Labor Supply of Mothers: The Role of Time Discounting

Peter Haan*†, Luke Haywood†, Ulrich Schneider*†

February 15, 2016

– Preliminary –

Abstract

We estimate a dynamic life-cycle model of labor supply with a focus on time preferences for women. We extend the dynamic discrete choice model to accommodate potentially non-exponential discounting. Variation in job protection regulations provides identifying variation to test time discounting, affecting future and not current payoffs. Reforms to job protection legislation in Germany constitute a natural experiment to identify the key time preference parameters of our model. We shed light on the importance of time-inconsistent preferences on maternal labor market return. The structure of time preferences will importantly affect cost and effectiveness of labor market policies.

JEL:

Keywords: time preferences; dynamic discrete choice; labor supply; maternity leave.

Acknowledgements: We are grateful to participants at seminars at Arizona State University and DIW Berlin for their comments. This research is partly founded by the DFG, project: SPP 1764. The usual disclaimer applies.

*Free University of Berlin

†Deutsches Institut für Wirtschaftsforschung Berlin

1 Introduction

Dynamic structural models of labor supply are used to analyze individual behavior over the life cycle and to evaluate a large array of policy reforms, see e.g. Eckstein and Wolpin (1989), Keane and Wolpin (1997), Adda et al. (2016), Blundell et al. (2015). In these models decisions at any point in time are made with respect to the discounted future stream of costs and benefits accruing throughout later periods of life. Therefore, assumptions about how individuals discount these future utility streams when making decisions, i.e. assumptions about individual time preferences, are crucial not only to describe the behavior of individuals but also for policy evaluation and optimal policy design.

In dynamic structural models assumptions about time preferences are typically restrictive. In particular, while there is considerable experimental and observational evidence that individuals deviate from exponential discounting of future utility streams (for a survey, see Frederick et al., 2002), with few exceptions (Fang and Silverman, 2009; Chan, 2014) models of labor supply continue to rely on the assumption of time consistent exponential discounting. One reason for this assumption is that identification of time preference parameters without variation in expected future streams of costs and benefits is not possible (Magnac and Thesmar, 2002). In more detail as shown by Fang and Wang (2015) time preference parameters are identified in addition to other structural parameters, such as preferences for consumption, if variation of expected future streams of costs and benefits can be exploited that does not affect per-period utilities. In other words, for the identification instruments are required that affect future transitions of individuals, but which do not directly affect the flow utility (Fang and Wang, 2015).

In this paper we exploit exogenous changes in the duration of job protection for mothers to identify time preferences in a dynamic model of female labor supply with labor market frictions. Job protection provides insurance against labor market frictions for mothers after parental leave, thereby influencing females labor supply choices. Crucially for our identification strategy, job protection does not affect directly the flow utility of mothers but only future employment transitions, i.e. it guarantees employment opportunities in future periods. We do not impose exponential discounting but specify time preferences with (quasi-)hyperbolic discounting as in Laibson (1997), Fang and Silverman (2009) or Chan (2014), a specification that allows for time consistent exponential discounting as a corner solution. For the identification of the time preferences we exploit exogenous variation in the duration of job protection in Germany which was increased in several steps from one to three years (Schoenberg and

Ludsteck, 2014). We allow for a flexible wage process including endogenous human capital accumulation and differential depreciation in full-time, part-time and non employment (similar to Adda et al. (2016) and Blundell et al. (2015)). Further we include a detailed specification of the relevant components of the tax and transfer system including joint taxation, unemployment benefits, social assistance and child-care costs. We take into account assortative matching and that partner's earnings affect mothers' labor supply decisions.

The structure of time preferences is crucial to understand maternal employment behavior and to evaluate the effects of family and labor market policies. Time inconsistent choices may partially explain the long career interruptions of mothers after childbirth which cause large career costs (Adda et al., 2016) and are an important determinant of the female-male wage gap (Kleven et al., 2015; Ejrnaes and Kunze, 2013; Anderson et al., 2002). Similarly, the employment and welfare effects of family and labor market policies depend on the time preference rate of mothers. Quasi-hyperbolic discounters' behavior is especially sensitive to current costs and benefits. Thus subsidies for child care even for a short period can induce large short run employment effects for mothers with time inconsistent behavior. In contrast for time consistent mothers even large subsidies for a short period should not lead to sizable employment effects as this reform has only a minor effect on the exponentially discounted life-cycle income. In this respect this paper can make an important contribution. Based on the estimated parameters of the model, including the time preferences, we will simulate the employment responses of various policy reforms aiming to improve the labor market situation of women including the introduction of subsidies for child care costs and wage subsidies for women.

Our paper relates to several strands of the literature. First, the structure of time preferences have been the object of numerous studies, e.g. by Strotz (1956), Phelps and Pollak (1968) and Pollak (1968) with recent interest sparked by Laibson (1997) and O'Donoghue and Rabin (1999) who formalize (quasi-)hyperbolic discounting. Since then applications of (quasi-)hyperbolic discounting have become popular in experimental and observational studies including structural models, e.g. Harris and Laibson (2002) for consumption decisions, (e.g. Diamond and Köszegi, 2003; Gustman and Steinmeier, 2012) for retirement and saving decisions. Fang and Silverman (2009) and Chan (2014) model labor supply and welfare program participation for present-biased individuals and we extend their work in several dimensions. By exploiting several policy reforms which provide exogenous variation in labor market incentives we can identify our time preference parameters based on observed labour market behaviour of a large sample. This allows us to build a richer model of labor market and childcare frictions, important alternative determinants of prolonged spells of maternal non-employment.

Second, we contribute to the study of female labor supply behavior in a structural life-cycle context, see e.g. Eckstein and Wolpin (1989), Keane and Wolpin (1997), Adda et al. (2016), Blundell et al. (2015).

Finally, this study contributes to the literature which evaluates the employment and welfare effects of family and labor market policies. The empirical evidence about the employment effects is mixed (see e.g. Baker et al. (2008) or Havnes and Mogstad (2011)). Several reduced form studies for Germany find sizable positive employment effect of family policy reforms, such as the introduction of subsidies of child care Bauernschuster and Schlotter (2015), or changes in the duration of payed parental leave.

The remainder of the paper is structured as follows. Section 2 presents the dynamic labor supply model. Section 3 presents the data. Section 4 provides background information on the institutions. Section 5 discusses identification with a special focus on the parameters describing the structure of time preferences. Section 7 concludes.

2 Economic Model

We model mothers' choice to work in full- or part-time as a function of current and future discounted utility from consumption and leisure. We allow for a flexible wage process that allows for endogenous human capital accumulation and depreciation (at different rates in full-time, part-time and non employment) following Adda et al. (2016) and Blundell et al. (2015). Furthermore, we take into account incentives to work originating in the tax and transfer system including joint taxation, unemployment benefits, social assistance, child-care costs and job protection. The latter provides insurance against labor market frictions for mothers after parental leave which we model as stochastic jobs . We allow for (quasi-)hyperbolic discounting as in Laibson (1997), Fang and Silverman (2009) or Chan (2014), thereby nesting both time-consistent exponential discounting and time-inconsistent hyperbolic discounting. In the following we first present the key features of our model, before we describe the functional specifications in detail.

2.1 Overview of the model

Each period, a woman chooses her labor supply among three alternatives, not to work, to work part-time or to work full-time.[1] However, due to job search frictions, unemployed women can only choose employment if they receive a job offer (which occurs at a stochastic rate). Employed women also face unemployment risk in form of an exogenous job separation rate.

We explicitly model the development of human capital as an process of accumulation and depreciation. This human capital influences wages and the amount of job offers, in line with strong persistence of labor market outcomes. While the human capital depreciates with a constant rate, the accumulation process depends on the employment status. We allow for the possibility that full-time workers will have relatively larger human capital accumulation than part-time workers, in line with the evidence provided by Blundell et al. (2015).

This is crucial in our context, since our focus lies mostly on how women incorporate the different future consequences of their possible choices into their current decision process. Labor supply determines future human capital which again determines future career paths and thus future consumption possibilities. In the short term, the employment status influences future job offer probabilities and therefore future employment possibilities.

The household context plays a major role when considering female labor supply. Although we do not model the marriage market and fertility choices endogenously, we capture the impact of the household context. Following Haan and Prowse (2015) and Blundell et al. (2015) we estimate the probabilities of marriage, divorce and childbirth in a first step. This allows us to take into account the role of assortative mating and generate different labor market incentives for different individuals according to their socio-economic variables.

2.2 The Structural Model

In each decision period t an employed individual i (and unemployed individuals who receive a job offer) can choose their level of labor supply $l_{i,t}$ from the choice set (i) non-employment ($l_{i,t} = 0$); (ii) part-time work ($l_{i,t} = 1$); (iii) full-time work ($l_{i,t} = 2$). Full-time workers are

[1] In our model education is exogenous and determines the labor market entry age from which on we model the decision process.

assumed to work twice as many hours as part-time workers.[2] We assume the decision period to be semi-annual.

Flow Utility. The per-period utility is similar to Adda et al. (2016) and is given by

$$
\begin{aligned}
u_{i,t} = \; & \frac{(c_{i,t}/\tilde{c})^{(1-\gamma_C)} - 1}{1-\gamma_C} \times \exp(\gamma^1_{PT}\mathbb{1}_{\{l_{i,t}=1\}} + \gamma^1_U \mathbb{1}_{\{l_{i,t}=0\}}) \\
& + \left[\gamma^1_{A,NC}\mathbb{1}_{\{NC_{i,t}=1\}} + (\gamma^2_{A,NC}\mathbb{1}_{\{NC_{i,t}=2\}}) \times \exp(\gamma_{NC,H}\mathbb{1}_{\{NC_{i,t}>0,\, H_{i,t}=1\}})\right] \\
& \quad \times \exp(\boldsymbol{\gamma_{A,NW} X_{age^{YC}}}) \times \exp(\boldsymbol{\gamma_{A,PT} X_{age^{YC}}}) \\
& + \epsilon_{i,t}
\end{aligned} \tag{1}
$$

$$
\text{with} \qquad \boldsymbol{X_{age^{YC}}} = (\mathbb{1}, \mathbb{1}_{\{age^{YC}_{i,t}\in[0,3]\}}, \mathbb{1}_{\{age^{YC}_{i,t}\in[3,6]\}}, \mathbb{1}_{\{age^{YC}_{i,t}\in[6,9]\}})'
$$

where $c_{i,t}$ denotes the consumption, $\tilde{c}$ an equivalence scale[3] which accounts for the number of household members, $NC_{i,t}$ the number of children[4], $H_{i,t}$ a dummy for the presence of a partner and $age^{YC}_{i,t}$ the age of the youngest child. Furthermore $\mathbb{1}_{\{condition\}}$ is an indicator function which equals 1 if the condition is true and 0 otherwise. It is assumed that the error term $\epsilon_{i,t}$ is independently and identically over time and labor supply choices distributed with a type-1 extreme value distribution.

The first line of equation (1) represents the basic trade-off between leisure and consumption, where we assume that the preferences between the two a non-separable. The utility derived from consumption is modeled using a standard CRRA function. Utility also depends on the number of children and the presence of a partner (line 2 of equation (1)). We allow for the possibility that utility derived from the presence of children and a husband might differ with respect to the amount of leisure time a woman has and the age of the children (line 3 of equation (1)). We discriminate between four different age groups of children, pre-kindergarden age, kindergarden age, elementary school age and all other ages. This specification allows for sophisticated leisure preferences which ensures that the length of career breaks can be driven by the combination of numerous factors.

Wages and Human Capital. The labor supply decision depends also strongly on consumption opportunities, for which wages are one of the pivotal factors. We model the wage

[2] We assume 226 working days in a given year, i.e. 113 working days in a half-year. Part-time is assumed to be 4 hours a working day (452 hours a half-year), full-time is 8 hours a working day (904 hours a half-year).

[3] We assume that $\tilde{c} = 1$ for singles, 1.6 for couples, 1.4 for lone mothers and 2 for couples with children.

