three, has been a major policy instrument to accomplish this goal. Figure 2 plots childcare enrollment shares for different child age brackets across recent years.¹²

¹²These enrollment patterns in the GSOEP data line up well with administrative data from the German Statistical Office, plotted in the corresponding Figure 12 in the Appendix.

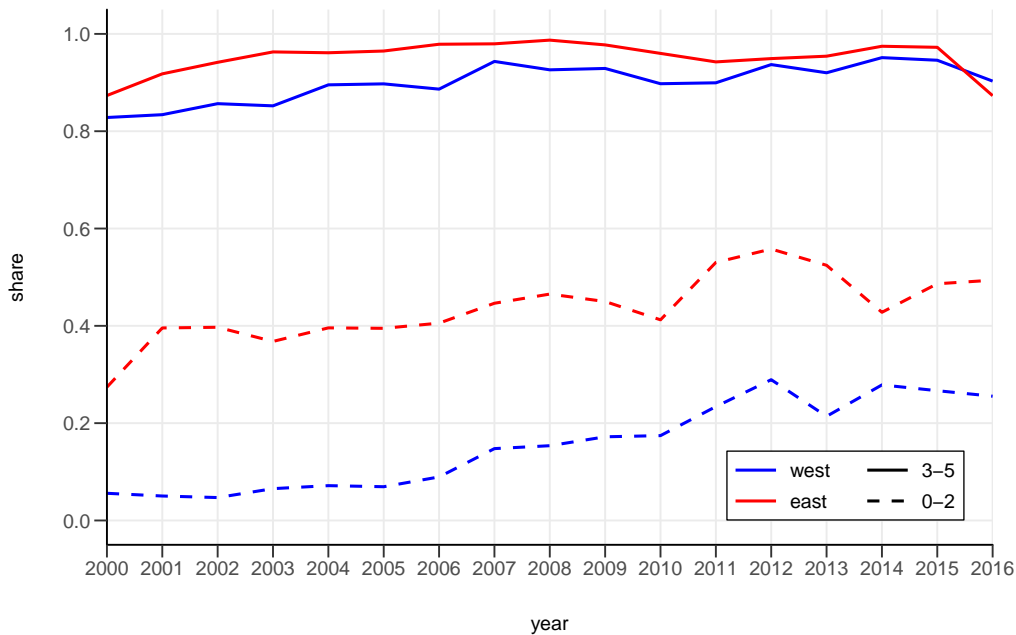


Figure 2: Childcare enrollment by region

Notes: Enrollment is binary in the sense that it is not conditional on a minimum number of hours. Sample definition: mothers aged 20 to 65 that are not in education. Source: 2000 to 2016 core GSOEP samples.

In 2000, more than 80 percent of 3 to 5 year olds were already attending childcare in West Germany, while the share was only six percent for children below 3. Over the years, during which subsidised childcare was substantially expanded, this enrollment share has increased to 26 percent in 2016. During this entire period, the childcare attendance shares for children below 3 have been considerably higher in the East, where the use of childcare is much more pronounced for historical reasons.¹³ Overall, childcare attendance of children below 3 has increased by more than 20 percentage points over last 16 years in both regions. Enrollment of 3 to 5 year olds has slightly increased over that time span in the West by five percentage points to 87 percent, while staying above 90 percent in the East for all periods except in 2016, when it dropped to 87 percent.

Maternal labour supply by age of the youngest child. To relate this pattern of increased childcare attendance to employment rates, Figures 3(a) and 3(b) plot maternal full-time and part-time participation rates by the age of the youngest child. During the 2000 to 2016 time

¹³During the division of Germany, the socialist GDR government established an advanced childcare system to both allow mothers to work as well as to insure that child education was in line with socialist ideology. This led to a high use of institutional childcare and to a higher availability of childcare slots, which both persisted after reunification.

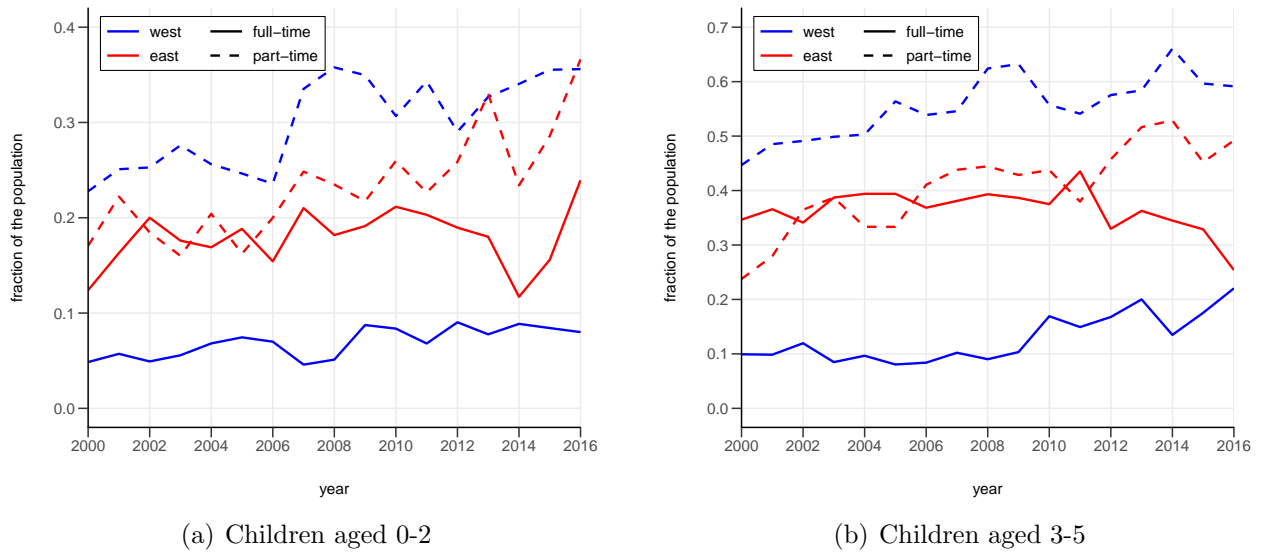


Figure 3: Employment of mothers by age of the child

Notes: Sample definition: mothers aged 20 to 65 that are not in education, conditional on having a child aged 0 to 2 for panel (a) or 3 to 5 for panel (b). The mothers in panel (a) mothers make up 7.6% (7.2%) of all West (East) German females in the sample, the mothers in panel (b) make up 6.2% (5.5%) of all West (East) German females in the sample. Source: 2000 to 2016 core GSOEP samples.

span, when childcare enrollment of children aged 0 to 2 increased substantially, mothers of children in this age bracket have been increasing their labour force participation both along the extensive and partly also the intensive margin. Part-time rates among these mothers have increased by more than ten percentage points across Germany, while full-time rates have also slightly increased. Looking at the 3 to 5 year olds in Figure 3(b), part-time employment shares and to some extent full-time shares, have been increasing between 2000 and 2016. In West Germany, both the full-time and the part-time share are 10 to 15 percentage points higher in 2016 than in 2000. While in the East, part-time employment shares have doubled from 25 to 50 percent, the full-time share has fallen recently to 25 percent (compared to 35 percent in 2000). For both age brackets, it is worth pointing out that full-time employment rates remain below 25 percent in West Germany throughout the entire time span.

Taken together, Figures 1 to 3(b) highlight that childbirth substantially impacts female employment shares over the life cycle. Nonetheless, these stylised facts also point to a potential positive impact that childcare has on the labour supply decision making of mothers.

Childcare prices. Figure 4 sheds some light on childcare prices by plotting for each income decile the share of household income spend on full-time equivalent childcare fees. Across the

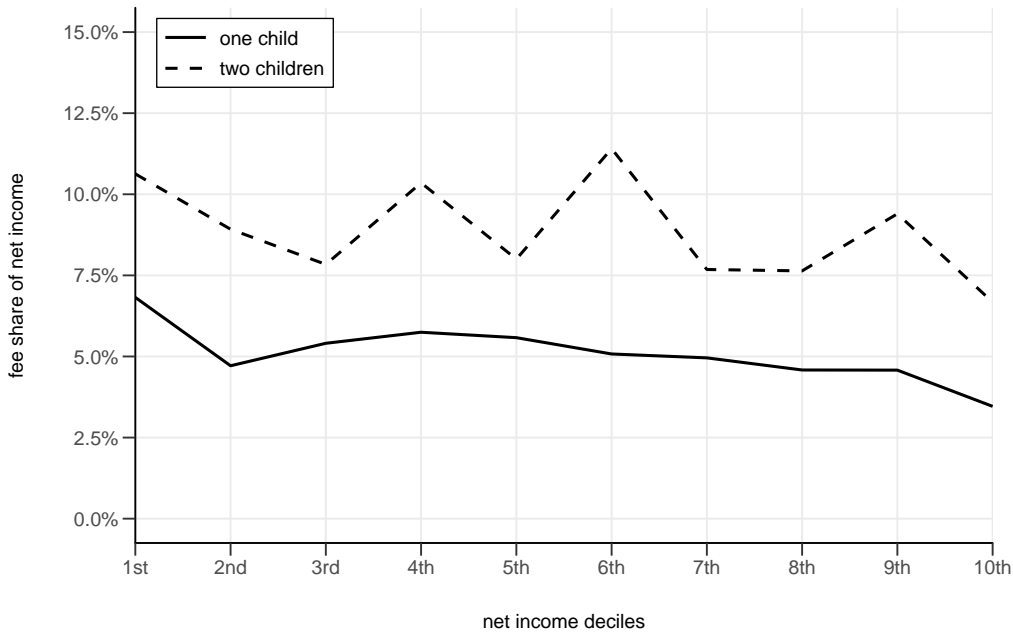


Figure 4: Income share of childcare fees by number of children

Notes: Full-time equivalent fee share of household net income. Sample definition: mothers aged 20 to 65 that are not in education, conditional on childcare attendance and observed childcare fees (incl. zero fee observations). Source: 2013 and 2015 core GSOEP samples, in 2011 prices.

income distribution, childcare fees for one child make up three to seven percent of households' monthly after-tax income. On average, households with lower incomes would need to spend a higher fraction of it for full-time childcare. Hence, the share of net income spent on childcare fees is regressive. Nonetheless, high income households pay higher fees in absolute terms, namely up to an average of 172 Euros per month in the tenth income decile. Low income households on the other hand only pay around 50 Euros.¹⁴ The additional childcare fee for a second child makes up around three percentage points for most income deciles, with some outliers in the 4th and 6th decile (possibly due to a low number of observations). Figure 4 also illustrates that childcare fees increase less than proportionally with family size.

Looking at how full-time childcare fees have evolved over time conditional on the age of the child, Figure 9 in the Appendix shows that average monthly fees for children below the age of 3 have always been higher than for children aged 3 to 5. From a difference of just 20 Euros in 2005, the fee gap between these two age brackets has widened to more than 80 Euros in 2015. It has therefore become relatively more costly for mothers to make use of childcare for children below the age of 3 compared to childcare for children aged 3 to 5.

¹⁴See Figure 10 in the Appendix for actual childcare fees per month, non-adjusted for attended hours.

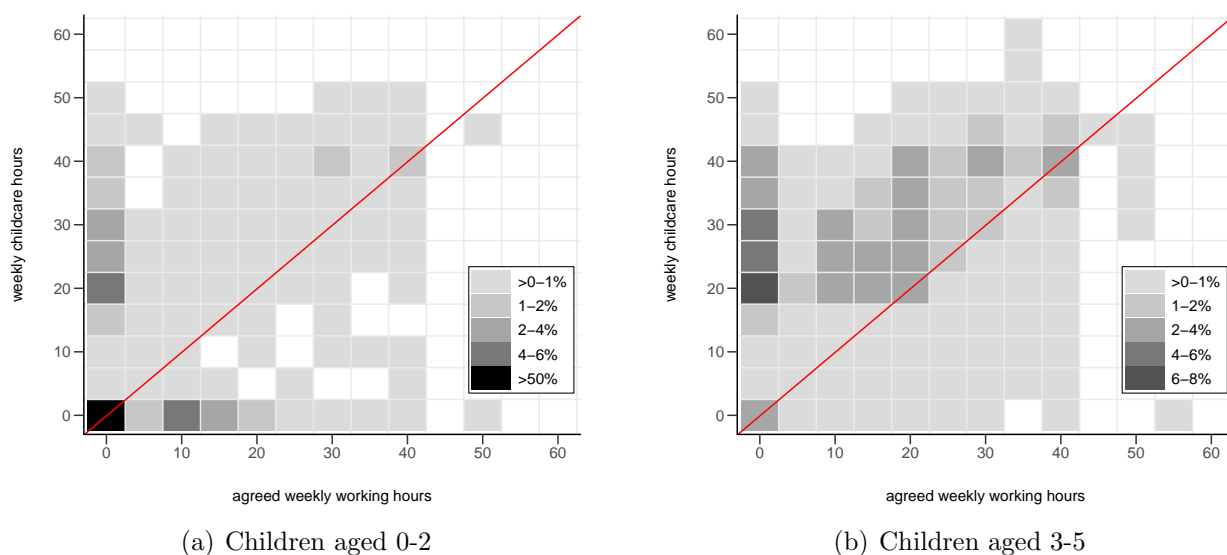


Figure 5: Mothers working hours vs. childcare hours by age of the child

Notes: Sample definition: married mothers aged 20 to 65 that are not in education, conditional on having a child aged 0 to 2 for panel (a) and 3 to 5 for panel (b), the child not being taken care of by a paid external care taker and the partner working full-time. Source: 2009 to 2016 core GSOEP samples.

