Time Preferences and Female Labor Supply

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

June 24, 2017

– Preliminary and Incomplete–

Abstract

We estimate a dynamic life-cycle model of labor supply with a focus on women's time preferences. We extend the dynamic discrete choice model of female labor supply to accommodate potentially time-inconsistent behavior. For the identification of the time preferences we exploit natural experiments: we use several parental leave reforms in Germany which extend parental leave from one to three years. Preliminary results provide evidence for significant time inconsistent behavior. Our approach allows us to shed light on the importance of time-inconsistent preferences in explaining important labor supply choices.

JEL: J24, D91, H31

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

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

*FU Berlin

†DIW Berlin

‡OECD

1 Introduction

Dynamic structural models of female labor supply are used to analyze individual behavior over the life cycle and to evaluate a large array of counterfactual policy reforms, see e.g. Eckstein and Wolpin (1989), Keane and Wolpin (1997), and more recently 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 utility accruing throughout life. Therefore, assumptions about how individuals discount these future streams when making decisions, i.e. assumptions about their time preferences, are crucial 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), models of labor supply continue to rely on the assumption of time consistent exponential discounting with only a few exceptions (Fang and Silverman, 2009; Chan, 2014). One reason for this assumption is that nonparametric identification of time preference parameters can generally not be achieved in a standard framework (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 transition probabilities can be exploited that does not affect per-period utilities. In other words, for the identification exclusion restrictions 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 possibilities. Crucially for our identification strategy, job protection does not affect directly the flow utility of mothers but only future employment transitions by guaranteeing 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 nests 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).

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 in-work benefits 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. Time preferences are thus of prime importance for predicting reactions to different policy instruments. Based on the estimated parameters of the model, we simulate the employment responses of various policy reforms aiming to improve the labor market situation of women including the introduction of in-work benefits.

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) and more recently Laibson (1997) and O'Donoghue and Rabin (1999). Besides laboratory experiments, economists have tried to estimate time-preferences with data from the field (see for example Laibson, 1997; Martinez et al., 2017; Della Vigna and Malmendier, 2006). One of the results seems that individuals discount differently in different domains as Augenblick et al. (2015) shows, that individuals exhibit a much stronger present bias in a real effort task than in a monetary setting. This is important for this study, because we focus on the trade-off of spending more time with young children or returning to employment to invest in your future working career.

Since applications of (quasi-)hyperbolic discounting have become more popular in experimental and observational studies, researchers started to model quasi-hyperbolic preferences also in structural models, e.g. Harris and Laibson (2002) for consumption decisions, and Diamond and Köszegi (2003) and 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 labor market behavior of a representative sample of women. This allows us to build a richer model of labor market frictions taking into account the role of family composition in labor supply as 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. Havnes and Mogstad (2011)). For example, several reduced form studies for Germany find sizable positive employment effects of family policy reforms, such as the introduction of subsidies of child care Bauernschuster and Schlotter (2015), or changes in the duration of paid parental leave. Given the large estimated career costs of prolonged interruptions, the way in which these expected costs are incorporated in women's labor market choices is especially important.

The remainder of the paper is structured as follows. Section 2 presents the dynamic labor supply model. Section 3 presents the data and introduces first suggestive evidence for time-inconsistent behavior in the context of maternal employment. Section 4 provides background information on the institutions. Section 5 discusses identification of key parameters and the estimation procedure. Section 8 concludes.

2 Economic Model

We model mothers' labor supply choices 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 stemming from 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 job offer arrivals . 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 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. Although we do not model fertility choices endogenously, we capture the impact on children on labor supply. Following Blundell et al. (2015) we estimate the probabilities of childbirth in a first step.

We explicitly model the development of human capital as a 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 human capital depreciates at 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. Furthermore, the employment status influences future job offer probabilities and therefore future employment possibilities.