[4] For reasons of tractability we restrict women to have a maximum of two children.

process (similar to Blundell et al., 2015):

$$\begin{aligned}\ln(w_{i,t}) = &\ln(\gamma_{w,sec}\mathbb{1}_{\{educ=secondary\}} + \gamma_{w,high}\mathbb{1}_{\{educ=high\ school\}} + \gamma_{w,uni}\mathbb{1}_{\{educ=university\}}) \\ &+ \gamma_{w,e}\ln(e_{i,t}+1) + \xi_{i,t}\end{aligned} \tag{2}$$

The hourly wage rate depends on the individual's highest education degree and accumulated human capital. $\xi_{i,t}$ is to be assumed a measurement error which follows a normal distribution with standard deviation σ_ξ. Since the education does not change over the life-cycle in our model, wage differences over time are mostly driven by on-the-job human capital. This evolves in the following manner:

$$e_{i,t} = e_{i,t-1}(1-\eta) + \begin{cases} 0 & \text{if } l_{i,t-1} = 0 \text{ (unemployed)} \\ \lambda & \text{if } l_{i,t-1} = 1 \text{ (part-time)} \\ 0.5 & \text{if } l_{i,t-1} = 2 \text{ (full-time)} \end{cases} \tag{3}$$

Human capital at the beginning of each period depends on the previous period's human capital and the employment status. There is depreciation with rate η every period[5] which can only be offset if the individual is employed. The possible accumulation depends on the working hours, i.e there are different gains for part-time and full-time employment. Since we assume the decision period to be semi-annual, we normalize the gain of full-time employment to be 0.5. We estimate the gain for part-time employment to not restrict ourselves to a specific ratio in wage growth between the two employment states.

Budget Constraint. Given the labor supply decision and the wage process, consumption is then given by:

$$\begin{aligned}c_{i,t} = &452 w_{i,t} \times (2 \times \mathbb{1}_{\{l_{i,t}=2\}} + \mathbb{1}_{\{l_{i,t}=1\}}) + \mathbb{1}_{\{H_{i,t}=1\}} earn^H_{i,t} \\ &- TT(earn^W_{i,t}, earn^H_{i,t}, H_{i,t}, ageC1_{i,t}, ageC2_{i,t}) \\ &- cc^E \times 452 \times NC_{i,t} \times (2 \times \mathbb{1}_{\{l_{i,t}=2\}} + \mathbb{1}_{\{l_{i,t}=1\}})\end{aligned} \tag{4}$$

where $earn^H_{i,t}$ denotes the gross earnings of the potential partner, $earn^W_{i,t}$ the gross earnings of the woman, TT for the German tax and transfer system, $ageCj_{i,t}$ the age of a potential child j and cc^E the expected cost of one hour of childcare. We assume 226 working days in a given year, i.e. 113 working days in one half-year. Additionally, we define part-time employment as 4 hours a working day (452 hours a half-year) and full-time as 8 hours a

[5] At the start of the working life, every individual is assumed to have no on-the-job human capital.

working day. Therefore, the first part line 1 of equation (4) describes the half-yearly labor earnings of women depending on the labor supply and wage rate. The second part of the sum are the half-yearly earnings of the partner, conditioned on the presence of a partner.

We model all key features of the German tax and transfer system which depends on the earnings and the presence of a partner and children. For example, we explicitly take into account the joint taxation present in Germany, the deductibility of social securities as well as child allowance. It is important to note that women who do not work are eligible for unemployment benefits and social securities which partly depend on the household income. Maternity benefits depend on the number and age of the children and the policy regime they were born in.

In Germany, subsidized childcare slots are rationed, but we assume that mothers who work have to find a childcare opportunity for all hours they are working. If they do not find a subsidized slot, they need to investigate private options which are often more costly. We approximate this process by modeling expected childcare costs similar to Wrohlich (2011):

$$cc^E = cc^S\pi + cc^{NS}(1-\pi) \tag{5}$$

where cc^E denotes the expected, cc^S the average subsidized, and cc^{NS} the average non-subsidized childcare costs per hour. π denotes the probability of being able to use subsidized childcare.

Labor Market Frictions. One reason for long non-employment spells can lie in the lack of employment opportunities. Provided an individual was not employed in the previous period, she receives an job offer with a probability Θ:[6]

$$\begin{aligned}\pi_{i,t}^{JO} = \Phi(&\gamma_{JO,sec}\mathbb{1}_{\{educ=secondary\}} + \gamma_{JO,high}\mathbb{1}_{\{educ=high\,school\}} \\ &+ \gamma_{JO,uni}\mathbb{1}_{\{educ=university\}} + \gamma_{JO,e}e_{i,t})\end{aligned} \tag{6}$$

Therefore, the job offer probability depends on education and the current amount of human capital. Furthermore, after child-birth, mothers' benefit from job protection which we model as a job offer probability of one. This allows mothers to return to employment and freely choose her hours at any time within the job protection period.

[6]The symbol Φ henceforth denotes the cumulative standard normal distribution.

In addition, we introduce a probability for involuntary job separations, $\pi_{i,t}^{JL}$, conditioned on being employed in the previous period. An individual who gets involuntary laid-off cannot work in the current period. In contrast, individuals who do not experience this job loss, can remain in employment if they desire to do so.

Dynamics of Family Composition. All family dynamics are modeled as an exogenous stochastic process. The probability with which a woman receives a child depends on her age, the presence of a partner, and the presence of other children and their respective age. We do not allow for more than two children. It is assumed that all children live with their mother until the age of 18.

The formation and destruction of partnerships are modeled in a similar manner. Both probabilities are estimated separately and depend on the age and education of the woman. The probability of a separation also depends on the presence and age of children. In line with assortative mating, we model the earnings of partners as a function of womens characteristics. Note that this implies women take into account the likelihood and material implications of meeting a partner when choosing their labor supply.

2.3 Intertemporal Optimization

In any given period t, an individual is assumed to maximize her expected life-time utility U_t:

$$\max_{\{l_t, l_{t+1}, \ldots, l_T\}} U_t(l_t, l_{t+1}, \ldots, l_T, \Omega_t) = u(l_t, \Omega_t) + \beta \mathbb{E}\left[\sum_{\tau=t+1}^{T} \delta^{\tau-t} u(l_\tau, \Omega_\tau) \middle| \Omega_t \right] \tag{7}$$

where we drop the index i for ease of notation. δ denotes the standard discount factor, β the present-bias factor (O'Donoghue and Rabin, 1999), $\mathbb{E}$ the expectations operator and Ω_t the state space at time t:

$$\Omega_t = \{age_t, e_t, NC_t, ageC1_t, ageC2_t, H_t, jp_t, \epsilon_t, l_{t-1}, jp_{t-1}\}.$$

The binary variable jp_t equals one when the women has job protection in period t.

Additionally to the standard discounting factor δ, the individual discounts all expected future values again by β. If $\beta = 1$, the standard exponential discounting framework applies, but if $\beta < 1$ the individual displays a present bias, indicating the impulse for immediate gratification.

The time-preferences we specify are also known as (β, δ)-preferences (O'Donoghue and Rabin, 1999). One typical characteristic of this specification is that the individual always discounts exponentially between future time-periods. For instance, from the perspective of period t, the individual discounts the utility of the first period $t+1$ by $1-\beta\delta$, while she discounts utility between any other two subsequent periods by $1-\delta$. The time-inconsistency in behavior might arises once the women progresses in time, for example to period $t+1$. While she discounted between period $t+1$ and $t+2$ with $1-\delta$ before, she now uses the discount factor $1-\beta\delta$. For instances, this can lead to women planning on returning to employment two years after child-birth, but delaying return once her child reaches age two.

These preference reversals generate inconsistencies which individuals may foresee. If individuals are aware of their inconsistencies and adapt their behavior accordingly, agents are called *sophisticated*. In absence of a commitment device, this requires sophisticated calculations, in particular. In contrast, individuals who are absolutely not aware of their time-inconsistencies are called *naïve*. In this paper, we assume that agents are fully naïve.

2.4 Solution of the Structural Model

To build the foundations for our identification strategy it is beneficial to understand how the model, described in the previous section, can be solved. With use of the assumption that agents are fully naïve, we can rewrite the long-run utility, i.e. the utility of exclusively future periods in a recursive manner (see Fang and Silverman, 2009; Chan, 2014):

$$\begin{aligned} & V_t(\Omega_t) = \max_{l_t \in \{0,1,2\}} u(l_t, \Omega_t) + \delta \mathbb{E}(V_{t+1}(\Omega_{t+1}) | \, \Omega_t) \qquad \text{for } t \neq T \\ \text{and} \quad & V_t(\Omega_t) = \max_{l_t \in \{0,1,2\}} u(l_t, \Omega_t) \qquad \text{for } t = T \end{aligned} \tag{8}$$

Note that because we are only looking at future periods, β is not included in (8). To simplify notation, we denote the term $\mathbb{E}(V_{t+1}(\Omega_{t+1}) | \, \Omega_t)$ henceforth with $\mathbb{E}\max_t$. We use the subscript t and not $t+1$ to emphasize that we are interested in the expected maximum from the perspective of period t. If we refer to a specific future realization of the state space in $t+1$, i.e. not the expected development of the state space, we will denote this by $\mathbb{E}\max_t(\widetilde{\Omega}_{t+1})$. The assumption of a finite horizon allows to solve the model by backwards induction.

If a woman loses her job or receives no job offer she has to remain out of employment for the current period. Taking into account that the preference shock is type-I extreme value distributed, her $\mathbb{E}\max_t$ is then given by:

$$\begin{aligned} &\mathbb{E}\max_t^{\text{non-emp}}(\widetilde{\Omega}_{t+1}) = \gamma + u(l_{t+1} = 0, \widetilde{\Omega}_{t+1}) + \delta \mathbb{E}\max_{t+1}(\widetilde{\Omega}_{t+1}) \quad \text{for } t < T-1 \\ \text{and} \quad &\mathbb{E}\max_t^{\text{non-emp}}(\widetilde{\Omega}_{t+1}) = \gamma + u(l_{t+1} = 0, \widetilde{\Omega}_{t+1}) \quad \text{for } t = T-1 \end{aligned} \tag{9}$$

where γ refers to the Euler-Mascheroni constant. Similarly, if the individual does not lose her job or receives a job offer, she has the possibility to choose among all three options. The $\mathbb{E}\max_t$ is then defined by:

$$\begin{aligned} \mathbb{E}\max_t^{\text{emp}}(\widetilde{\Omega}_{t+1}) =& \gamma + \log\left[\sum_{l_{t+1}=0}^{2} \exp\Big(u(l_{t+1}, \widetilde{\Omega}_{t+1})\Big)\right] \\ &+ \delta \mathbb{E}\max_{t+1}(\widetilde{\Omega}_{t+1}) \text{ for } t < T-1 \\ \text{and} \quad \mathbb{E}\max_t^{\text{emp}}(\widetilde{\Omega}_{t+1}) =& \gamma + \log\left[\sum_{l_{t+1}=0}^{2} \exp\Big(u(l_{t+1}, \widetilde{\Omega}_{t+1})\Big)\right] \text{ for } t = T-1 \end{aligned} \tag{10}$$

Building on equations (9) and (10) and the transition probabilities of the state space $\Pr\left(\widetilde{\Omega}_{t+1}\middle|\,\Omega_t\right)$, we can derive the final formula for the $\mathbb{E}\max$:

$$\begin{aligned} \mathbb{E}\max_t \mid (l_t = 0, \Omega_t) =& \sum_{\widetilde{\Omega}_{t+1}} \Pr\left(\widetilde{\Omega}_{t+1}\middle|\,\Omega_t\right) \Big[\pi_{i,t}^{JO} \times \mathbb{E}\max_t^{\text{emp}}(\widetilde{\Omega}_{i,t+1}) \\ &+ (1 - \pi_{i,t}^{JO}) \times \mathbb{E}\max_t^{\text{non-emp}}(\widetilde{\Omega}_{i,t+1})\Big] \\ \mathbb{E}\max_t \mid (l_t \in \{1,2\}, \Omega_t) =& \sum_{\widetilde{\Omega}_{t+1}} \Pr\left(\widetilde{\Omega}_{t+1}\middle|\,\Omega_t\right) \Big[(1 - \pi_{i,t}^{JL}) \times \mathbb{E}\max_t^{\text{emp}}(\widetilde{\Omega}_{i,t+1}) \\ &+ \pi_{i,t}^{JL} \times \mathbb{E}\max_t^{\text{non-emp}}(\widetilde{\Omega}_{i,t+1})\Big] \end{aligned} \tag{11}$$

With equation (11) we can rewrite equation (7) as

$$\max_{\{l_t, l_{t+1}, \ldots, l_T\}} U_t(l_t, l_{t+1}, \ldots, l_T, \Omega_t) = u(l_t, \Omega_t) + \beta\delta \mathbb{E}\max_t \tag{12}$$

For our identification strategy it is worth to point out that the job offer probability does not affect the flow utilities, although it is part of the state space. It only affects future employment possibilities and therefore exclusively the $\mathbb{E}\max$ in equation (12).