Childcare hours versus working hours. In order to further look at the joint decision making of mothers concerning the use of market childcare and labour supply, Figures 5(a) and 5(b) plot agreed weekly working hours of the mother against the number of hours the children spend in childcare. To focus on the trade-off faced by the female, the sample is restricted to married mothers whose partner works full-time.

As Figure 5(a) shows, more than half of mothers of below 3 year olds neither work nor send their child to childcare. Of those who do work, not everybody sends their child to childcare at the same time (mass at zero childcare hours and positive working hours). Furthermore, we observe some children in childcare whose mothers are not working at all.

For the 3 to 5 year olds in Figure 5(b), the picture looks entirely different. Most children attend childcare and many mothers work to some degree. Similar to the below 3 year olds, a number of mothers does not work while their child is in childcare (mass at zero working hours). Focusing on the mass close to the red 45-degree line, we observe that with increasing weekly working hours, children also spend more time in childcare. In fact, for 27 percent of the observations in Figure 5(b), childcare hours correspond to weekly working hours within a 10 hour interval (excluding mothers that do not work at all).¹⁵

¹⁵A ten hour interval corresponds to a three bin interval in agreed working hours. See Table 5 in the Appendix for summaries of the numbers underlying Figures 5(a) and 5(b).

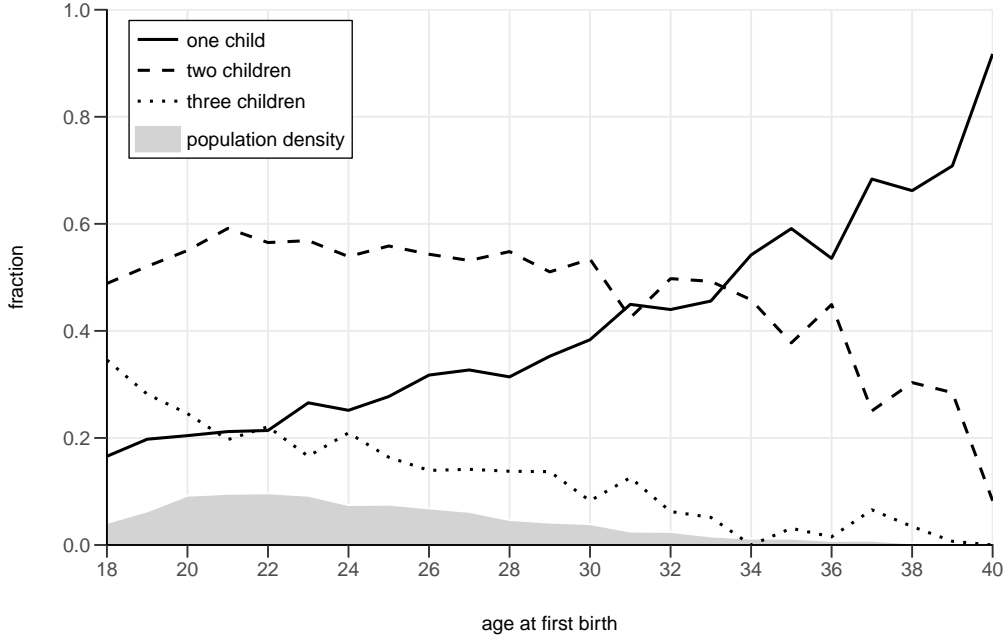


Figure 6: Completed fertility across age at first birth

Notes: Populations density illustrates the relative frequencies of age at first birth in the sample. Sample definition: mothers aged 41 to 65 that are not in education, conditional having given birth only between ages 18 and 40 to at most three children. Source: 2000 to 2016 core GSOEP samples.

Completed fertility. To complement the picture of the interplay of fertility, childcare, female labour supply and wages, Figure 6 plots completed fertility size across the age of the mother at first birth. It helps to understand when in their life-cycle mothers start having children and how many. Most mothers give birth for the first time in their early twenties and end up completing fertility with two children. Naturally, those who give birth later have a higher propensity to only have one child. Furthermore, having three children is more common if the first birth occurs early in life.

3 Model

We present here a dynamic model of households faced with labour supply and childcare choices. A household is composed of two adults with up to three children. Households' decision making is unitary and forward looking. The unit time period is 3 years and the period discount rate is denoted β . For simplicity, both spouses have the same age t . We assume that fertility is exogenous but allow households to be heterogeneous with respect to their fertility 'type', defined as the age at first child birth and completed family size. Households know their fertility type at all times. Households are also heterogeneous in their preference for home produced childcare,

their taste for the leisure of the female spouse and their access to free informal childcare (e.g. grandparents) during their children’s infancy (under 3 years). Our aim is to estimate these three dimensions of heterogeneity in order to gain a better understanding and an empirical estimate of the aggregate labour supply elasticity of mothers to the price of childcare.

Labour market choices are discrete: the female spouse can work full-time, part-time (half of the full-time hours) or not participate, while the male spouse always works full-time. The dynamic components of our framework come from the impact of labour supply choices on the growth rate of potential wages, which is positively affected by hours worked in the current period. Hence, career breaks imply dynamic wage penalties. There is however no instantaneous penalty of part-time work on the hourly wage. This is the only endogenous dynamic component of our model as childcare choices will not have any long-term impact on the household’s well-being, e.g. through positive effects on children. We rule out saving and borrowing.

We assume that marriages are formed at the age of 20, spouses retire at 65 and have a lifespan of 15 years after retirement. We rule out divorce. We formalise below the various components of our framework and the dynamic optimisation problem that households face. The choices to be made relate to the labour supply of the female spouse and the mode of childcare (if childcare is needed), to be chosen between home produced childcare, informal childcare (if available to the household), or use of childcare services.

3.1 Preferences.

Households value female leisure time L , consumption c and the time devoted by the mother to home produced childcare dcc (i.e. childcare in normal working hours). Even when both spouses work full-time, they enjoy a number \bar{L} of female leisure hours outside normal working hours. Preferences are reflected in the following instantaneous utility function:

$$u(c, L, dcc) = \frac{c^{1-\gamma}}{1-\gamma} + \alpha(L + \bar{L})^{1/2} + g \cdot dcc \quad (1)$$

Two elements of this utility function are heterogeneous across households: g and α which are the tastes for home produced childcare and female leisure. The coefficient of relative risk aversion γ is homogeneous.

3.2 Children

Each household's childbearing trajectory is deterministic and characterised by its 'fertility type', denoted f . This type comprises the age at first birth a and the number of children that the household will eventually have, n . Age at first birth can be in any of the first 7 periods, i.e. between the ages of 20 and 41. In addition, households can have between 0 and 3 children.

$$f = \begin{cases} (\cdot, 0) & \text{if no children} \\ (a, n) & a \in \{1, \dots, 7\}, n \in \{1, \dots, 3\} \text{ otherwise} \end{cases} \quad (2)$$

This setup leaves us with 22 different fertility types. The two underlying assumptions are that each household only has one child per 3-year period and that the children's age difference is equal to the period length (i.e. 3 years). Given the fertility type f , the household's number and ages of children is deterministic. $K(t, f)$ denotes a 3-element vector indicating the presence of a child in the first three age brackets (0 – 2), (3 – 5) and (6 – 8) in any household of type f and age t . For example, a household of type $f = (2, 1)$ is a household that will bear its first (and only) child in the second period, i.e. at age (23 – 25), and thus will have, in period 4, one child in the third age category and none in the other two categories: $K(4, (2, 1)) = (0, 0, 1)$. By assumption, each element of K can only be 0 or 1 since only one child can be born in each period.

3.3 Childcare

In each category $i = 1, 2, 3$ relating to the age ranges (0 – 2), (3 – 5) and (6 – 8), a child needs an age-specific number of hours of childcare within normal working hours, \bar{t}_i . For example, in the first and second categories, the infant needs care all of the time, whereas in the third category, the child needs care in the non-school hours.

There are three ways in which parents can fulfil this childcare need: maternal time, which we denote dcc and refer to as 'home produced' childcare, informal childcare, e.g. by grandparents, or formal childcare services, e.g. by a nursery, denoted mcc . Informal childcare is only available for infants, i.e. children in the youngest age range, and is free. It is only available to some households and this availability is denoted oth and known to the household. We will thus refer to it as unobserved heterogeneity. Formal childcare is always available at a price, normalised to

full-time use, which depends on the age of the child i , the family structure K and the household net income y^{net} (which will be derived below):¹⁶

$$p(i, K(t, f), y_t^{net})$$

Since households' preferences put value neither on informal childcare nor on market childcare, and informal childcare is free, households will always opt for informal childcare when it is available. The amount of market childcare necessary for a child of age i is thus:

$$mcc(i) = \max \{0, \bar{t}_i - dcc - oth * \mathbb{1}(i = 1)\} \quad (3)$$

and the household expenditure for childcare:

$$Ecc(t, f, y^{net}) = \sum_{i=1}^3 K_i(t, f) \cdot p(i, K(t, f), y^{net}) \cdot mcc(i).$$

where $K_i(t, f)$ is the i th element of the 3-dimensional vector $K(t, f)$ indicating if a child with age i currently lives in the household. Note that we assume that taking care of n children at home requires one hour of home produced childcare but n hours of market childcare.

We denote w^M and w^W the wages of the two spouses. As mentioned above, the male spouse always works full time and w^M refers to his actual wage. The female spouse may or may not work, in which case w^W refers to her potential wage. We assume the simplest wage dynamics in that wage processes are deterministic, whereby wages grow at a constant rate, which is specific to the labour supply choice of the current period.

Once the partners retire, they get a fraction \mathcal{B} of their last period's full-time earnings potential as retirement benefits.

3.4 Dynamic problem

We now formalise the dynamic optimisation problem faced by households. The state variables that determine the household's behaviour are its current age t , the male wage w_t^M and the female wage w_t^W . All three variables evolve deterministically, with age increasing by 1 every period and wages growing as described above. The unobserved and constant characteristics of the household that will also influence its choice are its fertility type f , the availability of

¹⁶Rationing will be considered in later versions of this paper.

informal childcare oth and its tastes for home produced childcare and for leisure, g and α respectively. We summarise these by a seven-dimensional vector $\Omega_t = (t, w_t^W, w_t^F, f, oth, g, \alpha)$.

At each age t , the household has to choose female labour supply (lm_t), female leisure (L_t), and the use of home produced (dcc_t) or market (mcc_t) childcare. The labour supply decision is discrete ($lm_t \in \{0, 0.5, 1\}$) and the other choices are continuous. The three constraints that the household faces are the time constraint for the female spouse in equation (4), the need for childcare of current children (represented by (3)) and the budget constraint (shown in (5)):

$$lm_t + L_t + dcc_t = 1 \quad (4)$$

$$c_t + Ecc(t, f, y_t^{net}) = y_t^{net} \quad (5)$$

Since we rule out saving and borrowing, households consume all their income net of childcare expenditures. The household net income is calculated as $y_t^{net} = y_t^{gross} - T(y_t^{gross})$, where $T(\cdot)$ is the income tax schedule and the gross household income, y_t^{gross} , is given by:

$$y_t^{gross} = w_t^M + lm_t \cdot w_t^W \quad (6)$$

The dynamic household problem is defined for a given state vector Ω_t as:

$$V(\Omega_t) = \max_{lm_t, c_t, L_t, dcc_t} u(c_t, L_t, dcc_t | \Omega_t) + \beta V(\Omega_{t+1} | \Omega_t, lm_t) \quad (7)$$

subject to the three constraints above. This problem can be broken down into an intra-temporal choice and an inter-temporal choice.