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 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

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

[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).

$$
\begin{aligned}
u_{i,t} = & \frac{[c_{i,t}/\bar{c}_{eq}(CD_{i,t})]^{(1-\gamma_C)} - 1}{1-\gamma_C} \\
& \times \exp\Bigg(\Big[\gamma_{PT_{low}} \mathbb{1}_{\{l_{i,t}=1\}} + \gamma_{FT_{low}} \mathbb{1}_{\{l_{i,t}=2\}}\Big] \times \mathbb{1}_{\{educ=low\}} \\
& \qquad + \Big[\gamma_{PT_{high}} \mathbb{1}_{\{l_{i,t}=1\}} + \gamma_{FT_{high}} \mathbb{1}_{\{l_{i,t}=2\}}\Big] \times \mathbb{1}_{\{educ=high\}} \Bigg) \\
& \times \Big(\gamma^{PT}_{ageYC,0} + \gamma^{PT}_{ageYC,1} ageYC_{i,t} + \gamma^{PT}_{ageYC,2} ageYC^2_{i,t}\Big)^{\mathbb{1}\{l_{i,t}=1, CD_{i,t}=1\}} \\
& \times \Big(\gamma^{FT}_{ageYC,0} + \gamma^{FT}_{ageYC,1} ageYC_{i,t} + \gamma^{FT}_{ageYC,2} ageYC^2_{i,t}\Big)^{\mathbb{1}\{l_{i,t}=2, CD_{i,t}=1\}} \\
\text{with} \qquad & \bar{c}_{eq}(NC_{i,t}) = 1 + 0.4\, CD_{i,t}
\end{aligned}
$$

where $c_{i,t}$ denotes the consumption, $\bar{c}$ an equivalence scale[3] which accounts for the number of household members, $CD_{i,t}$ indicates the presence of children, 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.

The first lines of equation (2.2) represents the basic trade-off between leisure and consumption, where we assume that the preferences between the two are non-separable. The utility derived from consumption is modeled using a standard CRRA function. We allow for the possibility that utility derived from the amount of leisure differs with the presence of children and their age.

Wages and Human Capital. The labor supply decision also strongly depends on consumption opportunities, for which wages are one of the pivotal factors. We model the wage process similar to Blundell et al. (2015):

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

The hourly wage rate depends on the individual's highest education degree and accumulated human capital and is measured with error, where $\xi_{i,t}$ follows a normal distribution with standard deviation σ_ξ. Since levels of education do not change over the life-cycle in our model, wage differences over time are mostly driven by accumulation of human capital on the job and deterioration out of employment. The process of human capital is given as

[3] We assume that $\bar{c} = 1$ for a single, we add 0.4 for every child.

follows:

$$e_{i,t} = e_{i,t-1}(1-\eta) + \begin{cases} 0 & \text{if } l_{i,t-1} = 0 \text{ (not employed)} \\ \lambda & \text{if } l_{i,t-1} = 1 \text{ (part-time)} \\ 0.5 & \text{if } l_{i,t-1} = 2 \text{ (full-time)} \end{cases} \tag{2}$$

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[4] which can only be offset if the individual is employed. The possible accumulation depends on whether a woman works for part-time or 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:

$$\begin{aligned} c_{i,t} =& w_{i,t} \times 452 \times (2 \times \mathbb{1}_{\{l_{i,t}=2\}} + \mathbb{1}_{\{l_{i,t}=1\}}) \\ & - TT(earn_{i,t}^{W}, ageYC_{i,t}) \\ & - cc^{E} \times 452 \times CD_{i,t} \times (2 \times \mathbb{1}_{\{l_{i,t}=2\}} + \mathbb{1}_{\{l_{i,t}=1\}}) \end{aligned} \tag{3}$$

where $earn_{i,t}^{W}$ the gross earnings of the woman, TT for the German tax and transfer system, $ageYC_{i,t}$ the age of a potential youngest child 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 working day. Therefore, the first part of line 1 of equation (3) 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 relevant features of the German tax and transfer system which depends on the earnings and children. In particular, we model income taxation, social security contributions as well as child benefits. It is important to note that women who do not work are eligible for unemployment benefits or social assistance benefits which partly depend on the household income. Maternity benefits depend on the age of the youngest child and the policy