3 Data and Descriptive Evidence

3.1 Data and Sample

For the estimation of our proposed model, we use longitudinal data from the German Socio-Economic Panel (SOEP) covering 1986-2006 (see Wagner et al., 2007, for a description of the SOEP).[7] While the SOEP interviews individuals on a yearly basis, it asks participants to fill out a monthly calendar of the previous year. Individuals are especially asked about their complete last year's employment history. This allows us to construct a semi-annual data set by combining the current year questionnaire with information from the questionnaire of the following year.

We restrict our sample to West German women between the age of 18 and 60.[8] We exclude women who ever worked as civil servants or were self-employed.The final data set is therefore an unbalanced panel in which individuals enter and leave the panel at various points in time. We observe over $6,200$ women, on average over five and a half years. Additionally, we observe $1,375$ births and a total of $3,861$ children in the age between 0 and 18. In total we have $419,855$ semi-annual observations.

The labor market experience for a given year is constructed by combining the answers of a working history questionnaire and the recorded employment status of follow up interviews. Wages are defined as gross monthly earnings divided by actual working hours during the same period. We express all nominal variables in year 2000 prices using the Consumer Price Index. [9]

4 Institutions

In this analysis we focus on Germany for two reasons. First, Germany has a very generous job protection system which allow mothers to return to their pre-birth job within 36 months. This generates large variation in the date of return. Second, during our observation several policy reforms changed the period of maternal job protection. In this paper we exploit this

[7]We use one additional wave more than we cover years. This is due to to the fact that some collected variables are looking back on the past year.

[8]For estimations for some exogenous processes, we include women until the age of 70 to have more robust estimates for the later years, see below.

[9]Organization for Economic Co-operation and Development, Consumer Price Index of All Items in Germany [DEUCPIALLMINMEI], retrieved from FRED, Federal Reserve Bank of St. Louis https://research.stlouisfed.org/fred2/series/DEUCPIALLMINMEI, February 3, 2016.

policy variation to identify the discount parameters of our model. We concentrate on six major expansions of maternity leave coverage between 1986 and 1993.[10] The objective of these reforms was twofold. First, they were intended to encourage mothers to spend more time with their children during their early development. Second, they sought to strengthen mothers' labor market attachment, since a longer job protection period was seen to ease the return to the labor market after maternity leave.

Since the late-1960s, mothers were entitled to have 14 weeks of paid leave around childbirth. In general, the time period was divided into six weeks before the birth date and eight weeks after and women were not allowed to work, especially for the time after childbirth. During this period, employees could not been dismissed and were guaranteed a comparable job to their previous held one, when returning within the limit. During the 14 weeks, women received the average income of the three month before entering maternity leave, i.e. they had an income replacement rate of 100%. The core of this law is still in place today. In the late-1970s, a first major reform was introduced that increased maternity leave coverage. The job protection period was extended to six months after childbirth, while a new maternity leave payment for the time between the 8$^{\text{th}}$ week and the end of the 6$^{\text{th}}$ month was introduced. In this period, women received DM 750 per month. It is important to note that these maternity benefits were only paid to women who were employed before childbirth.

The reforms we exploit started in January 1986. An overview of these can be found in table 1. The first reform expanded the job protection and maternity benefit period from six months to 10 months at the beginning of 1986 and then further to 12 months in January 1988.[11] Maternity payments from week six to eight remained at an income replacement of 100% or DM 600[12] if the mother was unemployed before. Between month three to six maternity benefits declined from DM 750 to DM 600[12] per month. From the seventh month to the 10$^{\text{th}}$ month (and later 12$^{\text{th}}$ month), the amount of maternity benefits was means tested and depended on the family income of the two years prior to childbirth. Around 84% of individuals were eligible for the full amount of the benefits (Schoenberg and Ludsteck, 2014).

A further increase in the job protection and maximum maternity benefit time period from

[10]The summary of the parental leave reforms are mainly based on Zmarzlik et al. (1999) and Bundeserziehungsgeldgesetz [BErzGG] [Federal Child-Raising Benefit Act], Dec. 6, 1985, BGBl.I at 2154 (F.R.G.) and its changes until its abolition in 2007.

[11]Additionally, parental leave for fathers was introduced. However, on average only around 1% of fathers took parental leave between 1987 and 1994 (Vaskovics and Rost, 1999).

[12]This is equivalent to $ 585 in 2016.

Table 1: Parental Leave Reforms from 1986 until 2006

Policy Regime	**Month, Year**	**Job Protection**	**Maternity Payment**
Regime I	January, 1986	10 months	3-6 month DM 600[12], 7-10 month means tested
	January, 1988	12 months	up to 12 months
Regime II	July, 1989	15 months	up to 15 months
	July, 1990	18 months	up to 18 months
Regime III	January, 1992	36 months	
	January, 1993	36 months	up to 24 months
	December, 2006		

month 12 to month 15 took place in July 1989, and another rise to 18 month in July 1990. In January 1992, the job protection period was further extended, to a total of three years. In contrast, the maximum maternity payment period stayed constant at 18 months until it was extended to two years in January 1993.[13] Minor changes in family policy were introduced in 2001, but the core regime of 1993 still continued.

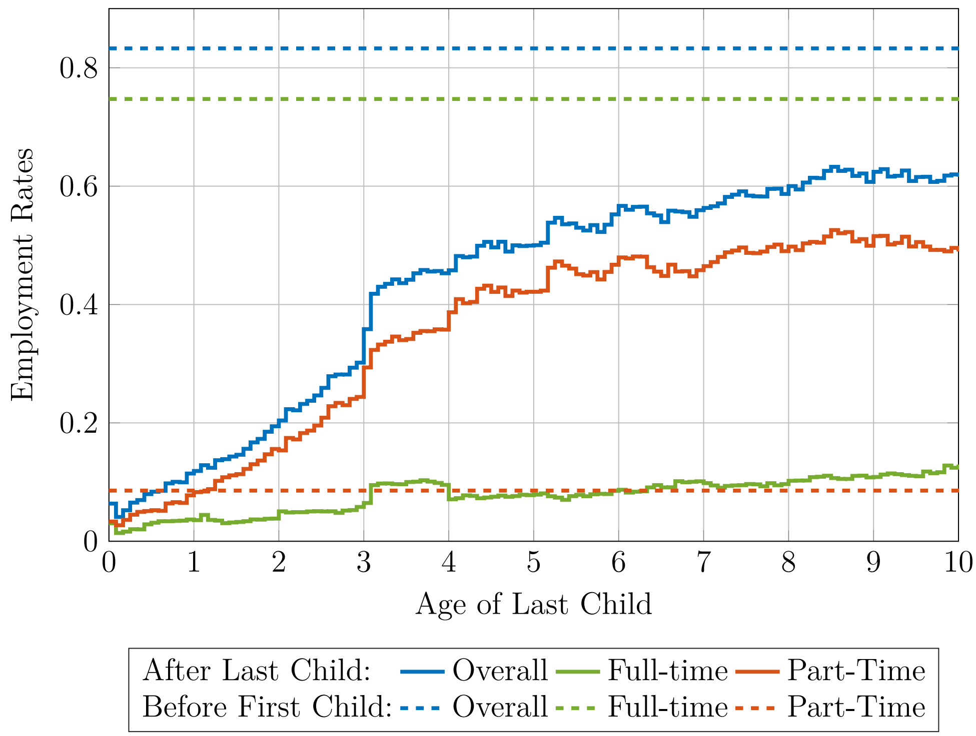
Table 1 categorizes these reforms into three periods, labelled Regimes I-III. First, tracking every policy change would not be computationally feasible: Each policy reform adds new circumstances and therefore increases the size of our state space. Second, since we allow mothers to revise their labor market choices only every 6 months, we cannot take into account changes in job protection from 10-12 or 15-18 months. Therefore, we approximate the duration of job protection to be one year for regime I, one and a half years for regime II and three years for regime III. Similarly, we assume the maternity benefits to be paid for one year for children born between January 1986 and July 1989, one and a half years for children born between July 1989 and January 1992, and two years for children born after January 1992, but before January 2007. These different regimes with different time spans, especially for the job protection periods, will help us identify the parameter of our structural model as we later explain in more detail.

[13]There was a minor change in the maternity benefits in 1994. For the first six months benefits were also means tested. For married couples the threshold was DM $100,000$, for singles DM $75,000$ for getting the full benefits in the first six months.

4.1 Descriptive Evidence

Given our selected sample, we pool all observations to get a first glimpse of how the employment status of women is affected by childbirth. Exactly one year before future mothers give birth to their first child, the majority, 74.7%, are working full-time, while 8.6% work part-time. Overall, the employment rate is 83.3%. This changes quite dramatically with the arrival of children, even in the long run. Figure 1 gives a detailed picture of the development of employment rates after the last child is born. We concentrate on the last child here, because we want to eliminate the possibility of a longer career break due to another child.

Figure 1: Emplyoment Rates After Last Child



Notes: Pooled sample of women who gave birth to their last child between 1986 and 2006. We expanded the sample period to 2012 to track also the long term employment status of women who gave birth to their last child shortly before 2006. *Before First Child* denotes the employment rates for future mothers one year before giving birth to their first child. Observations are weighted with SOEP sampling weights.

Note that we used a monthly version of our data set for this graph. We plot the above mentioned employment rates prior to the first child in dashed lines as an comparison. Three things stand out from this graph. First, employment rates rise only slowly over time, but never reach the pre-birth levels. Ten years after mothers received their very last child, only

61.9% are employed compared to over eighty percent pre-birth. Second, the difference in part-time employment before and after is especially striking. The majority of mothers, around fifty percent, do work part-time, ten years after their last child was born. Just about 12.6% are working in a full-time job at this time. Third, since the majority of our sample is observed in policy regime III, there is a strong jump in employment rates around the time the last child reached age 3 and hence job protection for most mothers ends. A non-negligible part of women return to employment just in-time to avoid risking to loose their previous held job.

Looking more closely at the individual career break lengths, we find that of all mothers who eventually will start working again, 50% start working after their last child reached the age of 2.9 years, 75% after their last child reached the age of 4.0 and 90% after their last child is 7.3. However, it seems that there is a noticeable gap between the preferred non-employment spell length and the realized one as shown in table 2.

Table 2: Return to the Labor Market after Last Child

	1st Year		2nd Year		3rd Year	
Time Period	Prefer.	Real.	Prefer.	Real.	Prefer.	Real.
Before Next Year	23.3%	12.5%	38.9%	20.2%	48.6%	40.3%
In the next 2 to 5 years	62.3%	62.0%	47.1%	51.7%	35.4%	19.5%
In more than 5 years	14.4%	25.5%	14.0%	28.1%	15.9%	40.2%
Observations	215		196		120	

Notes: Only women who are observed from the birth of their last child on until they enter the labor market and who have a job guarantee for three years, and stated that they want to be employed in the future are included in this table. Preferred length of career breaks are recorded in the year after the birth of the last child for category "1st Year", in the second year for category "2nd Year" and in the third year for category "3rd Year". The actual length is then computed from the month on the mother stated her preference to the month, the mother entered employment. Additionally, we included women whose non-employment spell was right censored if they stated their preference and did not work within the first five years. You can find the exact questions in appendix A. Frequencies are weighted with sampling weights.