The intra-temporal decision reflects the optimal time allocation between leisure and home produced childcare for a given female labour supply choice lm_t . The inter-temporal choice is to choose the optimal labour supply, lm_t , given the conditional optimal choices of leisure $L_t(lm_t)$ and home produced childcare $dcc_t(lm_t)$.

The optimal intra-temporal choice consists in solving the following static problem for a given discrete female labour supply decision $lm_t \in \{0, 0.5, 1\}$ and state vector Ω_t :

$$u^*(\Omega_t, lm_t) = \max_{c_t, L_t, dcc_t} u(c_t, L_t, dcc_t | \Omega_t, lm_t) \quad (8)$$

subject to the childcare constraint (3), the time constraint (4) and the budget constraint (5).

Finally, the optimal inter-temporal choice consists in maximising lifetime utility $V(\Omega_t)$ by choosing female labour supply lm_t :

$$V(\Omega_t) = \max_{lm_t} u^*(\Omega_t, lm_t) + \beta V(\Omega_{t+1}|\Omega_t, lm_t) \quad (9)$$

for given dynamics of the states $(\Omega_{t+1}|\Omega_t, lm_t)$. The model is solved by backward induction from retirement, where we assume that all income is used as consumption and the entire time endowed per period is used as leisure since, by construction, there are no children needing childcare in retirement.

4 Quantification of the Model

We quantify the model in five steps:

1. As the first step, we calibrate external parameter values.
2. Second, we estimate parameters of government policies that govern childcare prices and the tax system.
3. Then, we estimate the wage processes and the joint distribution of the household's female and male wage potentials given their age.
4. Fourth, we employ a multinomial logit model to estimate the distribution of fertility types.
5. Finally, we jointly estimate by maximum likelihood the distribution of heterogeneities in the preferences for home produced childcare and female leisure, as well as in the availability of informal childcare.

4.1 Calibration

In line with Blundell et al. 2016, we set the discount factor β to 0.94 and the leisure constant \bar{L} to 0.05. As outlined in the definition of the utility function (eq. (1)), we assume separability between consumption, leisure, and home produced childcare. We assume the coefficient of relative risk aversion γ to be 2, which is in the range of values found in the literature (e.g. Adda, Dustmann, and Stevens 2017).

Childcare need. Table 1 summarises the assumed childcare need for children of different ages: If the child is younger than 6 years, the childcare need (during working time) is set to 40 hours per week. For children aged 6 - 8, the need reduces to 15 hours per week because these children attend compulsory schooling for the remaining hours. We normalise these values by assuming that a normal working week has 40 hours. This yields for each child age i the age-specific time share of childcare needed within normal working hours, \bar{t}_i .

Table 1: Childcare need within normal working hours

Children's age interval	0 - 2	3 - 5	6 - 8	≥ 9
Hours of childcare needed per week	40	40	15	0
Normalized by 40h/week, \bar{t}_i	1	1	.375	0

4.2 Government policies

In this section, we calibrate the government policies that are required as exogenous inputs for our model: pension benefits, the German tax system and childcare prices.

4.2.1 Pensions

We approximate the German pension system by assuming that the households receive 50% of both partners' last period's potential gross full-time earnings throughout retirement, i.e. the pension share, \mathcal{B} , is equal to 0.5.

4.2.2 Taxes

In our estimation of German labour income taxes, we follow the approach of Lorenz and Sachs 2016. They map gross to net labour income on a monthly basis by applying the statutory German tax and transfer system of the year 2011 in detail. We adapt a functional form for the sum of taxes $T(y^{gross})$ that was originally introduced into the literature by Bénabou 2002 and is well-known as the HSV tax function from Heathcote, Storesletten, and Violante 2017. It maps gross to net income as follows: $y^{net} = y^{gross} - T(y^{gross}) = \lambda \cdot (y^{gross})^{1-\tau}$.

We determine the coefficients λ and τ based on the data for single households from Lorenz and Sachs 2016 by estimating the following model with Ordinary Least Squares:

$$\log(\tilde{y}^{net}) = \log(\lambda) + (1 - \tau) \cdot \log(\tilde{y}^{gross}) + \epsilon \quad (10)$$

where \tilde{y} denotes the income of single households in the data. We restrict our estimation to gross monthly labour incomes between 1,000 and 6,000 Euros. The estimated parameters $\lambda = 4.1259$ and $\tau = 0.2252$ are significantly different from zero and the estimation yields an R^2 of 0.995.

Since Lorenz and Sachs 2016 focus on single households, we need to adjust the tax function to take joint taxation of married couples' income in Germany into account. We assign both partners the average household income, $\tilde{y}^{gross} = y^{gross}/2$, and calculate their personal net income \tilde{y}^{net} by applying the estimated tax function for single households. Afterwards, we sum up both individual after-tax incomes.¹⁷ This yields the following formula for the *household net income of a married couple under joint taxation*:¹⁸

$$y_t^{net} = 2 \cdot \left[\lambda \cdot \left(\frac{y^{gross}}{2} \right)^{1-\tau} \right] \quad (11)$$

4.2.3 Childcare prices

In this section, we estimate childcare prices $p(i, K(t, f), y_t^{net})$ using GSOEP data. Recall that these prices depend on the age i of the child, the family structure represented by $K(t, f)$ and the household's net income y_t^{net} . A detailed description of the institutional background of the childcare price setting in Germany and its determinants can be found in Appendix C.1.1. With the estimation results at hand, we use the predicted childcare price functions as an input for our structural model.

Data We pool the core 2013 and 2015 GSOEP sample which contains data on childcare hours per day and monthly fees.¹⁹ ²⁰ Furthermore, we normalise the monthly fees by the reported childcare hours per day to extract the unit price of childcare normalised to 8 hours per day. For this purpose, we assume linearity of childcare prices with respect to hours.²¹

Empirical model. We propose a simple, preliminary linear model to estimate the childcare price function reflected in the model by $p(i, K(t, f), y_t^{net})$. We estimate equation (12) separately for each of the three age brackets $i = 1, 2, 3$, i.e. for children aged 0 - 2, 3 - 5 and 6 - 8. The

¹⁷Note that under joint taxation, tax functions are such that a married couple that earns together twice the amount of a single also faces twice the amount of income taxes.

¹⁸This approach is a first approximation. In a future version of this paper, we will estimate the tax function based on the statutory German tax system that (two-parent) households with children face.

¹⁹In terms of the sample construction, this estimation is based on the same sample as laid out in Section 2.2.

²⁰In Appendix C.1.3, we provide evidence for the data pooling.

²¹Figures 18, 19, and 20 in the Appendix C.1.4 show the dispersion of the normalised prices for the attended hours of childcare per day for children between 0 - 2, 3 - 5 and 6 - 8.

dependent variable p_{kjs} is the monthly price of full-time childcare paid by household j for child k in survey year s . The household indicator j summarises information about the family structure and household income.

$$\begin{aligned}
p_{kjs} = & \delta + \eta_1 \cdot \mathbb{1}\{\text{one sibling with age} < 17 \text{ in HH}\}_{js} \\
& + \eta_2 \cdot \mathbb{1}\{\text{two siblings with age} < 17 \text{ in HH}\}_{js} \\
& + \rho_0 \cdot y_{js}^{net} + \rho_1 \cdot y_{js}^{net} \cdot \mathbb{1}\{\text{one sibling with age} < 17 \text{ in HH}\}_{js} \\
& + \rho_2 \cdot y_{js}^{net} \cdot \mathbb{1}\{\text{two siblings with age} < 17 \text{ in HH}\}_{js} + \epsilon_{kjs}
\end{aligned} \tag{12}$$

As pointed out above, we estimate three sets of coefficients depending on the age bracket i . Furthermore, the predicted childcare prices depend not only on the child age, but also on the net household income and the number of siblings in the family. We restrict our estimation to children for which we observe positive childcare prices.

In the future, we intend to adjust this estimation to take selection into positive childcare prices into account. Furthermore, we plan to include regional price variation in the analysis. Since our current structural model does not cover either of these two aspects, for now, we adhere to the described setup.

Results The results of the estimations are summarised in Table 2. First of all, the constants reflecting base cost levels are significant except for the oldest children (6 - 8). Second, the coefficients on income (ρ_0) and the interaction terms between income and the number of siblings (ρ_1, ρ_2) are mostly significant which emphasises their importance as determinants of childcare prices. Third, the coefficients on the sibling dummies themselves, (η_1, η_2), are insignificant in most cases. However, if we estimate the same model without the interaction terms of income and sibling dummies, the remaining coefficients on income and sibling dummies (η_1, η_2, ρ_0) become highly significant. This clearly illustrates that family structure does matter for childcare prices, but the relationship appears to differ by income level. Therefore, the setup described in equation (12) remains our preferred specification for all children's age brackets.²²

Figure 7 visualises the estimated childcare price functions as well as the observed childcare prices for children below 3.²³ The effects of the covariates on the estimated childcare price functions are as expected: the price generally increases with income, but less so as the number

²²In Appendix C.1.3, we add a fixed year effect to the model which is insignificant for all age brackets.

²³The Figures 13 and 14 in the Appendix show the equivalent plots for children between 3 - 5 and 6 - 8.

Table 2: OLS estimation of childcare prices for different child age:

Dependent variable: p_{js}
Monthly price of market childcare normalised to full-time attendance (40h/week)

Child age	0 - 2	3 - 5	6 - 8
constant	66.34* (25.61)	78.71** (14.15)	54.70 (33.12)
1 sibling in HH (dummy)	10.20 (40.58)	19.86 (18.30)	90.81** (38.06)
2 siblings in HH (dummy)	3.66 (52.16)	-16.63 (22.22)	-10.68 (46.52)
Net HH income	0.0572** (.0077)	0.0282** (.0040)	.0435*** (.0110)
Net HH income x 1 other sibling in HH	-0.0187 (.0113)	-0.0132** (.0051)	-.0344** (.0117)
Net HH income x 2 other siblings in HH	-0.0257 (.0148)	-0.0128* (.0060)	-.0148 (.0136)
Observations	241	700	353
Adjusted R-squared	0.2685	0.1307	0.0835

Standard errors in parenthesis. ** p<0.01, * p<0.05, * p<0.1

Notes: Sample definition: Children attending daycare center for whom we also observe childcare fees (incl. 0) and attended childcare hours in 2013 and 2015. Monthly childcare fees are adjusted to full-time attendance of 40 hours per week. Household is abbreviated by HH. For age brackets 0 - 2, 3 - 5 and 6 - 8 we observe 12, 91 and 71 prices which are equal to 0. All prices are deflated to 2011.

of siblings increases, and it tends to be lower the more siblings live in the household.²⁴ The three estimations result in R^2 s ranging from 0.08 to 0.27. As mentioned above, the inclusion of regional indicators is expected to capture some of the remaining variation which is left for future work.

4.3 Wages: heterogeneity and dynamics

As a preliminary approximation, we calibrate the wage growth rates displayed in Table 3 as follows: In the GSOEP sample discussed in 2.2, we condition on observing a female in the same employment status for three consecutive years. Then we calculate their average growth rate in hourly wages over the entire three years, differentiating between full-time and part-time

²⁴The only estimated price function for which these results do not hold, is the one estimated on children aged 6 - 8 with 1 sibling which is very flat with a high constant.