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

regime the child was 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):

$$cc^E = cc^S\pi + cc^{NS}(1 - \pi) \tag{4}$$

where cc^E denotes the expected, cc^S the average subsidized, and cc^{NS} the average non-subsidized childcare costs per hour. The parameters π denotes the probability of being able to find 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 a job offer with probability:

$$\pi_{i,t}^{JO} = \gamma_{JO,low}\mathbb{1}_{\{educ=low\}} + \gamma_{JO,high}\mathbb{1}_{\{educ=high\}} \tag{5}$$

Therefore, the job offer probability depends on education and the current amount of human capital. Importantly, 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 their hours at any time within the job protection period.

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 is involuntarily laid off cannot work in the current period.

Dynamics of Family Composition. All family dynamics are modeled as exogenous stochastic processes. The probability with which a woman receives a child depends on her age, and the presence of other children and their respective age. It is assumed that all children live with their mother until the age of 18.

2.3 Intertemporal Optimization

We assume that the utility associated with different labor supply choices is only perceived by the researcher with error, adding an alternative-specific error term $\varepsilon_{i,l,t}$ which we assume to be independently and identically distributed over time and labor supply choices following a type-1 extreme value distribution, such that:

$$v_{i,l,t} = u_{i,t}(l_t) + \varepsilon_{i,l,t} \tag{6}$$

When considering utility streams arising at different points of time, individuals discount any future values by β additionally to the standard discounting factor δ. If $\beta = 1$, the standard exponential discounting framework applies, but if $\beta < 1$ the individual displays a present bias (O'Donoghue and Rabin, 1999), indicating the impulse for immediate gratification.

In any given period t, an individual then maximizes her expected life-time utility U_t:

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

where we drop the index i for ease of notation, $\mathbb{E}$ the expectations operator and Ω_t the state space at time t:

$$\Omega_t = \{age_t, e_t, CD_t, ageYC_t, jp_t, \varepsilon_t, l_{t-1}, jp_{t-1}\}.$$

where $\varepsilon_t \equiv (\varepsilon_{l_t=0}, \varepsilon_{l_t=1}, \varepsilon_{l_t=2})$ and jp_t is a dummy for a women having job protection in period t.

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 different time periods in the future. 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 instance, 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 contrast, individuals who are absolutely not aware of their time-inconsistencies are called *naïve*. In this paper, we assume that agents are fully naïve. This is in line with Fang and Wang (2015). While Chan (2014) finds some evidence of partial sophistication, he cannot reject a specification in which individuals are fully naïve.

2.4 Solution of the Structural Model

To build the foundations for our identification strategy we start with the solution of the model. We first focus on the long-run utility, i.e. the utility of exclusively future periods, contrasting utility streams arising in $t+j$ and $t+j+1$ (with $j \geq 1$) from the point of view of time t. Making use of the assumption that agents are fully naïve, we rewrite the value function in a recursive manner (see Fang and Silverman, 2009; Chan, 2014):

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

Note that since $j \geq 1$ we are considering only future periods, so that β is not included in (8). To simplify notation, we denote the term $\mathbb{E}(V_{t+j+1}(\Omega_{t+j+1}) | \, \Omega_{t+j})$ henceforth with $\mathbb{E}\max_{t+j}$. We use the subscript $t+j$ and not $t+j+1$ to emphasize that we are interested in the expected maximum based on the information set available in period $t+j$. If we refer to a specific future realization of the state space in $t+j+1$, i.e. not the expected development of the state space, we will denote this by $\mathbb{E}\max_{t+j}(\widetilde{\Omega}_{t+j+1})$. 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 $\mathbb{E}\max_{t+j}$ is then given by:

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

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

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

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

$$
\begin{aligned}
\mathbb{E}\max_{t+j} \mid (l_{t+j} = 0, \Omega_{t+j}) = & \sum_{\widetilde{\Omega}_{t+j+1}} \Pr\left(\widetilde{\Omega}_{t+j+1}\middle|\,\Omega_{t+j}\right) \Big[\pi^{JO}_{i,t+j} \times \mathbb{E}\max^{\text{emp}}_{t+j}(\widetilde{\Omega}_{i,t+j+1}) \\
& + (1-\pi^{JO}_{i,t+j}) \times \mathbb{E}\max^{\text{non-emp}}_{t+j}(\widetilde{\Omega}_{i,t+j+1})\Big] \\
\mathbb{E}\max_{t+j} \mid (l_{t+j} \in \{1,2\}, \Omega_{t+j}) = & \sum_{\widetilde{\Omega}_{t+j+1}} \Pr\left(\widetilde{\Omega}_{t+j+1}\middle|\,\Omega_{t+j}\right) \Big[\qquad (11) \\
& (1-\pi^{JL}_{i,t+j}) \times \mathbb{E}\max^{\text{emp}}_{t+j}(\widetilde{\Omega}_{i,t+j+1}) \\
& + \pi^{JL}_{i,t+j} \times \mathbb{E}\max^{\text{non-emp}}_{t+j}(\widetilde{\Omega}_{i,t+j+1})\Big]
\end{aligned}
$$

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

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

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).[5] While the SOEP interviews individuals on a yearly basis, it asks participants to fill out a monthly calendar of the previous year. In particular, individuals are asked about 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.[6] 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.[7]

[5] 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.

[6] 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.

[7] 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.

3.2 Suggestive evidence of time-inconsistent choices

Table 1: Return to the Labor Market after Last Child

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

Notes: Sample: Women in SOEP observed from the birth of their last child until they re-enter the labor market (or are right-censored but state their wish to return to the labor market) and who have a job guarantee for 3 years. "Prefer." refers to preferred length of career breaks as recorded in the "1st Year", "2nd Year" and "3rd Year" after the birth of the last child. "Real." refers to the (ex post) observed duration of career break based on later waves of the panel.

Time-inconsistent preferences such as those considered here generate systematic errors in predicting own behavior in the future: Plans about future decisions are continuously revised. Indeed, comparing survey data with later realizations reveals suggestive evidence of such regret. Table (1) presents data that we do not use for identification or estimation but which reveals a noticeable gap between the preferred date of return to employment after child-birth and the actual date.

The "1st Year" column of table 1 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 plan to be back in employment within five years or earlier. 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.

The previous trend continues in the second and the third year[8] which suggests that one-time errors like unexpected childcare availability or lacking support of the partner are unlikely to explain the gap fully. The mother should have learned about these errors in the first year and adapt her expectations for the second and third year. For models based on rational expectations and exponential discounting it is hard to explain the observed gap between preferences and realizations. Mothers with rational expectations should on average predict their behavior correctly. Time-inconsistent behavior can account for this pattern in a very natural way.

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 period, several policy reforms changed the period of maternal job protection which we exploit to identify the discount parameters of our model. We concentrate on six major expansions of maternity leave coverage between 1986 and 1993.[9] 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 be dismissed and were guaranteed a comparable job to their previous held one. 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

[8]By focusing on women with a guaranteed right to return to their previous employment we can be certain that labor demand plays no role in employment decisions. Delaying return within the first three years is also generally possible.

[9]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.

the 8th week and the end of the 6th 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 2. 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.[10] Maternity payments from week six to eight remained at an income replacement of 100% or DM 600[11] if the mother was unemployed before. Between month three to six maternity benefits declined from DM 750 to DM 600[11] per month. From the seventh month to the 10th month (and later 12th 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).