In this table, we exclusively consider women who have a job guarantee for three years after the birth of their last child. This allows us to ignore labor demand factors when we interpret differences between stated preferences and realizations of career breaks. It is also important to note that if an employee took less than the maximum time of the job protection, but decides to need more time, she has the possibility to extend her protection period. She can expand it even without needing the consent of the employer on the basis of an important ground,[14] in other cases an extension is always possible with the consent of the employer.

The "1st Year" column of table 2 presents the preferred return to the labor market stated at the first interview after the child is born. Almost a quarter of mothers do not want to interrupt their career for more than a year, 62.3% plan to be employed again in medium term, i.e. between two and five years. The rest considers to be back into employment in five years or later. Tracking the career breaks of all these women reveals a strong shift towards longer career breaks than initially stated in the first year. The fraction of mothers who return to the labor market within one year is only around half the fraction who wanted return within that time span. Additionally, the ratio of realized career breaks which lasts five years or longer (25.5%) is ten percentage points higher than the previously stated ones. This trend continues in the second and the third year.

After their last child has reached the age of one, 38.9% of women state that they want to return to the labor market within the next year, but we only observe a fraction of 20.2% of women who realize it. The percentage of women who stay longer out of employment than five years is with 28.1% twice as high as the fraction of women who planned to be out for so long. In the third year, the fractions for women who stated to be back in employment within a year and the fraction who actually realizes this time frame are closer together than in the two years before. This is most likely due to the fact that the stakes mothers can loose are much higher in the third year, since if mothers do not return within a year, the job guarantee ends. It seems that from those mothers who do not return within the job protection period, most stay out for a long time period before eventually returning to the labor market.

Although we do not use the numbers reported in table 2 in our structural model, but they can be used to provide a first glimpse that the assumptions standard dynamic models make, probably do not hold in this context. Especially for models who make the assumptions of rational expectations and exponential discounting, it is hard to explain the observed gap between preferences and realizations. If mothers had rational expectations they would on average predict their behavior correctly. Time-inconsistent behavior can account for this

[14] Such a cause can for example be that her husband got sick and therefore cannot take care of the children.

pattern in a very natural way.

We can observe time-inconsistent behavior in many different situations in reality. For example, individuals who buy a yearly gym pass, but then never go to the gym (Della Vigna and Malmendier, 2006). The same argument could be made about the mothers in our sample. They might plan to get back to their old job within a year after the birth of their last child, but when the first year comes to an end and it is time to return they decide to stay out of employment just a bit longer. We also observe some consistency over time in this behavior, since the gap between stated preferences and realizations holds over time. This indicates that the numbers of table 2 are not driven by a single shock in the first period, but are more likely an outcome of a consistent deviation from standard assumptions.

We do not claim that the descriptive evidence presented in this section is already a proof for time-inconsistent behavior, since various, mostly competing, theories could explain the presented descriptives. However, our identification strategy lied out in section 5 does not depend on the numbers presented here, but on exogenous variations in transition probabilities of individuals which allows us to test the hypothesis of time-inconsistent behavior.

5 Identification and Estimation

5.1 Identification

5.1.1 Intuition

Early work by Rust (1994) argues that dynamic discrete choice models a generally underidentified and that it is typically not possible to identify discount parameters. Magnac and Thesmar (2002) extend this work by deriving specific conditions which can lead to an identification of the exponential discounting parameter. Building on these conditions, Fang and Wang (2015) develop specific exclusive restrictions which allows researchers to identify parameters of a quasi-hyperbolic discounting model.

Their idea is to find variables that have no influence on flow utilities, but on the transition probabilities of at least one state variable. Consider two individuals who only differ in the values of such a variable. Although these two do not differ in their flow utilities, they do in their expected future utility streams, since the state space develops differently. Observed differences in choice probabilities are then only due to their expected futures, and therefore can inform the researcher how individuals weight these expected utility streams.

Our framework provides us with a natural instrument that fulfills the requirements described above. Women, who were employed before giving birth to a child, were protected from dismissal for a certain period of time. This job protection is reflected by a job offer probability of one in our model. This offer probability and therefore also the job protection does not enter directly into the flow utilities of our individuals. It exclusively affects expected future values. To recover the time-preference parameters, we can compare the choice probabilities of mothers with different job protection horizons. Hence, we exploit the various extensions in the length of job protection for identification of the time-preference parameters.

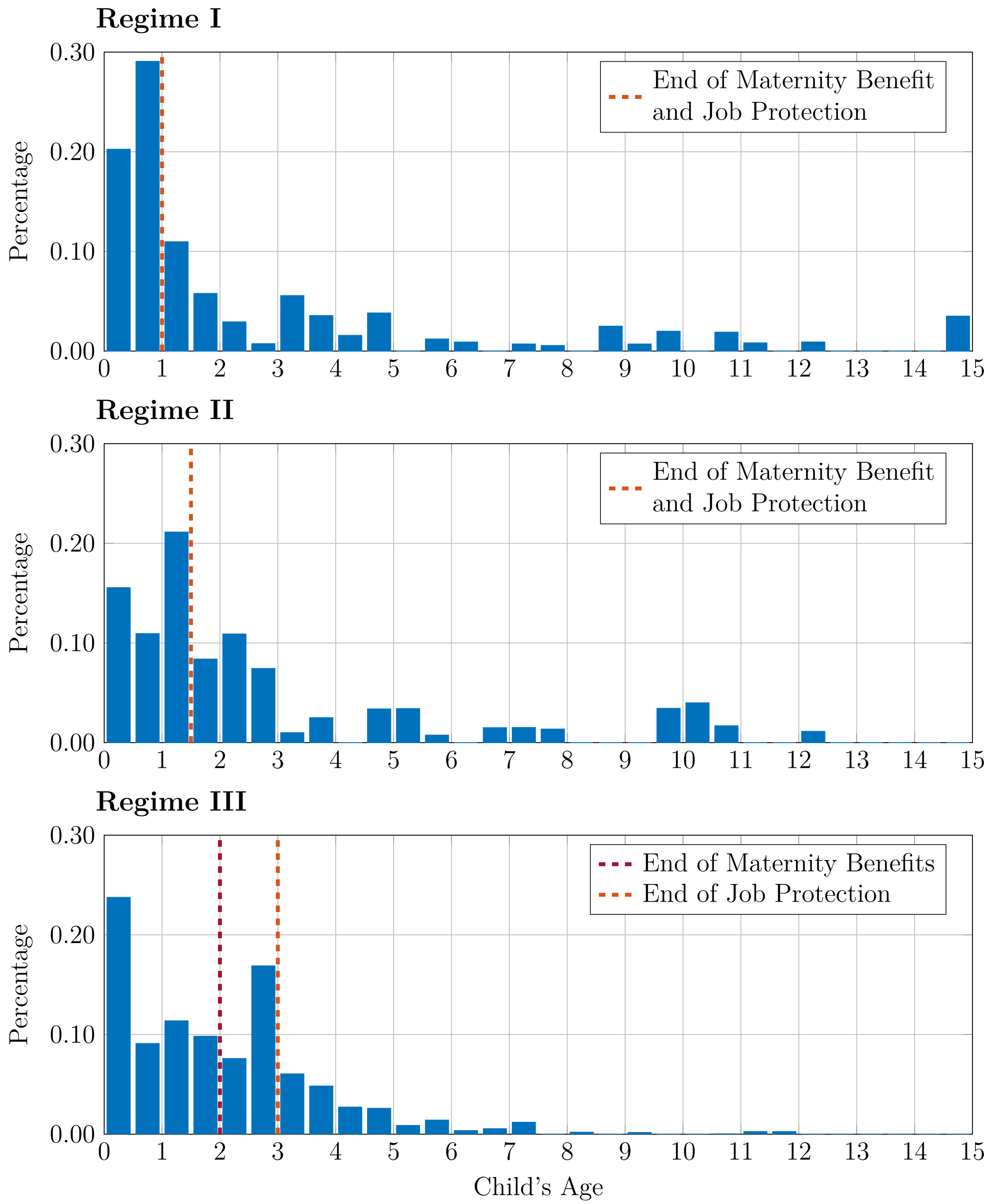
As argued by Schoenberg and Ludsteck (2014), these changes in the duration of the job protection for mothers are exogenous, which allows the authors to evaluate the causal employment effects of these reforms in a reduced form setting. Besides this exogeneity, the relevance of our instrument is important. This means that, under the assumption that our agents are not fully myopic, we should see different choice behavior of women in different job protection regimes. A first indication of the different behavior can be found in figure 2.

The figure shows the distribution of career breaks of mothers for different length of job protection. Most mothers return just in time to not lose their guarantee to be able to return to their former employer. That the behavior is strongly affected by the length of job protection becomes apparent when we compare the fraction of mothers who return to the labor market in the second half of the first year. While in regime I almost 30% of the mothers return at this time, less than 10% return in regime III. Note also that while these reforms might significantly influenced incentives to work for mothers, the group of young mothers constitutes only a very small fraction of the overall workforce, warranting our focus on changes in labor supply.

5.1.2 Formal Discussion

For simplicity, we restrict ourselves in the following example to three time periods and two types of individuals, A and B. The types are identical expect for the duration of their job protection. Type A has job protection for all three periods, while type B only for the first two periods. We will denote job protection in period t by jb_t, no job protection as $\overline{jb}_t$ and the state space without the current job protection status as $\breve{\Omega}_t = \Omega_t \setminus jp_t$. The life-time utility

Figure 2: Length of Career Breaks



Notes: Histogram of the length of career breaks after any child born in the respective policy regime. The length of a career break is defined as the time between the birth of a child and the time the mother starts working or receives another child. Hence, only mothers are included for whom we observe the employment status from the birth of a child until they are employed or receive another child. Observations are weighted with SOEP sampling weights.

in the second to last period is then specified as

$$\begin{aligned} U_2^A =& u(jp_2, \breve{\Omega}_2) + \beta\delta\mathbb{E}\left[\max u(jp_3, \breve{\Omega}_3)\middle| jp_2, \breve{\Omega}_2\right] \\ U_2^B =& u(jp_2, \breve{\Omega}_2) + \beta\delta\mathbb{E}\left[\max u(\overline{jp}_3, \breve{\Omega}_3)\middle| jp_2, \breve{\Omega}_2\right]. \end{aligned} \tag{13}$$

We have three main components that need to be identified in equation (13): the discount parameters β and δ, the flow utilities $u(\cdot)$ and the life-time utilities U_2. In general, under the regular conditions of estimating a dynamic discrete choice model, no additional problems should occur when identifying $u(\cdot)$ (Fang and Wang, 2015). For instances, it is possible to think of the identification of the flow utilities by using a two-step estimation approach which in a first step uses only observations from one type, say A. This first step would then correspond to a standard estimation of dynamic discrete choice model and should recover the parameters of the utility function. The discounting parameters could then be recovered in a second step.[15] Note that with the identification of $u(\cdot)$, the expected maximum is also identified.