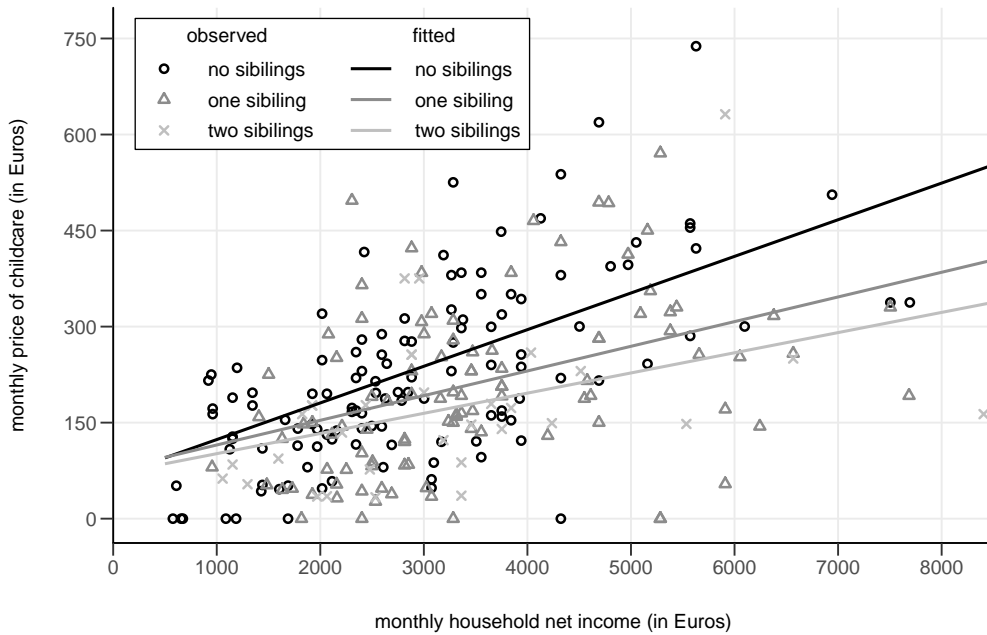


Figure 7: Observed unit prices of childcare and estimated price functions for child age 0 - 2

Notes: The price is normalised to 40h/week by the hours spent in childcare per week where 1 time unit equals 40h/week. Survey years 2013 and 2015 are pooled which leads to 253 observations of which 12 have a price of 0. We restrict the analysis to the core GSOEP sample. The prices are deflated to 2011.

work, to use them as per period wage growth in our model. Additionally, we set the per period wage growth if not working to -0.0333 , which corresponds to the per-period human capital depreciation rate found by Guner, Kaygusuz, and Ventura 2018.²⁵

Table 3: Wage growth

employment status	per period growth rate
full-time	0.0535
part-time	0.0332
not working	-0.0333

4.4 Distribution of fertility types

In our theoretical framework, we assume that households make choices in the knowledge of their fertility type, i.e. they anticipate accurately the mother's age at the time of the birth of her first child, a , and the number of children the household will have eventually, i.e. completed

²⁵The authors calibrate a human capital depreciation rate of 0.055 using US data in a setting where one period corresponds to five years. Assuming that human capital directly translates into wage growth, we convert their number into our three year setting.

fertility size, n . Each fertility type is defined as a pair (a, n) . We assume that first children are born to mothers aged 20 to 40. Since our unit time period is 3 years, we group values of a in ranges of three years. Besides, only few families have more than 3 children, so we focus on households with up to 3 children, i.e. $n \in \{0, \dots, 3\}$. Fertility type f will be defined as

$$f = \text{int} \left(\frac{a - 17}{3} \right) \cdot 3 \cdot \mathbb{1}_{n>0} + \min\{n, 3\} + \mathbb{1}_{n=0}$$

for households below the age of 35. Moreover, we assign to the classes $f = 17$ and 18 households who start having children in the age range 35-37 and have one and two children, respectively. The single class $f = 19$ contains all households with children who start childbearing in the age range 38-40 and have 1 child. This leaves us with 19 fertility classes that map directly into our structural model.

The fertility type is a constant and partially observed characteristic of households. More specifically, for households where the female spouse is aged over 41, f is perfectly observed. However, when she is under 40, we may know the age at which she bore her first child if it is already born, but we may not know the size of her completed fertility. Our strategy will thus be to estimate determinants of the probabilities to belong to each fertility class with the subsample of households aged over 41 and use these estimates to predict the fertility types of households aged under 40. This leaves us with a probability distribution for every household over the 19 fertility classes. Then, we also use the available information for households under 40 and exclude fertility classes that are impossible given the current age of the female spouse, her age at first birth and the presence of children.

To estimate the distribution of fertility types in the subsample of mothers aged 41 and older, we use a multinomial logit model to estimate the probabilities of belonging to each of the 19 possible fertility classes f for household i as:

$$\text{Prob}(f_i = f | \mathbf{w}_i) = \frac{\exp(\mathbf{w}'_i \alpha_f)}{1 + \sum_{k=1}^{19} \exp(\mathbf{w}'_i \alpha_k)}, \quad (13)$$

where \mathbf{w}'_i is a vector of observed characteristics for household i and α_f is a vector of parameters reflecting the impact of these characteristics on the probability of belonging to fertility type f .

We use four characteristics in the vector \mathbf{w} , namely an indicator for being religious (we assume that religion is constant over the lifecycle), an indicator for having completed A-levels,

a term in which we interact completed A-levels with having had at least 14 years of education throughout life, and a linear time trend.

We estimate this multinomial logit model with the subsample of households where the female spouse is aged 41 or above, which gives us a sample size of 5,291 women.²⁶ Figures illustrating the fit of the model are reported in Appendix C.2.1.²⁷ The main insights emerging from these results are threefold: first, households who report being religious tend to have more children. Second, mothers with more and longer education tend to have fewer children, to start considerably later. Finally, the time trend indicates that later cohorts tend to have children later and larger completed family size.

Table 4: Predicted fertility distributions in- and out-of-sample

Fertility class	a	n	in-sample (in %)	out-of-sample (in %)
1	-	0	17.26	17.50
2	20 - 22	1	4.78	1.08
3	20 - 22	2	13.17	4.78
4	20 - 22	3	5.03	5.15
5	23 - 25	1	5.41	1.55
6	23 - 25	2	11.72	6.38
7	23 - 25	3	3.78	4.37
8	26 - 28	1	4.78	2.56
9	26 - 28	2	9.13	10.12
10	26 - 28	3	2.78	4.39
11	29 - 31	1	3.84	3.54
12	29 - 31	2	6.20	13.28
13	29 - 31	3	1.40	2.41
14	32 - 34	1	3.10	4.49
15	32 - 34	2	2.97	5.95
16	32 - 34	3	0.34	2.07
17	35 - 37	1	2.14	3.30
18	35 - 37	2	1.15	3.99
19	38 - 40	1	1.04	3.09

Notes: In-sample definition: women aged 41 to 65, born between 1950 and 1975 and not in education, conditional on first birth (if any) only between 20 and 40 and having completed fertility by age 41. Out-of-sample definition: women aged 20 to 40, born between 1976 and 1995 and not in education, conditional on first birth (if any) only between 20 and 40. Source: 2000 to 2016 core GSOEP samples.

²⁶A detailed description of the precise sample underlying the estimation can be found in Appendix A.

²⁷We refrain from plotting estimation output tables because of their size. They are available upon demand.

We then use these estimates to predict fertility types for the households in which the female spouses are below the age of 41 when they are observed the last time in the data. One caveat of these predictions is that we have assumed that the time trend term is constant out-of-sample and equal to the year 1976, i.e. the birth year of the first out-of-sample cohort. In other words, conditional on religion and education, we have assumed that the predicted probabilities of belonging to each fertility type are constant for younger cohorts.²⁸

As detailed above, we next exclude fertility classes that are impossible given the available information on these younger households and normalise the remaining probabilities. For example, a mother who had a first child at the age of 28 can only belong to fertility types 8, 9 or 10. We keep only those probabilities and normalise them to add up to 1. Table 4 reports the resulting fertility distributions across the 19 types for the sample with age above 41 (in-sample) and those below (out-of-sample).

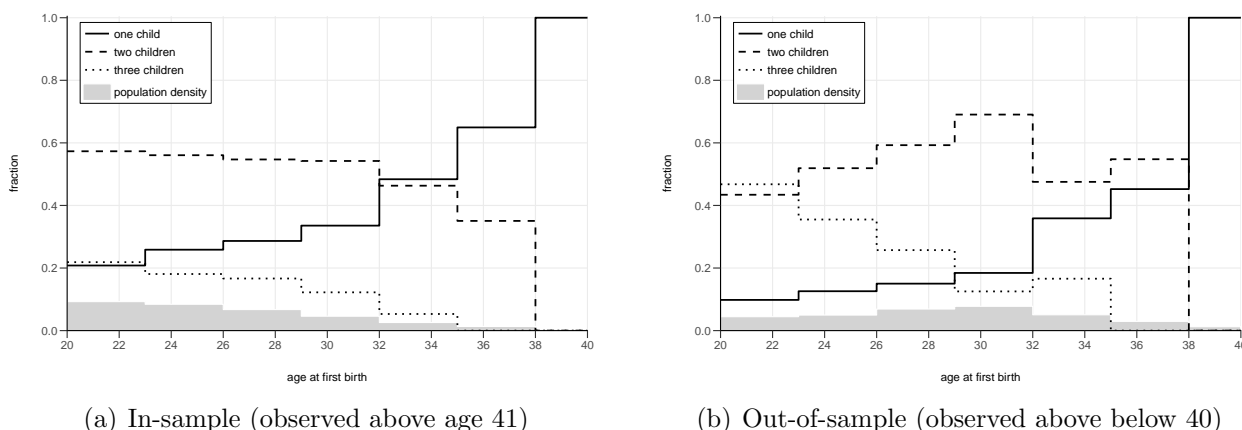


Figure 8: Predicted completed fertility across age at first birth

Notes: In-sample definition: women aged 41 to 65, born between 1950 and 1975 and not in education, conditional on first birth (if any) only between 20 and 40 and having completed fertility by age 41. Out-of-sample definition: women aged 20 to 40, born between 1976 and 1995 and not in education, conditional on first birth (if any) only between 20 and 40. Source: 2000 to 2016 core GSOEP samples.

Figures 8(a) plots the predicted completed fertility across age at first birth for the sample used in the estimation. We can compare this graph to Figure 6 in Section 2.3 and their overall patterns look very similar, the only difference occurring since we use age bins of 3 years in our estimation bins of 1 year in the earlier section. Additionally, 8(b) shows the predicted

²⁸The reason we have made this *ad hoc* assumption is that extrapolating the time trend would lead us to predict that a substantial share of households would have their first child after age 29 and have three children which did not seem credible.

completed fertility across age at first birth for the women observed under the age of 40. We predict that more women get their first child later.

Furthermore, we assign to each of these households the fertility type that is most likely. In the following, we will proceed with fertility types assigned to each household, defined as the actual fertility types for households aged over 41 and the predicted fertility type for younger households. This will be used as an input in the next section of this paper.

4.5 Maximum Likelihood Estimation

4.5.1 Methodology

In our framework, the preference for home produced childcare g , the availability of informal childcare oth and female leisure preference α are all driving forces of female labour supply and childcare decisions during parenthood. Therefore, they determine the female labour supply elasticities in which we are ultimately interested. In this subsection, we lay out an estimation procedure to identify the underlying joint distribution of unobserved heterogeneities.

We proceed as follows: we start out by making functional assumptions on the marginal distributions of the three heterogeneities as well as their correlation structure. Then we estimate by maximum likelihood (ML) the parameters governing the marginal distributions and their correlations, i.e. the joint distribution over (g, oth, α) .²⁹

Marginal distributions of each type of heterogeneity. To determine the marginal probability density functions (PDFs) h_g , h_{oth} and h_α , we need to make some functional form assumptions: currently, we assume first that h_{oth} can be approximated by a beta distribution where $oth \in [0, 1]$. Second, we assume that h_g and h_α are distributed following (truncated) gamma distributions.³⁰

We allow the distribution parameters to depend on a vector of covariates x . As an example, consider the estimation of h_{oth} using one covariate x_{oth} : The parameters θ_{oth} , estimated by

²⁹In a future version of the paper, we will employ a dynamic estimation approach by matching the labour supply and childcare choices over consecutive periods.

³⁰These functional form assumptions are preliminary and might very well change in future versions of the paper.

maximum likelihood, are used to calculate the shape parameters a_{oth} and b_{oth} of the beta distribution h_{oth} conditional on x_{oth} :

$$a_{oth}(x_{oth}, \theta_{oth,1}^a, \theta_{oth,2}^a) = (\theta_{oth,1}^a + \theta_{oth,2}^a \cdot x_{oth})^2 \quad (14)$$

$$b_{oth}(x_{oth}, \theta_{oth,1}^b, \theta_{oth,2}^b) = (\theta_{oth,1}^b + \theta_{oth,2}^b \cdot x_{oth})^2 \quad (15)$$

The covariate x_{oth} takes the value one if the couple lives in one of the partners' home towns, and zero otherwise. This is a proxy for the availability of grandparents, other relatives and friends who might provide informal and free childcare. For the estimation of the marginal PDFs of g we use a dummy x_g which reflects the education level of the woman's mother: we indicate with 1 if she has either A-levels without studies or a university degree and a 0 otherwise.³¹ Finally, for the marginal distribution of the leisure preference α , we are not using covariates in the current version.

Joint distribution. For the determination of the conditional joint PDF h , we follow two different alternative approaches. First, we assume independence of the marginal PDFs. Second, we assume that the marginal PDFs are non-linearly correlated via a Gaussian copula.