Table 2: Parental Leave Reforms from 1986 until 2006

	Month, Year	**Job Prot. (Law)**	**Job Prot. (Model)**	**Maternity Benefits**
Regime I	January, 1986	10 months	12 months	3-6 months DM 600,[11] 7-10 months means tested
	January, 1988	12 months		up to 12 months
Regime II	July, 1989	15 months	18 months	up to 15 months
	July, 1990	18 months		up to 18 months
Regime III	January, 1992	36 months	36 months	up to 18 months
	January, 1993			up to 24 months
	January, 2007	maternal benefits are related to previous earnings		

A further increase in the job protection and maximum maternity benefit time period from 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.[12] Minor changes in family policy were introduced in

[10] 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).

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

[12] 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

2001, but the core regime of 1993 still continued.

Table 2 categorizes these reforms into three periods, labeled Regimes I-III. There are several reasons why we summarize the reforms. 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.

5 Identification and Estimation

5.1 Identification

5.1.1 Intuition

Early work by Rust (1994) argues that dynamic discrete choice models are generally under-identified and that it is typically not possible to identify discount parameters. Magnac and Thesmar (2002) derive specific conditions for identification of the exponential discounting parameter. Building on these conditions, Fang and Wang (2015) develop specific exclusion restrictions which allow researchers to identify the 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 a natural instrument, the length of job protection, that exploits

benefits in the first six months.

the same intuition as in Fang and Wang (2015) although it does not change exogenous transition probabilities, but restricts future choices. Remember that women, who were employed before giving birth to a child, were protected from dismissal for a certain period of time with the length of this period varying through the different policy regime (see table 2). The likelihood of being able to work does not directly enter the utility function (see section (2.2)). A comparison of two groups of individuals which differ only by the job protection regime their child is born in, can then help to identify how individuals value future expected utilities.

Assume that one group has job protection for 3 years after childbirth, while another group is protected for only 1.5 years. Although their probabilities of finding a job in 2 years time differ due to the different regimes, their choice-specific utility half a year after childbirth does not differ on average.[13] Therefore, if the choice behavior differs between the two groups within the first 1.5 years, it can only be because of their different futures. Stated differently, if we do not observe differences in their choice behavior within that time, despite their difference futures, individuals must be fully myopic.

Comparing groups of individuals that only differ in their employment probability in the very next period,[14] informs us about the one-period ahead discount factor, which in our model is given by $\beta\delta$. Comparing groups that differ in employment rates further in the future,[15] can then help to identify the discounting necessary in addition to the one-period ahead discount factor. Such a comparison can then identify δ when we already know $\beta\delta$.

For identification purposes it is important to point out that Schoenberg and Ludsteck (2014) argue that the changes in the duration of the job protection for mothers were unexpected. This allowed them to evaluate the causal employment effects of these reforms in a reduced form setting. Besides this exogeneity, it is also important that the policy was salient to the individuals. Figure 1 shows that women behaved differently across different regimes, which is a strong indicator of the salience of the policy.