While we cannot identify U_2 directly, the differences between the two types can be identified by comparing the choice probabilities of the two types (Hotz and Miller, 1993). We can transform (13) to:

$$\beta\delta = \frac{U_2^A - U_2^B}{\mathbb{E}\left[\max u(jp_3, \breve{\Omega}_3)\middle| jp_2, \breve{\Omega}_2\right] - \mathbb{E}\left[\max u(\overline{jp}_3, \breve{\Omega}_3)\middle| jp_2, \breve{\Omega}_2\right]} \tag{14}$$

which gets all elements that can be directly identified to the right hand side. Since the duration of job protection changes the possibility of receiving a job offer and therefore changes the expected utility, it is plausible that

$$\mathbb{E}\left[\max u(jp_3, \breve{\Omega}_3)\middle| jp_2, \breve{\Omega}_2\right] - \mathbb{E}\left[\max u(\overline{jp}_3, \breve{\Omega}_3)\middle| jp_2, \breve{\Omega}_2\right] \neq 0.^{16} \tag{15}$$

Therefore allows us to identify the product of $\beta\delta$. Equation (14) also provides the intuition for identifying totally myopic individuals. Conditioned on the differences between the two types with respect to their expected future, equal choice probabilities indicate that individuals do not account for their future when making decisions and therefore at least one of the discounting parameters has to be equal zero.

[15] We will not use this described two-step procedure in our estimation. It should only illustrate a possibility one way to think about identification.

[16] Indeed, only when the utility of not working, working part-time and working full-time are equal or there are no labor market frictions, equation (15) equals zero.

Assuming that our individuals are at least to some degree forward looking, i.e. not myopic, we can separately identify δ by investigating the life-time utilities of the first period

$$\begin{aligned} U_1^A =& u(jp_1, \breve{\Omega}_1) + \beta\delta\mathbb{E}\left[\max u(jp_2, \breve{\Omega}_2)\Big| jp_1, \breve{\Omega}_1\right] + \beta\delta^2\mathbb{E}\left[\max u(jp_3, \breve{\Omega}_3)\Big| jp_1, \breve{\Omega}_1\right] \\ U_1^B =& u(jp_1, \breve{\Omega}_1) + \beta\delta\mathbb{E}\left[\max u(jp_2, \breve{\Omega}_2)\Big| jp_1, \breve{\Omega}_1\right] + \beta\delta^2\mathbb{E}\left[\max u(\overline{jp}_3, \breve{\Omega}_3)\Big| jp_1, \breve{\Omega}_1\right] \end{aligned} \tag{16}$$

As before, we can arrange the difference between the two equations to have all identified components on the right hand side:

$$\delta = \frac{U_1^A - U_1^B}{\beta\delta\left(\mathbb{E}\left[\max u(jp_3, \breve{\Omega}_3)\Big| jp_1, \breve{\Omega}_1\right] - \mathbb{E}\left[\max u(\overline{jp}_3, \breve{\Omega}_3)\Big| jp_1, \breve{\Omega}_1\right]\right)} \tag{17}$$

The same arguments as above can be made about the identification of the differences in life-time utilities and the expected future utilities. Additionally, as shown in equation (14) we have identified the product $\beta\delta$. Therefore, (17) identifies δ, and thus indirectly also β.

It is straightforward to extend this short example to multiple periods. The crucial element for identification is that the difference of the equations in (13) and the difference of the equations (16) can be transformed into two equations with two unknowns independently of the number of periods.

5.2 Estimation

We follow a two-step procedure to estimate the parameters of our model. In a first stage, we estimate the parameters of all exogenous processes, including the partnership development, the arrival of children, childcare costs and the job destruction rate.[17] Details and estimation results can be found in Appendix D.

The time-preference parameters, the parameters of the flow utility function, of the wage process and the job offer probability are estimated in a second-step via general indirect inference (Bruins et al., 2015).[18] This second-step is implemented as an iterative procedure which

[17]The SOEP data allows us to explore the reasons why an individual lost her job. From this we are able to construct a probability of involuntary job loss.

[18]General Indirect Inference is primarily a smoothed version of the indirect inference estimator, especially for dynamic discrete choice models. For Indirect Inference see Smith (1990), Gourieroux et al. (1993) and Gallant and Tauchen (1996), for an application of general indirect inference see for example Altonji et al. (2013).

involves numerous steps. First, for a given set of parameters, we solve the model as described in section 2.4, we then simulate the life-cycles of $31,020$ women which corresponds to five times the number of women we observe in our data. For each simulated individual, we only keep the observations of the time periods for which we also observe the respective women in our data set. We account for missing wages by only recording wages when simulated individuals are in employment and the SOEP interview was conducted in the respective period.[19]

In a standard Indirect Inference estimation, we would then compute moments of the simulated and the observed data and minimize the distance between them following an objective function. However, since we are estimating a dynamic discrete choice model, the objective function would be a step-function. Small changes in a parameter of our model will result in changes in discrete outcomes which leads to discrete changes of the objective function. It is therefore not possible to use gradient-based optimization algorithms. Additionally, since we are aiming to estimate over 30 parameters in a model of 85 time periods which needs to be solved in each iterative step, non-gradient based search algorithms are prohibitively demanding in terms of computational time. Therefore, we use Generalized Indirect Inference as proposed by Bruins et al. (2015).

Generalized Indirect Inference smooths the objective function by smoothing the moments of the simulated data. Instead of using the discrete choice outcomes (henceforth denoted by $y_{i,j,t}(\Theta)$) to compute the moments, a smoothed function of the latent utilities is applied. We follow Altonji et al. (2013) and use

$$g(u_{i,j,t}(\Theta), \lambda) = y_{i,j,t}(\Theta; \lambda) = \frac{\exp(u_{i,j,t}(\Theta)/\lambda)}{\sum_{k=0}^{2} \exp(u_{i,k,t}(\Theta)/\lambda)} \tag{18}$$

as a kernel. In (18) we denoted latent utilities of alternative j, individual i and time period t by $u_{i,j,t}$. The parameter λ is called smoothing parameter which we will set to a similar value as Altonji et al. (2013) with $\frac{300}{N}$ where N denotes the sample size.[20] Bruins et al. (2015) show that under some regularity assumptions, the smoothed estimator converges to the the standard indirect inference estimator in the limit.[21]

We then minimize the distance between the moments of our observed data set and the

[19] Additionally we account for the probability that the person might not have answered the questions about her salary or weekly working hours. A more detailed discussion of how we handle the observability of wages can be found at the end of Appendix E.

[20] Note that for $N \to \infty$, $\lambda \to 0^+$ and $y_{i,j,t}(\Theta; \lambda) \to y_{i,j,t}(\Theta)$.

[21] See Appendix E for a detailed discussion of how we smooth different components of our model.

smoothed moments of the simulated data set according to the following formula:

$$\hat{\Theta} = \arg\min_{\Theta} \left\{ \sum_{k=1}^{K} \left[\frac{\left(M_k^O - M_k^S(\Theta)\right)^2}{\mathbb{V}\mathrm{ar}(M_k^O)} \right] \right\} \tag{19}$$

where M_k^O denotes the k-th moment of the observed data set, $M_k^S(\Theta)$ the same moment of the simulated data set with parameters Θ, and $\mathbb{V}\mathrm{ar}(M_k^0)$ the variance of the same observed moment. Note that we do not use the optimal weighting matrix due to its poor small-sample properties (Altonji and Segal, 1996).

An overview of the moments we use for estimation is provided in Appendix F.

6 Results

6.1 First Stage Results

In this section we present a short summary of the estimation results of the first stage.

Family Dynamics
Figure 3 plots the probability that a single women finds a partner in the next period. For all three education levels, this probability has a concave shape over the life-cycle. While at the beginning the partner arrival probability for women with secondary education is much higher compared to the other education groups, all three groups are getting closer and are basically similar for women aged 40 or older. Note that we set the probability equal to zero for women in their mid-50s.

Similar to the partner arrival probabilities, the shape of the partner separation (figure 4) probabilities have a comparable shape over all education groups. Separations seem to be higher for women who only have a secondary education. With children, the separation probabilities sink compared to partnerships without children.

Figure 5 shows that in Germany, single mothers are usually an exemptions. Most women receive a child only in a steady partnership, while again the fertility is especially high for women without a high school diploma compared to the other education groups.

Using the above described estimated probabilities, we simulate the the family dynamics