If we assume *independent* marginal PDFs, the conditional joint PDF is simply given by

$$h(g, oth, \alpha | x_n, \theta) = h_g(g | x_{n,g}, \theta_g) \cdot h_{oth}(oth | x_{n,oth}, \theta_{oth}) \cdot h_\alpha(\alpha | x_{n,\alpha}, \theta_\alpha)$$

To allow for *non-linearly correlated distributions of heterogeneities*, we include a copula C to link the marginal PDFs of g , oth and α .³² By Sklar's Theorem, we can describe the joint (conditional) cumulative distribution functions (CDFs) H by the marginal CDFs H_g, H_{oth}, H_α , and a copula C , which we assume to be Gaussian, the joint PDF h is given by:³³

$$h(g, oth, \alpha | x_n, \theta, \Sigma) = h_g(g | x_{n,g}, \theta_g) \cdot h_{oth}(oth | x_{n,oth}, \theta_{oth}) \cdot h_\alpha(\alpha | x_{n,\alpha}, \theta_\alpha) \cdot \underbrace{\phi_\Sigma \left[\Phi^{-1}(H_g(g | x_{n,g}, \theta_g)), \Phi^{-1}(H_{oth}(oth | x_{n,oth}, \theta_{oth})), \Phi^{-1}(H_\alpha(\alpha | x_{n,\alpha}, \theta_\alpha)) \right]}_{\text{derived from the Copula } C = \Phi_\Sigma}$$

³¹We observe 54 grandmothers (out of 250) with high education. In addition, for 179 of the 250 households in our sample we observe the couple to live in one of the partners' home town. For the sample definition, see Section 4.5.3 and more detailed in the Appendix A.

³²A copula C is a joint CDF of random variables that are marginally uniformly distributed on $[0, 1]$. Using the marginal CDFs as inputs to the copula allows us to express the joint distribution of (g, oth, α) as a function of the marginals.

³³Note that C, H_g, H_{oth} , and H_α are all differentiable.

Note that Σ represents the covariance matrix of the three-dimensional normal distribution with PDF ϕ_Σ , while Φ^{-1} is the inverse of the CDF of the one-dimensional standard normal distribution.

Predicted choices. In the data, we can directly observe the female labour market choice lm_n and the total amount of market childcare used by the household mcc_n . To relate the household choices in the data to the model, we also need to determine the household's remaining state variables, denoted by $s_n = (t_n, w_{t_n,n}^M, w_{t_n,n}^W, f_n)$. We can directly observe the age of the household t_n as well as the male wage $w_{t_n,n}^M$ because we condition on men working full-time. In contrast, the female wage potential $w_{t_n,n}^W$ and fertility type f_n might be unobserved. We lay out our estimation below and will use their predicted values when these are unobserved.

Given s_n , the vector of heterogeneities (g, oth, α) and the optimal policies from the model, we denote the model-predicted household behaviour in terms of female labour supply and use of market childcare as $\widehat{lm}(s_n, g, oth, \alpha)$ and $\widehat{mcc}(s_n, g, oth, \alpha)$. Then, the model-data match is defined as an indicator function³⁴

$$m(lm_n, mcc_n | s_n, g, oth, \alpha) = \begin{cases} 1 & \text{iff } \widehat{lm}(s_n, g, oth, \alpha) = lm_n \text{ and} \\ & \widehat{mcc}(s_n, g, oth, \alpha) = mcc_n \\ 0 & \text{otherwise.} \end{cases} \quad (16)$$

Likelihood. Given the components introduced above, we can now write the *household likelihood contribution*:

$$l(lm_n, mcc_n | x_n, s_n, \theta, \Sigma) = \int \int \int_{\mathcal{S}} \underbrace{h(g, oth, \alpha | x_n, \theta, \Sigma)}_{\text{joint PDF of } g, oth \text{ and } \alpha} \cdot \underbrace{m(lm_n, mcc_n | s_n, g, oth, \alpha)}_{\text{model-data match}} dg \cdot doth \cdot d\alpha \quad (17)$$

where \mathcal{S} is the state space of (g, oth, α) . The likelihood contribution consists of two components: first, the joint PDF h , our outcome of interest, and second, an indicator for a model-data match.

³⁴Note that it does not matter if we somehow normalize the current weights (even household-specific normalization works) because in the end, we take the logarithm of l and therefore normalization only shifts the objective function up or down, but it does not affect the estimated parameter values. However, this is not true if the change in weights is not multiplicative.

The parameters θ and Σ are chosen to maximise the *sample log-likelihood* given by

$$\mathcal{L}(\theta, \Sigma; lm, mcc | x, s) = \sum_{n=1}^N \log l(lm_n, mcc_n | x_n, s_n, \theta, \Sigma) \quad (18)$$

Note that the estimation is conditional on all households' observed or estimated characteristics and choices s , x , lm , and mcc where as an example $s = [s_1, \dots, s_N]$, $n = 1, \dots, N$.

4.5.2 Identification

We do not have a formal proof of identification, but we give here an intuition for identification of the parameters that govern the distribution of heterogeneities instead. For simplicity, let us consider a household with one child below 3. Given observed female labour market choices and market childcare used by the household, we observe four different groups of households that allow us to disentangle the different preference heterogeneities:

1. First, the woman works and the amount of market childcare is sufficient to cover her working time.
2. Second, the woman does not work and, nonetheless, the household consumes a positive amount of market childcare.
3. Third, the woman works, but the household does not buy any or only an insufficiently low amount of market childcare.
4. Finally, the woman does not work, and does not buy any market childcare, i.e. home produced childcare covers the full childcare need.

Otherwise equivalent households that buy the same amount of market childcare, but choose different female labour supply (1. vs. 2.) help us to identify female leisure heterogeneity. To identify informal childcare availability, the setup allows us to compare households with the same female work decisions, but who choose different amounts of market childcare (1. vs. 3.). Lastly, home produced childcare preferences can be identified by comparing households with the same female labour supply, but different market childcare choices (2. vs. 4.) as well as by contrasting households with different labour supply and market childcare choices (1. vs. 4.).

4.5.3 Sample definition and auxiliary regressions

Sample definition. We conduct the ML estimation with the 2016 cross-section of the sample described in Section 2.2, using all females with full-time working husbands whose fertility type we can classify and that have children with childcare need according to Table 1.³⁵

Estimation of female wage potential. Since we do not have information about wages for all women, we need to run auxiliary regressions to predict wages for these observations. First, we calculate for all women for whom we observe monthly labour income and agreed working hours per week, the hourly wage rate:³⁶

$$\text{hourly wage} = \frac{\text{income per month}}{\text{agreed working hours per week}} \frac{5 \cdot 12}{250} \quad (19)$$

Then, by using OLS, we regress log hourly wages onto a second-order polynomial of past full-time and part-time experience, different education levels, and dummies that indicate first, whether a child below the age of 6 lives in the household and second, if either one or two (and more) children below the age of 17 are present:³⁷

$$\begin{aligned} \log w_j = & \delta + \beta_1 \cdot FTexp_j + \beta_2 \cdot FTexp_j^2 + \gamma_1 \cdot PTexp_j + \gamma_2 \cdot PTexp_j^2 \\ & + \mu_1 \cdot \mathbb{1}\{\text{Some school degree and vocational training}\}_j \\ & + \mu_2 \cdot \mathbb{1}\{\text{A-Levels without studies}\}_j + \mu_3 \cdot \mathbb{1}\{\text{University degree}\}_j \\ & + \kappa_1 \cdot \mathbb{1}\{\text{at least one child} < 6 \text{ years old in household}\}_j \\ & + \kappa_2 \cdot \mathbb{1}\{\text{one child} < 17 \text{ years old in household}\}_j \\ & + \kappa_3 \cdot \mathbb{1}\{\text{two or more children} < 17 \text{ years old in household}\}_j + \epsilon_j \end{aligned} \quad (20)$$

where j indicates the woman. The baseline education in this regression is 'no or some school degree (without vocational training)'. The results are shown in the Appendix C.3.1 in Table 6.

For those women for whom we do not observe either gross monthly labour income or agreed hours of work per week (as they are not in work), we predict their potential hourly wage using the coefficient estimates obtained above. Otherwise, we use the observed hourly wage directly. To generate gross monthly labour income in case of full-time work as our ingredient for the

³⁵A detailed description of the precise sample underlying the estimation can be found in Appendix A.

³⁶We normalize the gross monthly labour income by agreed working hours per day assuming 250 working days per year.

³⁷In a future version, we will take selection into account.

structural model, we use the inverse of the normalization in equation (19), replacing agreed hours with 40 hours per week.³⁸

Estimation of fertility types. To predict the fertility type f_n for all households, we use the estimation results from the multinomial logit model in Section 4.4 to predict the probability to belong to each of the fertility classes for every household. Then, we exclude some classes that are not feasible given the already realised fertility. Finally, among the remaining classes we pick for each household the one with the highest probability.

4.5.4 Results

Preliminary estimation results show that, as intuitively expected, if the couples live in one of the partners' home towns, the distribution of informal care availability features more mass for higher values of oth compared to the distribution of couples that do not live in either of their home towns. Furthermore, we also observe that for higher education levels of the female's mother, women tend to have a lower preference for home produced childcare. These estimated distributions are essential inputs into the fiscal calculation which are work in progress.

[WORK IN PROGRESS]

5 Model Fit

[WORK IN PROGRESS]

6 Fiscal calculation and Policy Experiments

[WORK IN PROGRESS]

7 Conclusion

[WORK IN PROGRESS]

³⁸This is a very preliminary first step. In a future version, we will estimate a more elaborate wage process in Section 4.3 and use it to assign hourly wages for the ML-estimation as well.

References

- Adda, J., C. Dustmann, and K. Stevens (2017). “The Career Costs of Children”. *Journal of Political Economy* 125.2, 293–337.
- Angelov, N., P. Johansson, and E. Lindahl (2016). “Parenthood and the Gender Gap in Pay”. *Journal of Labor Economics* 34.3, 545–579.
- Authoring Group Educational Reporting (2018). *Education in Germany 2018*.
- Baker, M., J. Gruber, and K. Milligan (2008). “Universal child care, maternal labor supply, and family well-being”. *Journal of Political Economy* 116.4, 709–745.
- Bastani, S., S. Blomquist, and L. Micheletto (2017). “Child Care Subsidies, Quality, and Optimal Income Taxation”. Working Paper Series 6533. CESifo.
- Bauernschuster, S. and M. Schlotter (2015). “Public child care and mothers’ labor supply - Evidence from two quasi-experiments”. *Journal of Public Economics* 123, 1–16.
- Bénabou, R. (2002). “Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?” *Econometrica* 70.2, 481–517.
- Bettendorf, L. J., E. L. Jongen, and P. Muller (2015). “Childcare subsidies and labour supply? Evidence from a large Dutch reform”. *Labour Economics* 36, 112–123.
- 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.
- Blundell, R., M. Costa Dias, C. Meghir, and J. Shaw (2016). “Female Labor Supply, Human Capital, and Welfare Reform”. *Econometrica* 84.5, 1705–1753.
- Cascio, E. U., S. J. Haider, and H. S. Nielsen (2015). “The effectiveness of policies that promote labor force participation of women with children: A collection of national studies”. *Labour Economics* 36, 64–71.
- Colas, M., S. Findeisen, and D. Sachs (2018). “Optimal Need-Based Financial Aid”. Discussion Paper 13196. CEPR.
- Domeij, D. and P. Klein (2013). “Should Day Care be Subsidized?” *The Review of Economic Studies* 80.2, 568–595.
- Gathmann, C. and B. Sass (2018). “Taxing Childcare: Effects on Childcare Choices, Family Labor Supply, and Children”. *Journal of Labor Economics* 36.3, 665–709.
- Geyer, J., P. Haan, and K. Wrohlich (2015). “The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach”. *Labour Economics* 36, 84–98.