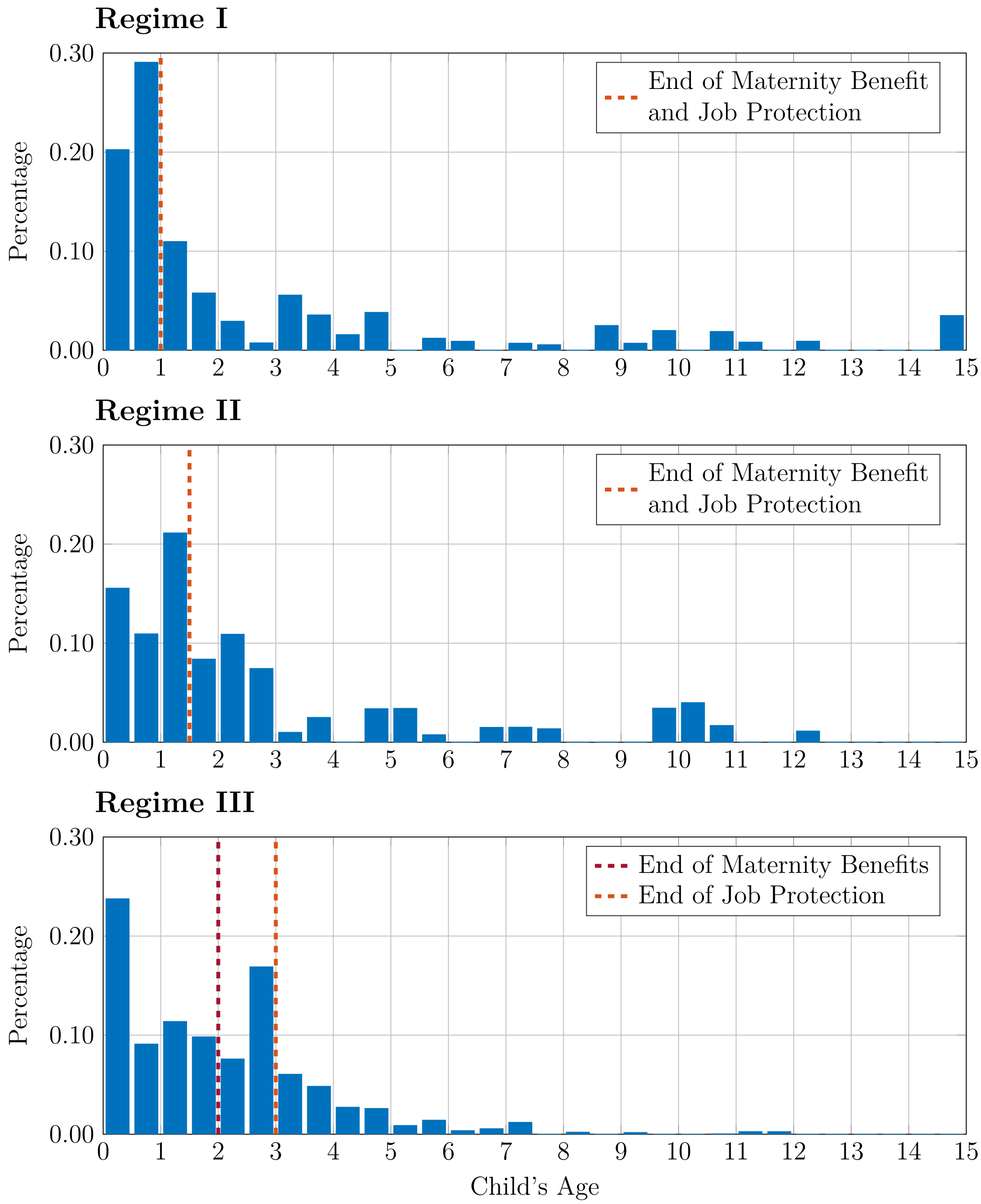
Figure (1) shows the distribution of career breaks of mothers for different lengths of job protection. Most mothers return just in time to not lose their guarantee to be able to return to their former employer. That the behavior is strongly affected by the length of job protection becomes apparent when we compare the fraction of mothers who return to the labor market

[13] We assume that the distribution of observed and unobserved heterogeneity across these groups is equal.

[14] For instance, compare groups in *regime I* and *regime II* half a year after childbirth.

[15] For instance, compare in *regime II* and *regime III* half a year after childbirth.