Figure 3: Partner Arrival Probability

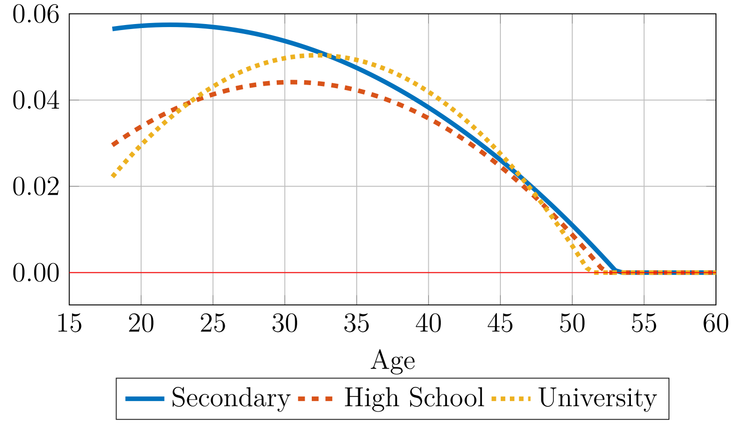


Figure 4: Partner Separation Probability

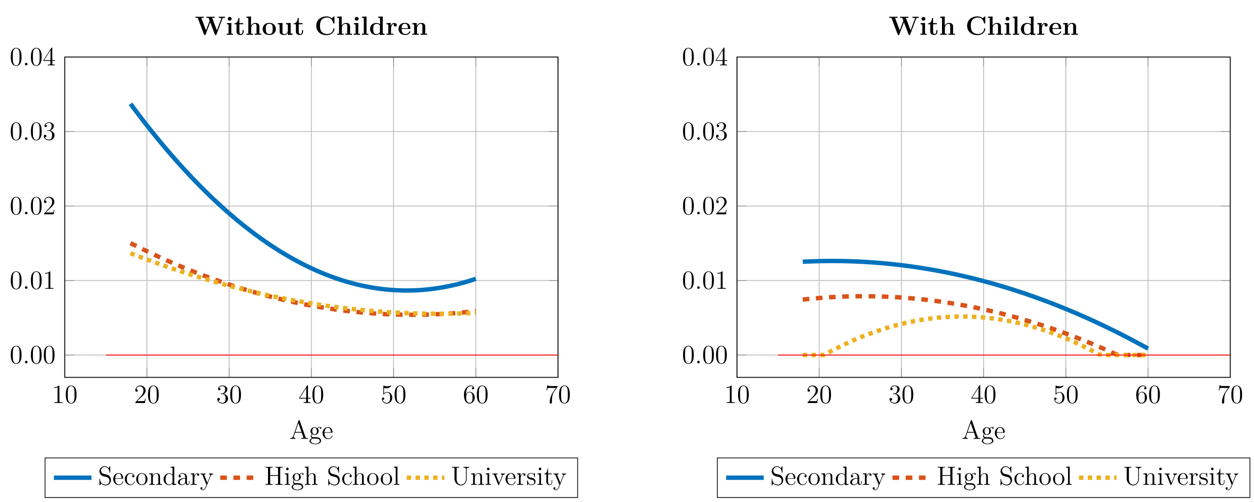
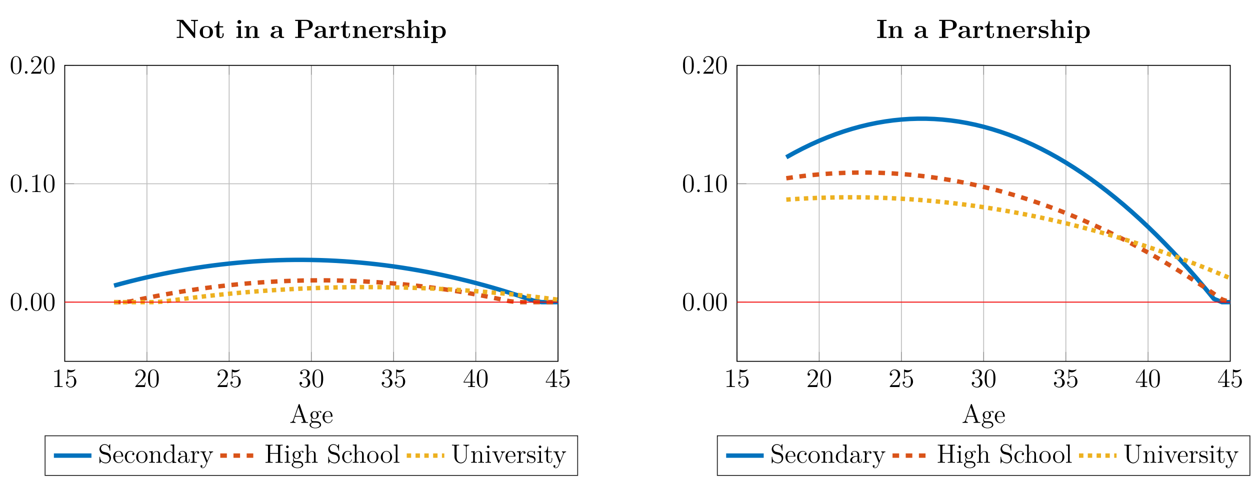
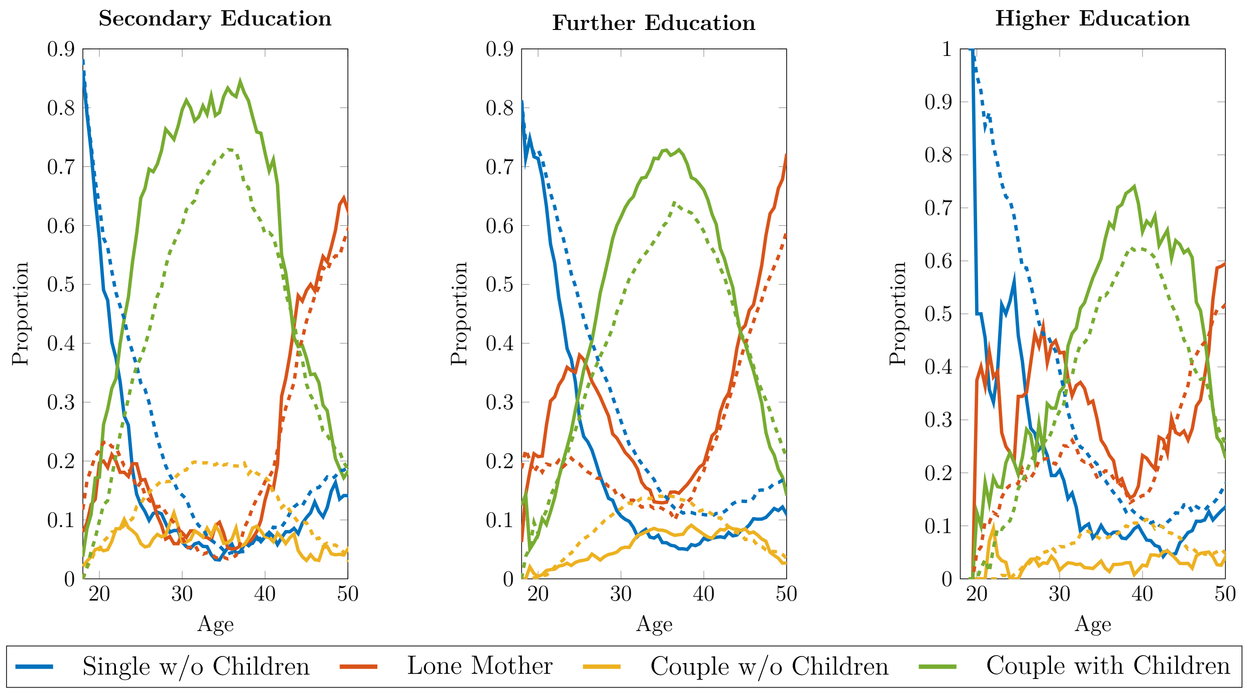


Figure 5: Child Arrival Probability



for our sample. We differ between four groups, singles without children, single with children, partnerships without children and partnerships with children.

Figure 6: Family Dynamics



7 Conclusion

We propose a dynamic discrete choice model of labor supply allowing for time-inconsistent preferences. We use actual large-stakes choices regarding employment observed over many years to estimate time preferences in the labor market. Our identification strategy is based on policy reforms affecting mothers with young children. Mothers' reactions to changes in job protection allow us to determine to what extent mothers take into account changes in future labor market opportunities in their current choices.

Time-inconsistent preferences allow us to explain the gap between planned and actual length of maternal leave: Time-inconsistent agents feel regret when judging previous choices using their ex-post preferences. They are also highly relevant for evaluating policies to support young families and maternal employment. Reactions to policies affecting instantaneous utility of working (e.g. subsidized child care) would be predicted to differ significantly from policies with similar public cost affecting labor market incentives later in life (e.g. a change in income tax or a pension bonus for mothers).

References

Adda, J., C. Dustmann, and K. Stevens (2016): "The Career Costs of Children," *Journal of Political Economy*.

Altonji, J. G. and L. M. Segal (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 14, 353–366.

Altonji, J. G., A. A. Smith, and I. Vidangos (2013): "Modeling Earnings Dynamics," *Econometrica*, 81, 1395–1454.

Anderson, D. J., M. Binder, and K. Krause (2002): "The Motherhood Wage Penalty: Which Mothers Pay It and Why?" *American Economic Review*, 92, 354–358.

Baker, M., J. Gruber, and K. Milligan (2008): "Universal Child Care, Maternal Labor Supply, and Family Well-Being," *Journal of Political Economy*, 116, 709–745.

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.

Blundell, R., M. Costas-Dias, C. Meghir, and J. M. Shaw (2015): "Female Labour Supply, Human Capital and Welfare Reform," .

Bruins, M., J. A. Duffy, M. P. Keane, and A. A. Smith, Jr. (2015): "Generalized Indirect Inference for Discrete Choice Models," .

Chan, M. K. (2014): "Welfare Dependence and Self-Control: An Empirical Analysis," .

Della Vigna, S. and U. Malmendier (2006): "Paying Not to Go to the Gym," *American Economic Review*, 96, 694–719.

Diamond, P. and B. Köszegi (2003): "Quasi-Hyperbolic Discounting and Retirement," *Journal of Public Economics*, 87, 1839–1872.

Eckstein, Z. and K. I. Wolpin (1989): "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience," *The Review of Economic Studies*, 56, 375–390.

Ejrnaes, M. and A. Kunze (2013): "Work and Wage Dynamics around Childbirth," *Scandinavian Journal of Economics*, 115, 856–877.

Fang, H. and D. Silverman (2009): "Time-inconsistency and welfare program participation: Evidence from the NLSY," *International Economic Review*, 50, 1043–1077.

Fang, H. and Y. Wang (2015): "Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions," *International Economic Review*, 56, 565–596.

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002): "Time Discounting and Preference : A Critical Time Review," *Journal of Economic Literature*, 40, 351–401.

Gallant, R. and G. Tauchen (1996): "Which Moments to Match?" *Econometric Theory*, 12, 657–681.

Gourieroux, C., A. Monfort, and E. M. Renault (1993): "Indirect inference," *Journal of Applied Econometrics*, 8, S85—-S118.

Gustman, A. L. and T. L. Steinmeier (2012): "Policy effects in hyperbolic vs. exponential models of consumption and retirement," *Journal of Public Economics*, 96, 465–473.

Haan, P. and V. Prowse (2015): "Optimal Unemployment Insurance and Welfare Benefits in a Life-Cycle Model of Family Labor Supply and Savings," .

Harris, C. and D. I. Laibson (2002): "Hyperbolic Discounting and Consumtpion," in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, Cambridge University Press, chap. 7, 258–298.

Havnes, T. and M. Mogstad (2011): "No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes." *American Economic Journal: Economic Policy*, 3, 97–129.

Hotz, V. J. and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 60, 497–529.

Keane, M. P. and K. I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105, 473.

Kleven, H. J., C. Landais, and J. E. Sogaard (2015): "Parenthood and the Gender Gap: Evidence from Denmark," .

Laibson, D. (1997): "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 112, 443–478.

Magnac, T. and D. Thesmar (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 70, 801–816.

O'Donoghue, T. and M. Rabin (1999): "Doing It Now or Later," *American Economic Review*, 89, 103–124.

Phelps, E. S. and R. A. Pollak (1968): "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies*, 35, 185–199.

Pollak, R. A. (1968): "Consistent Planning," *Review of Economic Studies*, 35, 201–208.

Rust, J. (1994): "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics*, vol. 4, chap. 51, 3081–3143.

Schoenberg, U. and J. Ludsteck (2014): "Expansions in Maternity Leave Coverage and Mothers' Labor Market Outcomes after Childbirth," *Journal of Labor Economics*, 32, 469–505.

Smith, A. A. (1990): "Three Essays on the Solution and Estimation of Dynamic Macroeconomic Models," Ph.d. thesis, Duke University.

Strotz, R. (1956): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 165–180.

Vaskovics, L. A. and H. Rost (1999): *Väter und Erziehungsurlaub*, 179, Stuttgart: Kohlhammer.

Wagner, G. G., J. R. Frick, and J. Schupp (2007): "The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements," *Schmollers Jahrbuch*, 127, 139–169.

Wrohlich, K. (2011): "Labor Supply and Child Care Choices in a Rationed Child Care Market," .

Zmarzlik, J., M. Zipperer, and H. P. Viethen (1999): *Mutterschutzgesetz, Mutterschaftsleistungen, Bundeserziehungsgeldgesetz - mit Mutterschutzverordnung*, Köln: Heymanns, 8 ed.

Appendix A: Survey Questions related to Table 2

In the SOEP questionnaire, individuals are asked if they are currently engaged in paid employment. In general, mothers who are in maternity leave or non-employed do answer that they are *not employed*. All individuals who answer that they are currently not employed will be ask the following question:

Do you intend to engage in paid employment (again) in the future?

- No, definitely not . . . . □
- Probably not . . . . □
- Probably . . . . □
- Yes, definitely . . . . □

Only if individuals do not answer the question with "*No, definitely not*", they will get this follow-up question:

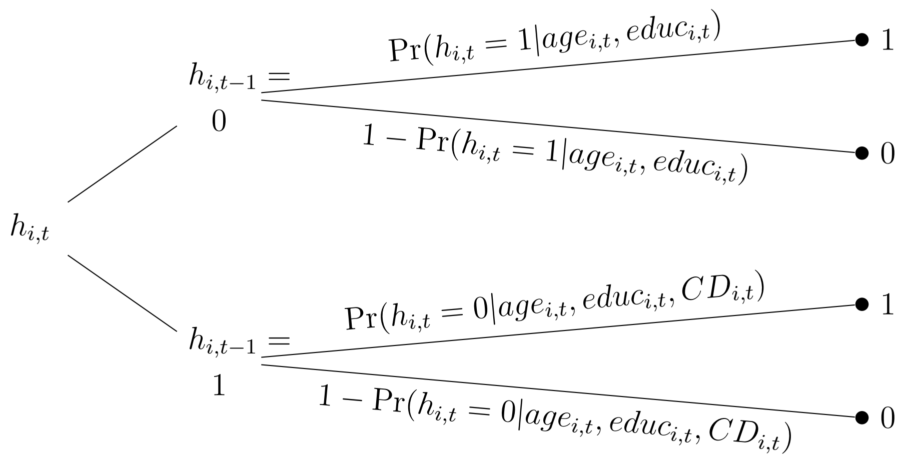
When, approximately, do you want to to start with paid employment?