- Givord, P. and C. Marbot (2015). “Does the cost of child care affect female labor market participation? An evaluation of a French reform of childcare subsidies”. *Labour Economics* 36, 99–111.
- Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2018). “The German Socio-Economic Panel (SOEP)”. *Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics* aop.
- Guner, N., R. Kaygusuz, and G. Ventura (2018). “Child-Related Transfers, Household Labor Supply and Welfare”.
- Heathcote, J., K. Storesletten, and G. L. Violante (2017). “Optimal Tax Progressivity: An Analytical Framework”. *The Quarterly Journal of Economics* 132.4, 1693–1754.
- Ho, C. and N. Pavoni (2016). “Efficient Child Care Subsidies”. Working Paper 572. IGIER, Bocconi University.
- Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimueller (2019). “Child Penalties Across Countries: Evidence and Explanations”. Discussion Paper 13474. CEPR.
- Kleven, H., C. Landais, and J. E. Søgaard (2018). “Children and Gender Inequality: Evidence from Denmark”. Working Paper 24219. National Bureau of Economic Research.
- Lorenz, N. and D. Sachs (2016). “Identifying Laffer Bounds: A Sufficient-Statistics Approach with an Application to Germany”. *The Scandinavian Journal of Economics* 118.4, 646–665.
- Nollenberger, N. and N. Rodríguez-Planas (2015). “Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain”. *Labour Economics* 36, 124–136.
- Rainer, H., W. Auer, S. Bauernschuster, N. Danzer, A. Fichtl, T. Hener, C. Holzner, J. Reinkowski, and M. Werding (2013). “Öffentlich geförderte Kinderbetreuung in Deutschland: Evaluierung der Auswirkungen auf die Arbeitsmarktbeteiligung von Müttern”. *ifo Schnelldienst* 66.07, 31–40.
- Turon, H. (2018). “Home production of childcare and labour supply decisions in a collective household model”.
- Wang, H. (2019). “Fertility and Family Leave Policies in Germany: Optimal Policy Design in a Dynamic Framework”.

Appendix

A Estimation Samples

To conduct the ML estimation with the 2016 cross-section of the sample described in Section 2.2, we require non-missing information for childcare take-up, hours, and fees for the first three children that live in the household. Additionally, we condition on observing the education of the woman's mother, whether the family lives in one of the partners' home town, past full-time and part-time experience, maternal religious status as well as education and length of education of the woman. Finally, we drop households that are either in the top 1% or bottom 1% of the male or female wage percentiles and non-married couples.

To conduct the fertility type estimation, we restrict the sample described in Section 2.2 to women born between 1950 and 1995. We keep only the last observation for each woman, i.e. we restrict the sample to all women observed in 2016 regardless of their age plus those that were observed the last time before 2016 being at least 41 at that point. In addition, we restrict ourselves to women who had their first child between 20 and 40 (if they had one) and who had completed fertility at the age of 41. The fertility type prediction is done using the restrictions on the sample as the fertility type estimation, except that we include women younger than 41 that may still have children in the future.

B Additional Tables and Figures - Stylised Facts

Table 5: Population shares of mothers' working hours vs. childcare hours (per week)

	children aged 0-2	children aged 3-5
zero childcare, zero working hours	0.5152	0.0370
childcare = working hours +/- 5	0.0918	0.2673
childcare < working hours - 5	0.1159	0.0575
childcare > working hours + 5	0.2771	0.6382

Notes: Sample definition: mothers aged 20 to 65 that are not in education, conditional on having a child aged zero to five, the child not being taken care of by a paid external care taker and the partner working full-time. Source: 2009 to 2016 core GSOEP samples.. Source: 2000 to 2016 core GSOEP samples.

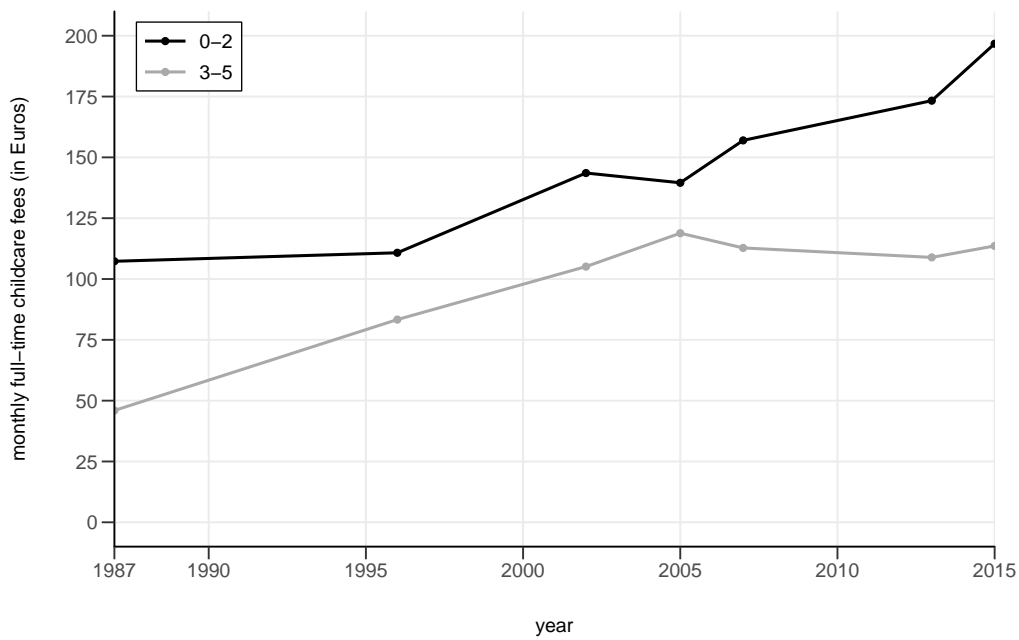


Figure 9: Monthly full time childcare fees over time by age of children

Notes: Sample definition: Children attending a public childcare institution for whom we also observe childcare fees (incl. zero fee observations). From 1987 - 2007, conditional on full-time attendance (no hours specified). In 2013 and 2015, conditional on attending ≥ 30 hours per week. Fees per month are actual fees, and not adjusted by hours. Source: core GSOEP samples from years 1987, 1996, 2002, 2005, 2007, 2013 and 2015, in 2011 prices.

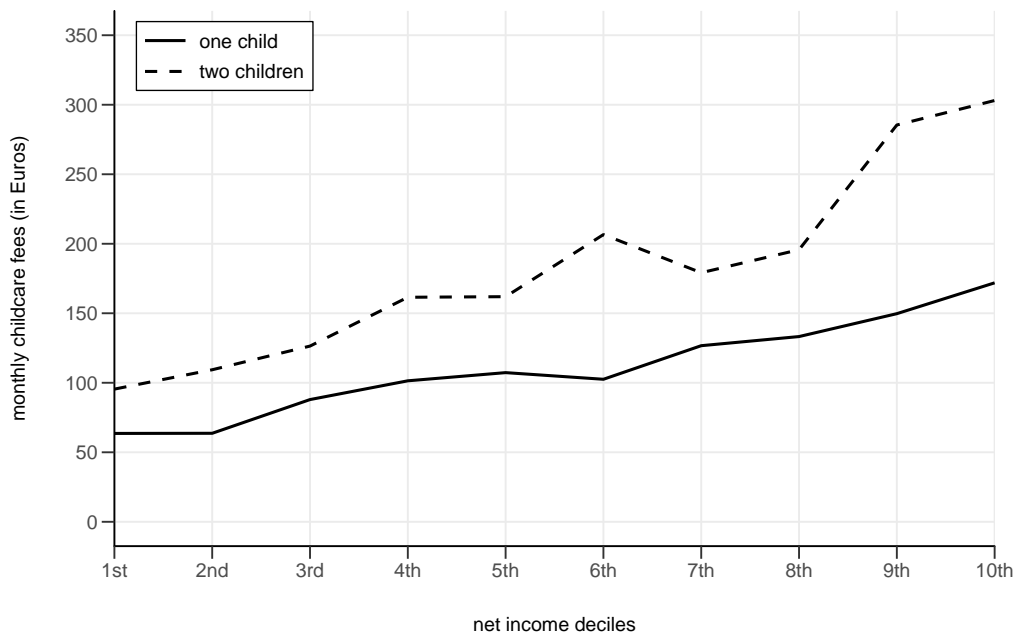


Figure 10: Actual childcare fees by number of children

Notes: Sample definition: mothers aged 20 to 65 that are not in education, conditional on childcare attendance and observed childcare fees (incl. zero fee observations). Source: 2013 and 2015 core GSOEP samples, in 2011 prices.

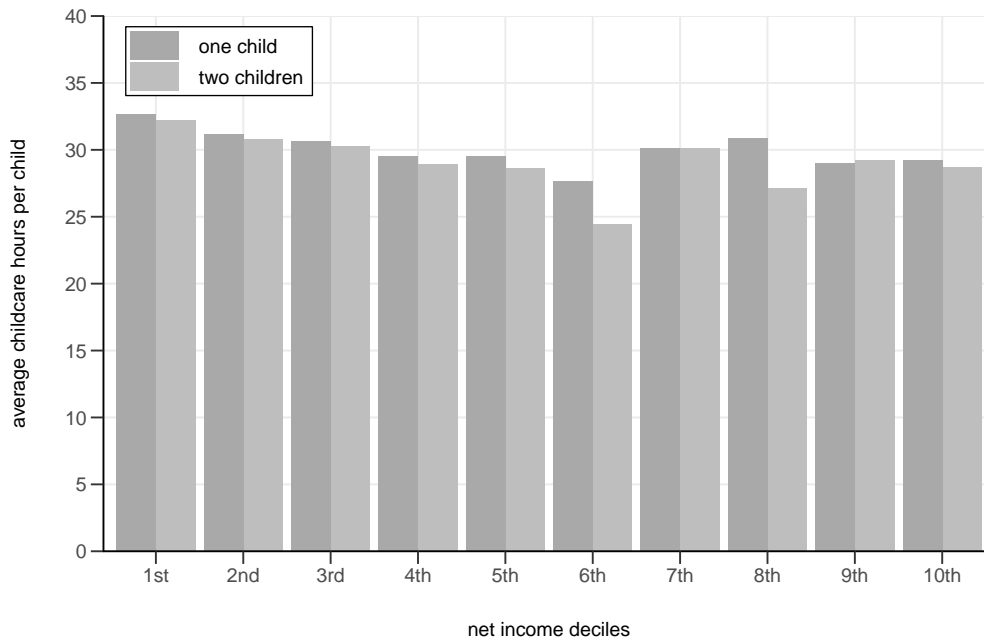


Figure 11: Average childcare hours by number of children

Notes: Sample definition: mothers aged 20 to 65 that are not in education, conditional on childcare attendance and observed childcare fees (incl. zero fee observations). Source: 2013 and 2015 core GSOEP samples, in 2011 prices.

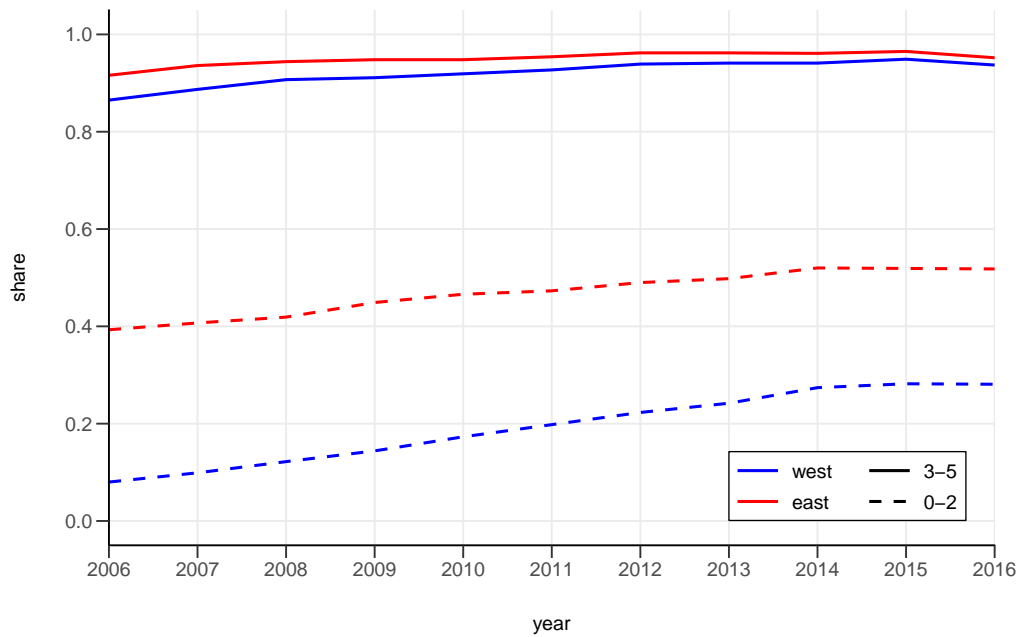


Figure 12: Childcare enrollment by region - based on administrative data

Notes: Enrollment is binary in the sense that it is not conditional on a minimum number of hours. Sample definition: German population of children aged 5 and below. Source: Authoring Group Educational Reporting 2018, data from the German Statistical Office.

C Quantification of the Model

C.1 Childcare price estimation

C.1.1 Determinants of childcare prices

Child age. One of the important determinants of childcare prices in Germany is the age of the child. This is due to the fact that children of different ages visit different childcare institutions, as laid out in Section 2.1. Prices are usually higher for younger children since costs of nurseries are higher than those of kindergartens.