- As soon as possible . . . . □
- Within this year . . . . □
- In the next two to five years . . . . □
- In the distant future, in more than five years . . . . □

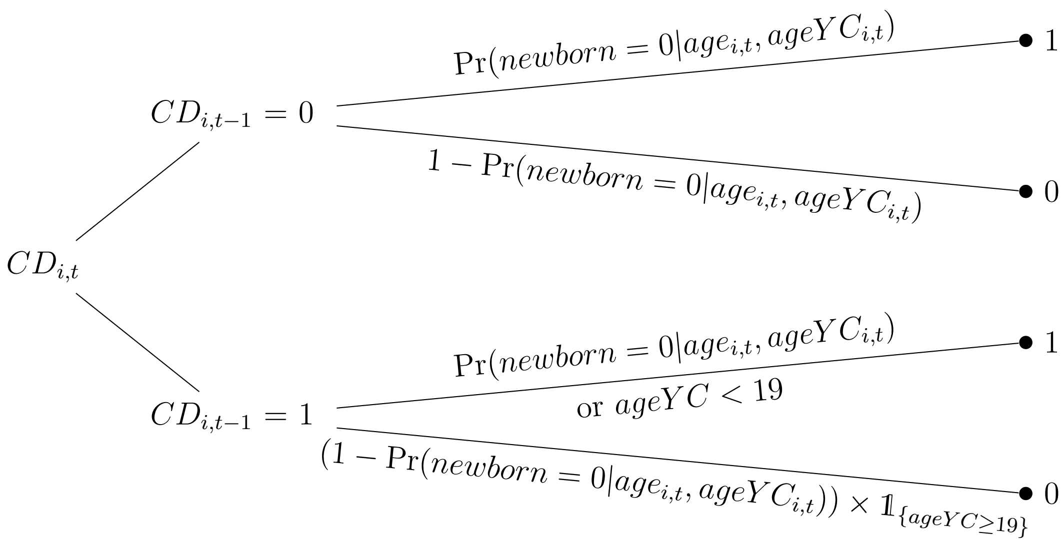
Note that we combined the first two answers ("*As soon as possible*" and "*Within this year*") of this question for table 2.

Appendix B: Exogenous Processes

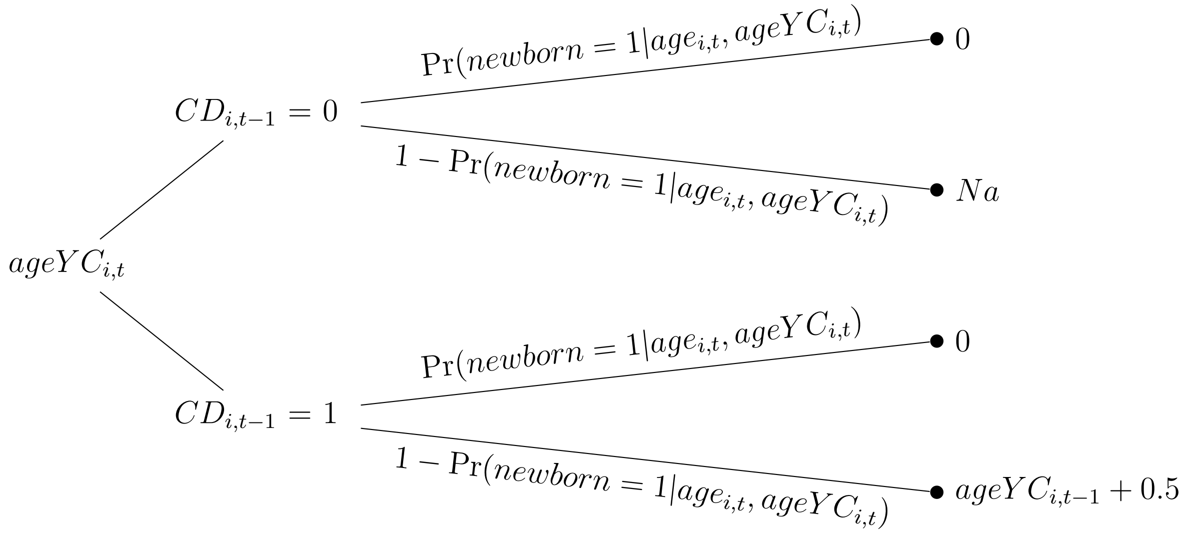
B.0.1 Marriage Status



B.0.2 Child Dummy



B.0.3 Age of Youngest Child



B.0.4 Partner's income

First-stage estimated.

$$wage_{partner;\, i,t} = wage_{partner}(educ_{i,t},\ age_{i,t},\ work\, experience_{i,t})$$

Appendix C: Law of Motion for Job Protection

Table 3 demonstrate how we derive the job protection status from our state variables. We demonstrate this for the policy regime which provides three years of job protection.

Table 3: Law of motion of job protection

age child 1	age child 2	employed_{t-1}	l_t	job prot_{t-1}	job prot_t	note
NaN	NaN	$0 \vee 1$	$0 \vee 1 \vee 2$	0	0	no children
0	NaN	0	$0 \vee 1 \vee 2$	0	0	not eligible
0	NaN	1	$0 \vee 1$	$0 \vee 1$	1	eligible
$0-3$	NaN	0	2	$0 \vee 1$	0	started full-time work
$0.5-3$	NaN	0	$0 \vee 1$	1	1	job protection
$0.5-3$	NaN	0	$0 \vee 1$	0	0	either not eligible or already worked afterwards
$>= 3.5$	NaN	$0 \vee 1$	$0 \vee 1 \vee 2$	$0 \vee 1$	0	child too old
$<= 3.5$	0	$0 \vee 1$	$0 \vee 1$	1	1	next child is born
$<= 3.5$	0	0	$0 \vee 1 \vee 2$	0	0	still not eligible
$<= 3.5$	0	1	$0 \vee 1$	$0 \vee 1$	1	now eligible
> 3.5	0	0	0	0	0	not eligible
> 3.5	0	1	0	0	1	eligible
> 3.5	$0-3$	0	2	$0 \vee 1$	0	started full-time work
> 3.5	$0.5-3$	0	$0 \vee 1$	1	1	job protection
> 3.5	$0.5-3$	0	$0 \vee 1$	0	0	either not eligible or already worked afterwards
> 3.5	$>= 3.5$	$0 \vee 1$	$0 \vee 1 \vee 2$	$0 \vee 1$	0	child too old

Appendix D: Estimation of Exogenous Parameters

Appendix D.1 Job Destruction

If the woman is employed, she will looses her job with a pre-estimated probability of 3.20% each period which corresponds to a yearly job destruction rate of 6.29%. The computation of this rate is the following:

$$job\,separation\,rate = \frac{all\,separations_t}{working\,population_{t-1}} \times \frac{involuntary\,separations}{voluntary\,separations}$$

The involuntary separations include the categories listed in table 4. Women who were in job protection in $t-1$ are excluded from the computations. The probability to loose a job will be denoted by π^{JL}.

Table 4: Distribution of Job Loss Reasons

Reason of Job Loss	Percentage
Company Shut Down	19.89%
Dismissal	49.54%
Temporary Contract Expired	23.20%
Relocation by Employee	7.37%

Appendix E: Smoothing of the Objective Function

To be able to use gradient-based optimization algorithms, we smooth the moments we estimate from the simulated data set. In this appendix we discuss further details of how we adjust the data. As stated above, we use the same kernel as Altonji et al. (2013), with a similar value for lambda. However, we only clone our observed data set five times, instead of twenty times, mainly to shorten the computational burden.

The life-time utility of any choice depends on values of previous periods, especially human capital, the possibility of employment and the job protection status depend on previous made choices. Smoothing choices leads automatically to a smoothed version of the job protection status of last period, since an individual looses the her job guarantee when working full-time. In general, smoothing complicates the usage of state variables which depend on the last period.

To cope with this problem, we will slightly extend the method proposed in Bruins et al. (2015) to handle such cases. Given the three possible employment states of last periods and the possibility to have job protection in the current period, we compute for each of the six possible state combinations the utilities for all three choices. Let those be denoted by $u(jp_{t-1}, l_{t-1})$, $jp_{t-1} \in \{0, 1\}$, $l_{t-1} \in \{0, 1, 2\}$). Then the smoothed choice is computed in the following manner:

$$choice_{i,j,t} = \sum_{p=0}^{1}\sum_{q=0}^{2}\left[\frac{\exp\left(\frac{u_{i,j,t}(\Theta|jp_{t-1}=p,l_{t-1}=q)}{\lambda}\right)}{\sum_{k=0}^{2}\exp\left(\frac{u_{i,k,t}(\Theta|jp_{t-1}=p,l_{t-1}=q)}{\lambda}\right)}\right] \times \widetilde{jp}_{i,p,t-1} \times y_{i,q,t-1}(\Theta,\lambda)$$

where $\widetilde{jp}_{i,p,t-1}$ denotes the smoothed job protection status of last period and $\widetilde{jp}_{i,p,0} = 0, y_{i,q,0} = 0$ and $t = 1 \dots T$. We use the variable $choice_{i,t}$ here, because to derive the final smoothed employment status, we additionally have to incorporate the employment possibilities. It is important to note that women do have an individual labor market entry age, i.e. not every women enters the labor market at time $t = 1$. When ever an individual enters the labor market we set the necessary last periods states to the values designed in $t = 0$.

The event that an unemployed individual will receive a job offer or an employed person will be laid-off is discrete and influences the particular choice an individual can make. Any small change in the parameters influencing the job offer probability will lead to changes in discrete outcomes. Hence, we also smooth the job offer probability. We do this similar to Altonji et al. (2013).

Starting at equation (6), we can describe the event of getting a job offer as

$$index_{i,t}^{JO} = \gamma_{JO,sec}\mathbb{1}_{\{educ_i=1\}} + \gamma_{JO,high}\mathbb{1}_{\{educ_i=2\}} + \gamma_{JO,uni}\mathbb{1}_{\{educ_{ki}=3\}} + \gamma_{JO,e}e_{i,t} + \epsilon_{i,t}^{JO}$$
$$\text{with } \epsilon_{i,t}^{JO} \sim \mathcal{N}(0,1).$$

which allows us to express the discrete event of a job offer as

$$JO = \mathbb{1}_{\{index_{i,t}^{JO}>0\}} \text{ given } l_{i,t-1} = 0.$$

Using the same kernel as in before, we can derive the smooth event of a job offer as

$$\widetilde{JO}_{i,t} = \frac{\exp\left(\frac{index_{i,t}^{JO}}{\lambda}\right)}{1+\exp\left(\frac{index_{i,t}^{JO}}{\lambda}\right)} \times y_{i,0,t-1}(\Theta;\lambda)$$

Note that $\lim_{\lambda \to 0^+} \widetilde{JO} = JO \in 0, 1$ holds.[22]

To compute the final smooth employment possibility for each individual we have to factor in the job protection status as well. We do this in the following manner (here only the probability for having job protection this period):

$$\widetilde{jp}_{i,1,t} = \sum_{p=0}^{1} \sum_{q=0}^{2} \left[jp(\Theta | jp_{i,t} = p, l_{i,t} = q, \Omega_{i,t} \setminus \{jp_{i,t}, l_{i,t}\}) \times \widetilde{jp}_{i,p,t-1} \times y_{i,q,t-1}(\Theta; \lambda) \right]$$

The overall probability that an individual will have the choice possibility of employment in the current period is then given by:

$$\Pr(Employment)_{i,t} = [\widetilde{JO}_{i,t} + (1-\pi)(1 - y_{i,0,t-1}(\Theta; \lambda))] \times (1 - \widetilde{jp}_{i,1,t}) + 1 \times \widetilde{jp}_{i,1,t}$$

Finally, the smoothed labor market outcome of period t can be expressed as:

$$y_{i,j,t} = \begin{cases} choice_{i,0,t} + [1 - \Pr(Employment)] \times (choice_{i,1,t} + choice_{i,2,t}) & \text{if } j = 0 \\ \Pr(Employment) \times choice_{i,1,t} & \text{if } j = 1 \\ \Pr(Employment) \times choice_{i,2,t} & \text{if } j = 2 \end{cases}$$

The on-the-job human capital will also be treated differently. We will add human capital according to the “distribution” of choices:

$$\tilde{e}_{i,t} = \tilde{e}_{i,t-1}(1 - \eta) + \lambda y_{i,1,t-1} + 0.25 y_{i,2,t-1}$$
$$\text{with } \tilde{e}_{i,t} = 0$$

Again, note that $\lim_{\lambda \to 0^+} \tilde{e}_{i,t} = e_{i,t}$ holds.

Observability of Wages

We use the approach by Altonji et al. (2013) to handle missing observations in the observed data for all exogenous missing observations. For endogneously missing observations, i.e. for missing wages, we use a different approach, since the observability is strongly influenced by the labor choice.

Without smoothing, we only would include wages for simulated individuals, if the simulated individual is employed and the respective observed observation had an interview in

[22]Since, we estimate the job destruction rate outside the model, we do not need to smooth it.

the period. The smoothing leaves us with a distribution over the employment status which prohibits to use the previous described procedure. Instead we apply the following algorithm:

$$w_{i,t}^{observed} = \begin{cases} w_{i,t} & \text{if } (1 - y_{i,0,t} > z_{i,t}) \wedge (interviewD_{i,t} = 1) \\ NaN & \text{else} \end{cases}$$
$$\text{with } z_{i,t} \sim \mathcal{U}(0,1)$$

where $\mathcal{U}$ denotes a uniform distribution and $interviewD$ equals one if an interview took place in the given period. Note that for $\lambda \rightarrow 0^+$, we end up with the wages we would observe without smoothing.

Additional to the mentioned algorithm in section, wages are only observed if women answer questions about their weekly working hours and their monthly net income. We do control for this by estimating in the first stage a probit model of the following form:

$$\begin{aligned}\Pr(No\ Answer) = \Phi\Big(&\beta_{const}^{NA} + \beta_{educ}^{NA} educ_{i,t} + \beta_{age}^{NA} age_{i,t} + \beta_{NC}^{NA} NC_{i,t} \\ &+ \beta_{partner}^{NA} partner_{i,t} + \beta_{ageYC}^{NA} ageYC_{i,t}\Big)\end{aligned}$$

where $ageYC$ is set to zero if no child is present. For the estimation we include only observations for individuals who are working and have an interview in the given period.

After simulating the data for the given set of parameters, we compute for each woman who is working and having an interview in the given period, the probability of a missing answer. Is the computed probability lower than a individual and period specific uniformly distributed draw, we set the wage to a missing.[23]

Appendix F: Overview of Moments

Appendix F.1 Employment Rates

To identify the parameters of the utility function and the job offer rate. We use the following formula for the moments:[24]

[23] These draws are also computed in the first stage to ensure convergence.

[24] Note that we smooth the discrete outcomes of our simulation for being able to use gradient-based optimization algorithms

- All employment:

$$\mathbb{E}[(1 - y_{i,0,t}) = 1] = \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{1}{\sum_{i=1}^{N} \sum_{t=1}^{T} \mathbb{1}_{y_{i,0,t}}} (1 - y_{i,0,t})$$

- Part-Time Employment

$$\mathbb{E}[y_{i,1,t} = 1] = \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{1}{\sum_{i=1}^{N} \sum_{t=1}^{T} \mathbb{1}_{y_{i,1,t}}} y_{i,1,t}$$

and for the variance of these moments:

- All employment:

$$\mathbb{V}\text{ar}(\mathbb{E}[(1 - y_{i,0,t}) = 1]) = \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\frac{1}{\sum_{i=1}^{N} \sum_{t=1}^{T} \mathbb{1}_{y_{i,0,t}}} \right)^2 \mathbb{V}\text{ar}(1 - y_{i,0,t})$$

- Part-Time Employment

$$\mathbb{V}\text{ar}(\mathbb{E}[y_{i,1,t} = 1]) = \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\frac{1}{\sum_{i=1}^{N} \sum_{t=1}^{T} \mathbb{1}_{y_{i,1,t}}} \right)^2 \mathbb{V}\text{ar}(y_{i,1,t})$$

Table 5: Employment Rates

Description	Number of Moments
All Employment (separately for all 3 education groups)	
all	3
single women, no child	3
married women, no child	3
lone mothers	3
married mothers	3
age youngest child $\in [0, 3[$	3
age youngest child $\in [3, 6[$	3
age youngest child $\in [6, 11[$	3
age youngest child $\in [11, 18]$	3
mothers with multiple children	3
Part-Time Employment (separately for all 3 education groups)	
all	3
single women, no child	3
married women, no child	3
lone mothers	3
married mothers	3
age youngest child $\in [0, 3[$	3
age youngest child $\in [3, 6[$	3
age youngest child $\in [6, 11[$	3
age youngest child $\in [11, 18]$	3
mothers with multiple children	3
Total	60

Appendix F.2 Regime Specific Employment Rates

We use these to identify the discounting parameters. The same formulas as above apply here.

Table 6: Employment Rates when potentially in Job Protection

Description	Number of Moments
All Employment (for all regimes separately)	
age youngest child $\in [0, 0.5[$	3
age youngest child $\in [0.5, 1[$	3
age youngest child $\in [1, 1.5[$	3
Part-Time Employment (for all regimes separately)	
age youngest child $\in [0, 0.5[$	3
age youngest child $\in [0.5, 1[$	3
age youngest child $\in [1, 1.5[$	3
Total	18

Appendix F.3 Transitions into Employment

For identifying the parameters of the utility function and the job offer rate. We estimate the following Linear Probability Model for the groups *all*, and *women with children*:

$$\begin{aligned}(1 - y_{i,0,t})y_{i,0,t-1} = & \beta_{sec}^{Tran} \mathbb{1}_{educ_i=1} + \beta_{high}^{Tran} \mathbb{1}_{educ_i=2} + \beta_{uni}^{Tran} \mathbb{1}_{educ_i=3} \\ & + \beta_{PT}^{Tran} experPT_{i,t-1} + \beta_{FT}^{Tran} experFT_{i,t-1} \\ & + \beta_{PT^2}^{Tran} experPT_{i,t-1}^2 + \beta_{FT^2}^{Tran} experFT_{i,t-1}^2\end{aligned}$$

It is important to note that these are not the classical transition rates usually used in this field, when indirect inference is applied. We use a linear probability model here, since it seems important to include experience as an approximation of human capital which enters the job offer rate. We use the squared heteroskedasticity-consistent standard errors for the parameters as variances for the moments.

Table 7: Transitions From Out of Work into Work (Linear Probability Model)

Description	Number of Moments
All Women	
Coefficients of Linear Probability Model	7
Women with Children	
Coefficients of Linear Probability Model	7
Total	14

Appendix F.4 Transitions into Non-Employment

We would like to compute $\mathbb{E}[y_{i,0,t}|y_{i,0,t-1}=0]$ for several subgroups. With smoothed choices, we can derive the following:

$$\begin{aligned}\mathbb{E}[y_{i,0,t}|y_{i,0,t-1}=0] &= \mathbb{E}[y_{i,0,t}|(1-y_{i,0,t-1})=1] \\ &= \sum_{i=1}^{N}\sum_{t=2}^{T}\frac{(1-y_{i,0,t-1})}{\sum_{i=1}^{N}\sum_{t=2}^{T}(1-y_{i,0,t-1})}y_{i,0,t}\end{aligned}$$

and we use the following to estimate the variance:

$$\mathbb{V}\text{ar}(\mathbb{E}[y_{i,0,t}|y_{i,0,t-1}=0]) = \sum_{i=1}^{N}\sum_{t=2}^{T}\left(\frac{(1-y_{i,0,t-1})}{\sum_{i=1}^{N}\sum_{t=2}^{T}(1-y_{i,0,t-1})}\right)^2\mathbb{V}\text{ar}(y_{i,0,t})$$

Table 8: Transitions From Work to Out of Work (separately for all 3 education groups)

Description	Number of Moments
all women	3
single women with no children	3
married women with no children	3
lone mothers	3
married mothers	3
past wage in bottom percentile ($wage_{i,t-1} < P_{10}$)	3
$wage_{i,t-1} < P_{50}$	3
$wage_{i,t-1} < P_{90}$	3
Total	24

Appendix F.5 Log Wages Regressions on Accumulated Experience and Lagged Wages

We run the regressions of the following form:

$$\begin{aligned} wage_{i,t} =& \beta^{W1}_{sec}\mathbb{1}_{educ_i=1} + \beta^{W1}_{high}\mathbb{1}_{educ_i=2} + \beta^{W1}_{uni}\mathbb{1}_{educ_i=3} \\ &+ \beta^{W1}_{L\_Wage} wage_{i,t-1} \\ &+ \beta^{W1}_{Exper} ln(experPT_{i,t} + experFT_{i,t}) \\ &+ \beta^{W1}_{L\_Exper} ln(experPT_{i,t-1} + experFT_{i,t-1}) \end{aligned}$$

and add the variance of the residuals as an extra moment.

Table 9: Wage Regression

Description	Number of Moments
regression coefficients	6
variance of the residual	1
Total	7

Appendix F.6 Distribution of Log Wages During Working Life

We use the following formula for the moments:

- mean:

$$\mathbb{E}[wage_{i,t}|y_{i,1,t} = 1] = \sum_{i=1}^{N}\sum_{t=1}^{T} \frac{y_{i,1,t}}{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,1,t}} wage_{i,t}$$

- percentiles:

$$\sum_{i=1}^{N}\sum_{t=1}^{T} \frac{y_{i,1,t}}{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,1,t}} \mathbb{1}_{wage_{i,t}<P_\alpha}$$

  where P_α denotes the α-percentile of the distribution of wages.

and the following formulas for their variance:

- variance of the mean:

$$\text{var}(wage_{i,t}) \times \sum_{i=1}^{N}\sum_{t=1}^{T}\left(\frac{y_{i,1,t}}{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,1,t}}\right)^2$$

- variance of percentiles:

$$\text{var}(\mathbb{1}_{wage_{i,t}<P_\alpha}) \times \sum_{i=1}^{N}\sum_{t=1}^{T}\left(\frac{y_{i,1,t}}{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,1,t}}\right)^2$$

Description	Number of Moments
Full-Time Employment (separately for all 3 education groups)	
mean	3
$log(wage) < 10^{th}\ percentile$	3
$log(wage) < 25^{th}\ percentile$	3
$log(wage) < 50^{th}\ percentile$	3
$log(wage) < 75^{th}\ percentile$	3
$log(wage) < 90^{th}\ percentile$	3
Part-Time Employment (separately for all 3 education groups)	
mean	3
$log(wage) < 10^{th}\ percentile$	3
$log(wage) < 25^{th}\ percentile$	3
$log(wage) < 50^{th}\ percentile$	3
$log(wage) < 75^{th}\ percentile$	3
$log(wage) < 90^{th}\ percentile$	3
Total	36

Appendix F.7 Further Moments in Log Wages

We run two additional regressions, one in which we regress $\log(wage_{i,t})$ on $\frac{age_{i,t}}{10}$ and of $\Delta\log(wages_{i,t})$ on $\Delta\log(experFT_{i,t} + experPT_{i,t})$. We run separate regressions for the different education levels. Note that we do not use a constant in these regressions. We also compute the mean yearly change in $\log(wage_{i,t})$, once for women who worked full-time the year before and once for women who worked part-time the year before.

For the regressions, we again use the squared heteroskedasticity-consistent standard errors as variances for the moments. For the mean yearly changes we use the following formulas:

$$\sum_{i=1}^{N}\sum_{t=2}^{T}\frac{y_{i,j,t}}{\sum_{i=1}^{N}\sum_{t=2}^{T}y_{i,j,t}}\left(log(wage_{i,t})-log(wage_{i,t-1})\right)\ \text{ with } j\in\{1,2\}$$

for the mean and

$$\mathbb{V}\text{ar}\left(log(wage_{i,t})-log(wage_{i,t-1})\right)\sum_{i=1}^{N}\sum_{t=2}^{T}\left(\frac{y_{i,j,t}}{\sum_{i=1}^{N}\sum_{t=2}^{T}y_{i,j,t}}\right)^2\ \text{ with } j\in\{1,2\}$$

for the variance of the mean.

Description	Number of Moments
Regression on *age*	3
Regression on $log(experience)$	3
Changes in $log(wage)$ if working full-time before	3
Changes in $log(wage)$ if working part-time before	3
Total	12