Regional variation. Childcare prices in Germany differ further on a regional level because of two reasons: First, the price schedules are set discretionally on a municipality level. Second, the German constitution ensures that the federal states bear the political responsibility for their respective education systems. Since childcare is part of the education system, different federal states have implemented different childcare pricing regulations. So far, we do not include regional variation in our empirical and structural model, but we intend to do so in the future. Currently, 11 of the 16 federal states allow for partial or full reduction of parental contributions conditional on the children's age cohort and hourly use of day care.³⁹

Further determinants of childcare prices. Despite their autonomy, the different states define in their legislation vastly similar determinants of childcare prices besides child age:⁴⁰

1. *Household income:* In 11 out of 16 states the household income is used as a determinant and in 2 out of 16 states it can be used.⁴¹
2. *Number of children in the household:* In 12 out of 16 states, childcare prices are determined conditional on the number of children in the household. Furthermore, in one additional state it can be used optionally as a determinant and two further states condition on the number of children in the household that attend nursery or kindergarten.

³⁹See Authoring Group Educational Reporting 2018: *Education in Germany 2018*, Section C2, p. 70–71 and Tables C2-3a.

⁴⁰See Authoring Group Educational Reporting 2018: *Education in Germany 2018*, Section C2, p. 70–71 and Table C2-15web.

⁴¹Note that different municipalities differ in their definitions of household income: First, municipalities differ in using the net or gross household income. Second, they also condition on household income of different years (e.g. current year versus previous years).

3. *Agreed hours of day care:* In 8 out of 16 states, the legislation also requires the prices to be determined on the agreed hours of day care between the parents and the day care centre/kindergarten.

C.1.2 Results: Childcare prices estimation for children 3 - 5 and 6 - 8

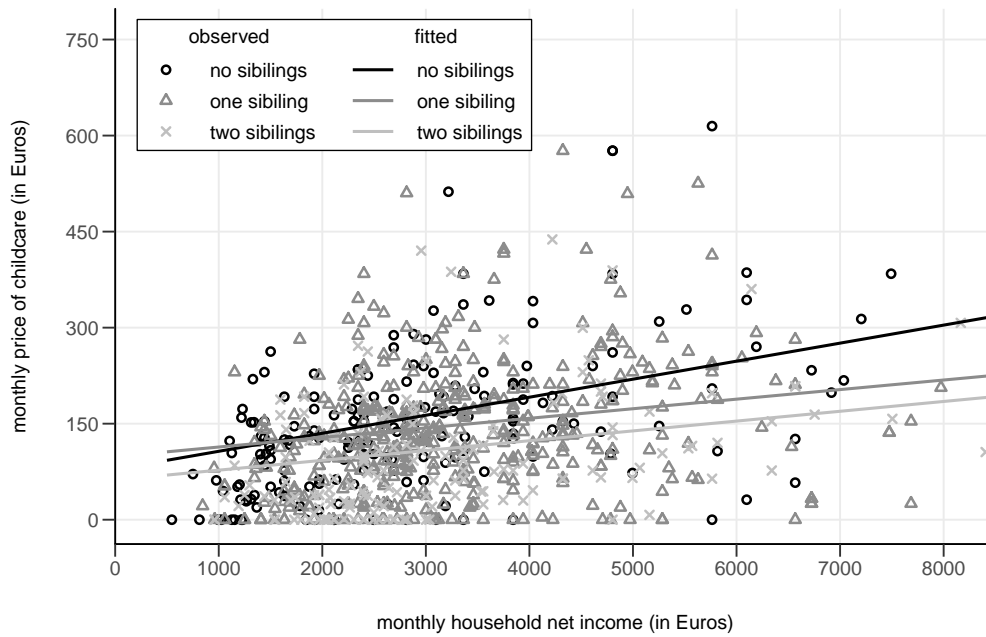


Figure 13: Observed unit prices of childcare and estimated price functions for child age 3 - 5

Notes: The price is normalised to 40h/week by the hours spent in childcare per week where 1 time unit equals 40h/week. Survey years 2013 and 2015 are pooled which leads to 791 observations of which 91 have a price of 0. We restrict the analysis to the core GSOEP sample. The prices are deflated to 2011

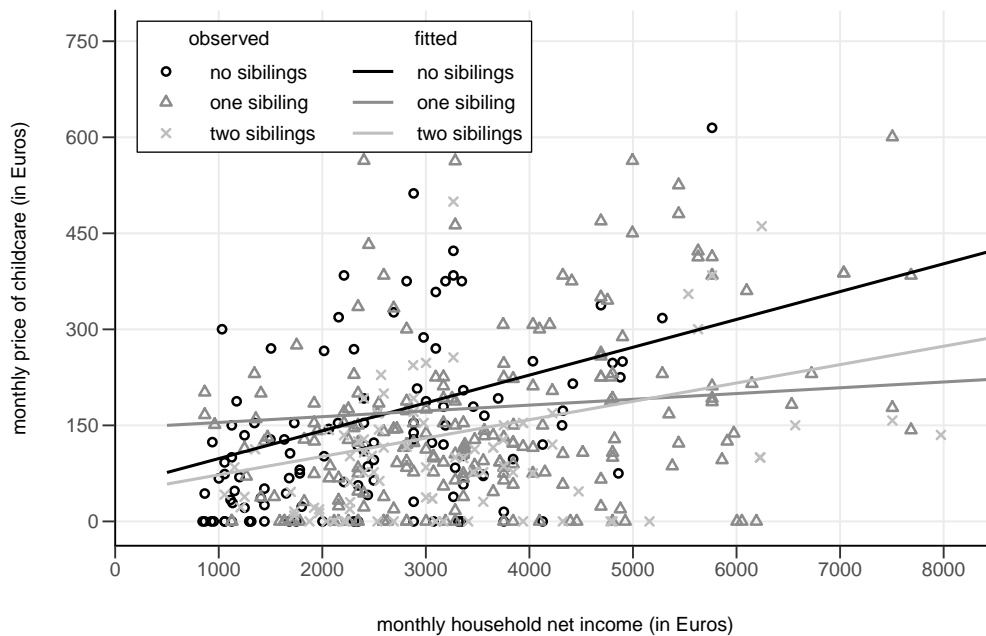


Figure 14: Observed unit prices of childcare and estimated price functions for child age 6 - 8

Notes: The price is normalised to 40h/week by the hours spent in childcare per week where 1 time unit equals 40h/week. Survey years 2013 and 2015 are pooled which leads to 424 observations of which 71 have a price of 0. We restrict the analysis to the core GSOEP sample. The prices are deflated to 2011

C.1.3 Data pooling – Unit childcare prices conditional on years 2013 and 2015

To rationalise the data pooling of the years 2013 and 2015, we first show in Figures 15, 16, and 17 that the prices in 2013 and 2015 exhibit no visible differences between the years. In Figure 9, we also plot the average childcare prices over time. Even there, we see no effect on price for children aged 3 - 5, but an increase for younger children. Nonetheless, this can be due to a change in the composition of households. And indeed, if we add a year fixed effect $\kappa \cdot \mathbb{1}\{t = 2015\}_t$ to the empirical model, the coefficient κ remains insignificant for all age groups indicating that there is no significant difference of childcare prices between the years 2013 and 2015.

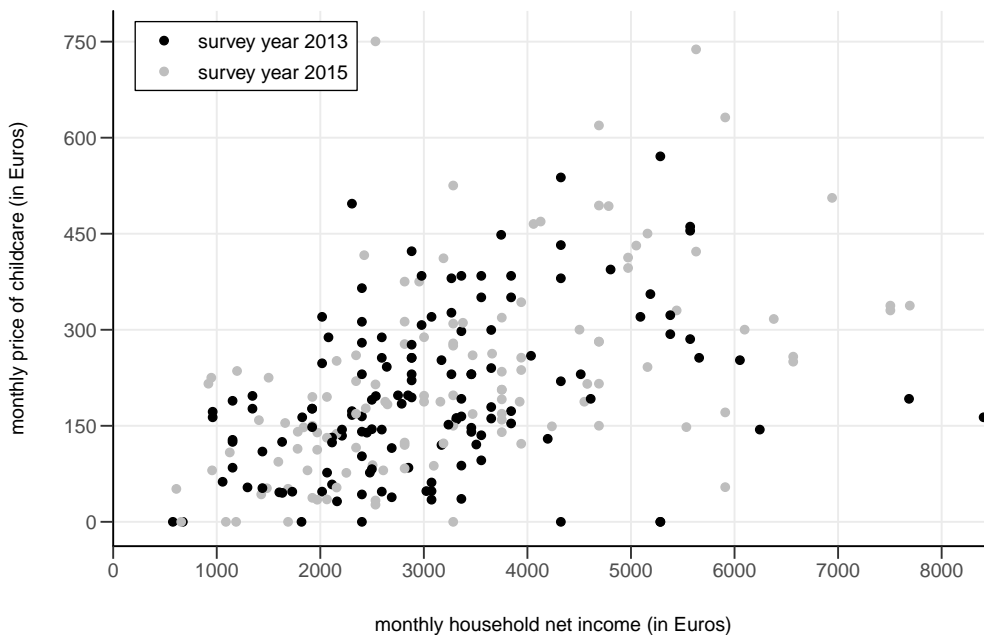


Figure 15: Unit childcare prices for children aged 0 - 2 conditional on years 2013 and 2015

Notes: The price is normalised to 40h/week by the hours spent in childcare per week where 1 time unit equals 40h/week. The survey years 2013 (134 observations) and 2015 (119 observations) are separated in this graph. Of those observations, 7 (2013) and 5 (2015) have a price of 0. We restrict the analysis to the core GSOEP sample. The prices are deflated to 2011

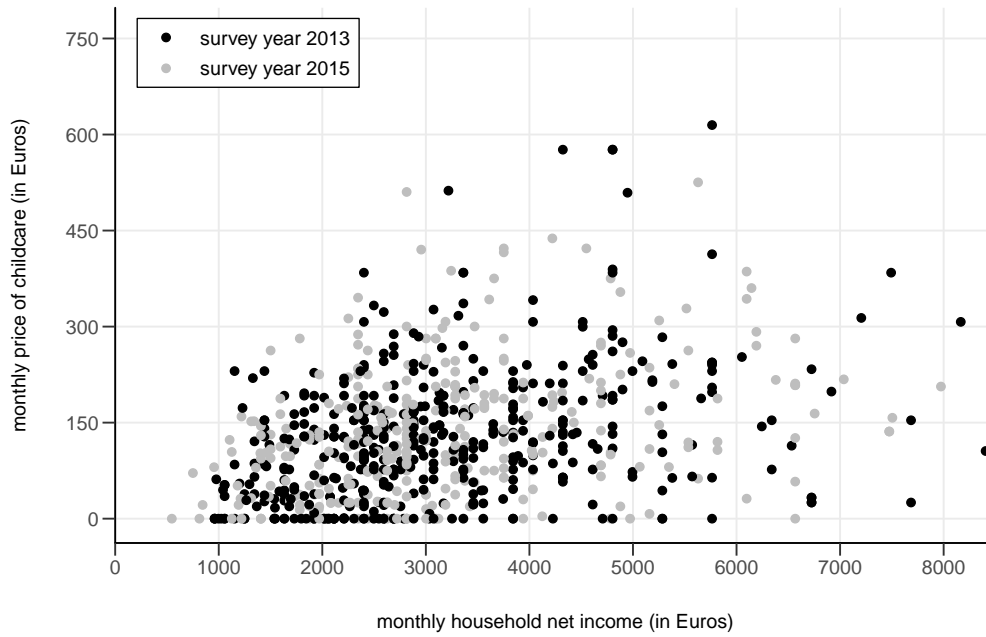


Figure 16: Unit childcare prices for children aged 3 - 5 conditional on years 2013 and 2015

Notes: The price is normalised to 40h/week by the hours spent in childcare per week where 1 time unit equals 40h/week. The survey years 2013 (454 observations) and 2015 (337 observations) are separated in this graph. Of those observations, 59 (2013) and 32 (2015) have a price of 0. We restrict the analysis to the core GSOEP sample. The prices are deflated to 2011

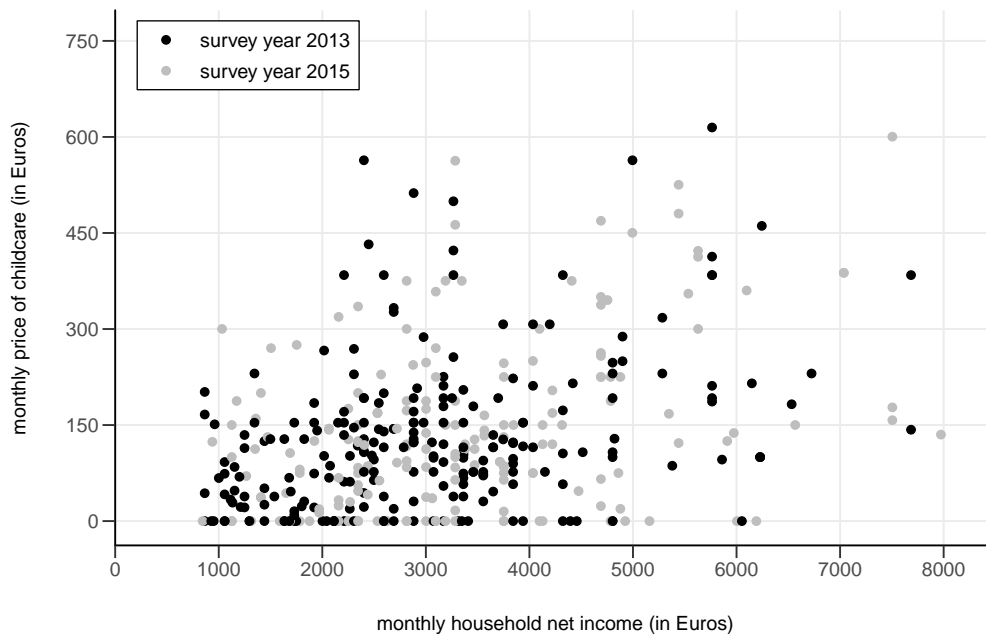


Figure 17: Unit childcare prices for children aged 6 - 8 conditional on year 2013 and 2015.

Notes: The price is normalised to 40h/week by the hours spent in childcare per week where 1 time unit equals 40h/week. The survey years 2013 (238 observations) and 2015 (186 observations) are separated in this graph. Of those observations, 38 (2013) and 33 (2015) have a price of 0. We restrict the analysis to the core GSOEP sample. The prices are deflated to 2011

C.1.4 Dispersion of unit childcare prices

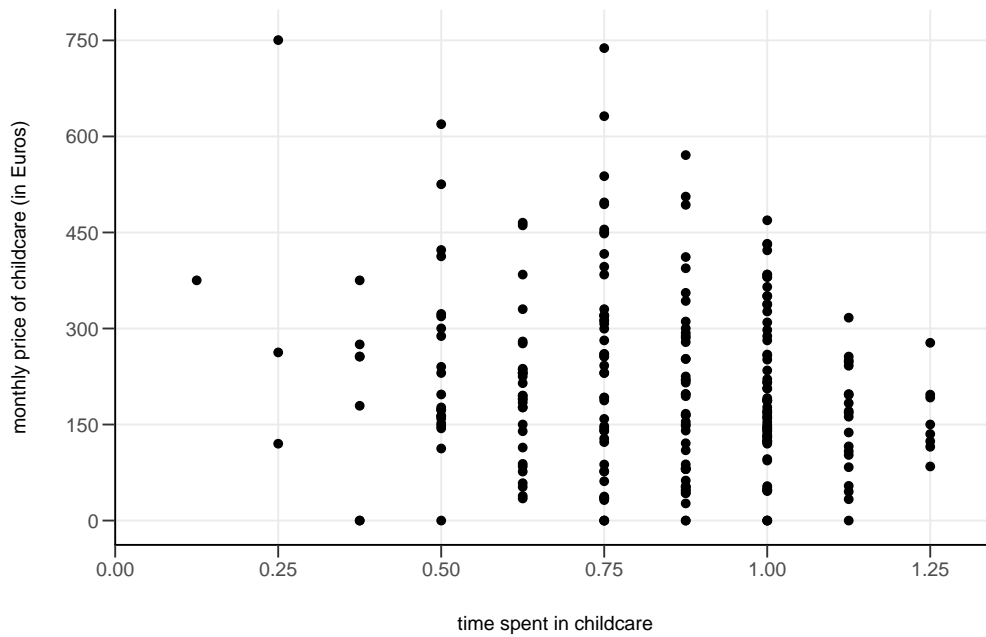


Figure 18: Dispersion of the unit price of childcare for children aged 0 - 2 by childcare hours

Notes: The price is normalised to 40h/week by the hours spent in childcare per week where 1 time unit equals 40h/week. Survey years 2013 and 2015 are pooled which leads to 253 observations of which 12 have a price of 0. We restrict the analysis to the core GSOEP sample. The prices are deflated to 2011

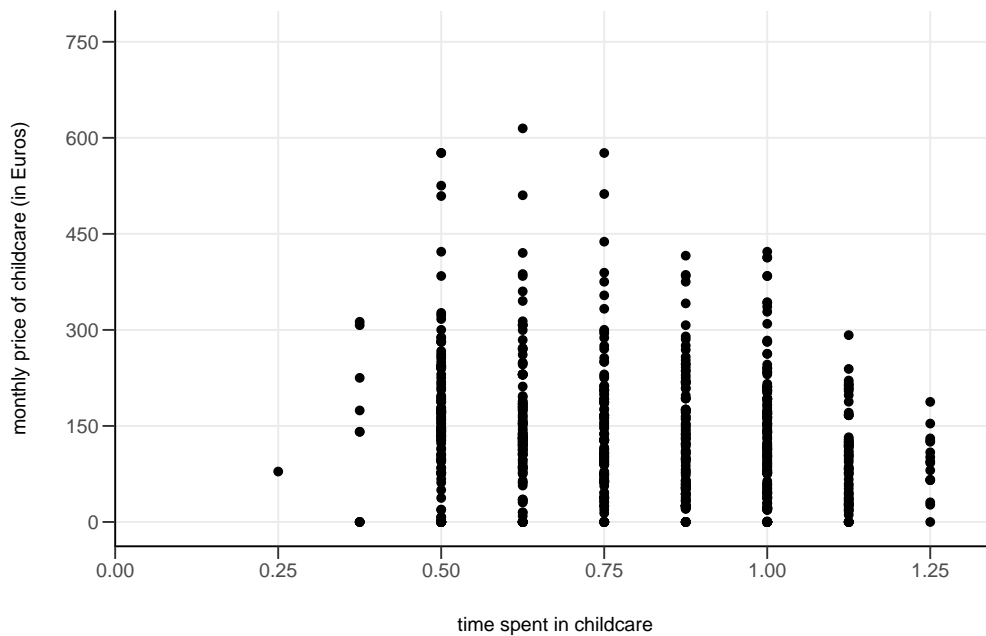


Figure 19: Dispersion of the unit price of childcare for children aged 3 - 5 by childcare hours

Notes: The price is normalised to 40h/week by the hours spent in childcare per week where 1 time unit equals 40h/week. Survey years 2013 and 2015 are pooled which leads to 791 observations of which 91 have a price of 0. We restrict the analysis to the core GSOEP sample. The prices are deflated to 2011

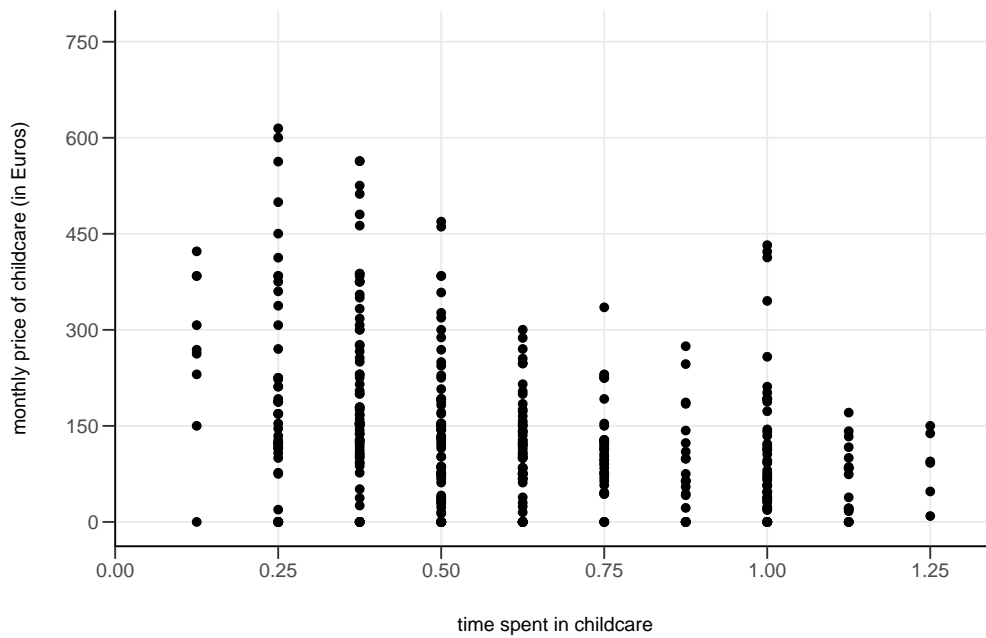
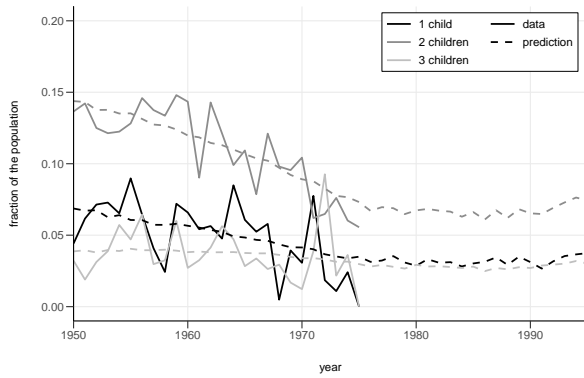


Figure 20: Dispersion of the unit price of childcare for children aged 6 - 8 by childcare hours

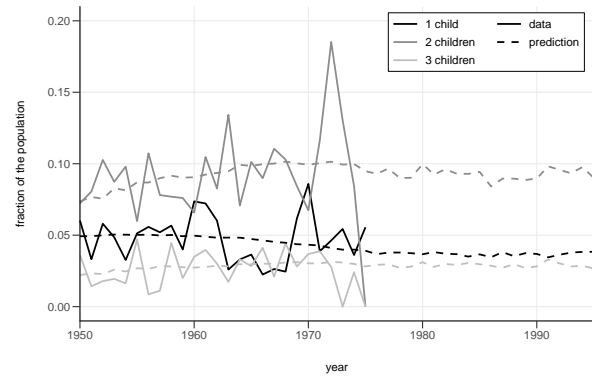
Notes: The price is normalised to 40h/week by the hours spent in childcare per week where 1 time unit equals 40h/week. Survey years 2013 and 2015 are pooled which leads to 424 observations of which 71 have a price of 0. We restrict the analysis to the core GSOEP sample. The prices are deflated to 2011

C.2 Distribution of fertility types

C.2.1 Model fit: Predicted probabilities of fertility classes for birth year



(a) Fertility class 5 - 7: Age at first birth 23 - 25



(b) Fertility class 8 - 10: Age at first birth 26 - 28

Figure 21: Average predicted probabilities and observed data shares of fertility classes across birth years

Notes: In-sample definition: women aged 41 to 65, born between 1950 and 1975 and not in education, conditional on first birth (if any) only between 20 and 40 and having completed fertility by age 41. Out-of-sample definition: women aged 20 to 40, born between 1976 and 1995 and not in education, conditional on first birth (if any) only between 20 and 40. Source: 2000 to 2016 core GSOEP samples.

C.3 Maximum Likelihood Estimation

C.3.1 Estimation of female hourly wages

Table 6: OLS estimation of female hourly wages: *Dependent variable: log w*

Child age	0 - 2
constant	2.250*** (0.076)
Full-time experience	0.0180*** (0.0042)
Full-time experience squared	-0.0002* (0.0001)
Part-time experience	0.0021 (0.0047)
Part-time experience squared	0.0002 (0.0002)
Some school degree and vocational training	0.1202* (0.0646)
A-Levels without studies	0.2858*** (0.0700)
University degree	0.4778*** (0.0671)
Child below age 6 in HH	-0.0239 (0.0469)
1 child below 17 in HH	0.0925*** (0.0346)
2 children below 17 in HH	0.1577*** (0.0393)
Observations	935
Adjusted R-squared	0.1811

Standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Notes: Sample definition: Selected sample contains 1,308 married couples in core GSOEP 2016 with husbands working full-time and woman being between 20 and 65 and not having her first child before 20 or after 41. We drop the top and bottom 1% of the male earnings and female wage distribution. All wages are deflated to 2011.