Figure 1: 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 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 influence 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 Illustration

We illustrate our identification argument in a three period model with two groups of individuals. Assume two groups, A and B, only differ with respect to the length of job protection, i.e. the time within they are legally certain to be employed by their previous employer. Group A benefits from three periods of job protection after childbirth, while group B only has two periods. We denote job protection in period t by jb_t, no job protection by $\overline{jb_t}$.[16]

Consider the decision in the second period of individuals who have not started to work after having their last child. Their respective life-time utilities are given by

$$\begin{aligned} U_2^A =& v(l_2;\Omega_2) + \beta\delta\mathbb{E}\max\left(l_2, \Omega_2, jp_3\right) \\ U_2^B =& v(l_2;\Omega_2) + \beta\delta\mathbb{E}\max\left(l_2, \Omega_2, \overline{jp_3}\right) \end{aligned} \tag{13}$$

Note that there is no difference in the choice specific utility functions between groups, only in the expected future utilities. Because our "instrument"[17], the length of job protection, does not directly influence the flow utilities, we follow the statement of Fang and Wang (2015) that the utility function is identified in such a model.[18] In combination with the assumption about the distribution of the choice-specific error terms, the expectation of the maximum utility in period 3 is also identified, since it is only a function of the flow utilities. Differences in the life-time utilities U_2^A and U_2^B correspond to the choice probabilities of the two groups. Because all parts of equation (13) besides the product $\beta\delta$ are directly identified, these choice probabilities identify the product $\beta\delta$.

As a thought experiment, assume individuals are fully myopic, then the product of $\beta\delta$

[16]Note that the state-space of the model presented in section 2 includes only the past status of job protection, because last period's employment status and the policy regime jointly determine the current period's job protection status. Here we explicitly indicate the potential job protection status of future periods.

[17]We use the term instrument only in quotation marks, since it does not directly resemble the classic interpretation of an instrument in a regression framework.

[18]Another way of looking at the identification of the utility function is that its parameters should be identified by using only a single policy regime. The policy regime should have no influence on the preferences on consumption and leisure and thus the "instrument" is orthogonal to these preferences.

must be zero. Equation (13) then simplifies to

$$\begin{aligned} U_2^A =& v(l_2; \Omega_2) \\ U_2^B =& v(l_2; \Omega_2) \end{aligned} \tag{14}$$

and thus we would expect that the choice probabilities of the two groups would not differ. On the other hand, if they do differ, they inform us about how they value the one period ahead future.

Having identified the product of $\beta\delta$, we still need a condition to identify one of the two parameters separately. This can be accomplished by looking at the choice behavior in the very first period:

$$\begin{aligned} U_1^A =& v(l_1; \Omega_1) + \beta\delta\mathbb{E}\left[\ln\left(\sum_{l_2} \exp(v(l_2, \Omega_2) + \delta\,\mathbb{E}\max(l_2, \Omega_2, jp_3))\right)\middle| l_1, \Omega_1\right] \\ U_1^B =& v(l_1; \Omega_1) + \beta\delta\mathbb{E}\left[\ln\left(\sum_{l_2} \exp(v(l_2, \Omega_2) + \delta\,\mathbb{E}\max(l_2, \Omega_2, \overline{jp}_3))\right)\middle| l_1, \Omega_1\right] \end{aligned} \tag{15}$$

Again, both groups have exactly the same choice-specific utility in the first two periods and only differ with respect to their job offer probability in the third period. As before, flow utilities and the expected maximum utility in the third period are identified; and from equation (13) also the product $\beta\delta$. Therefore, δ is the only unknown parameter in equation (15) and can thus be identified by the difference of choice probabilities of the two groups, A and B.

Comparing the differences in choice probabilities in period 1 to the differences in period 2 identifies δ, while the differences in choice probabilities in period 2 alone identify the product $\beta\delta$. It is important to note that in a model with more than three periods, it is always possible to construct similar equations to equations (13) and (15) with different polynomials of the discount factors. It should therefore also be possible to identify these parameters within longer time horizons.

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 arrival of children, childcare costs and the job destruction rate.[19]

In a second step, we use the method of simulated moments[20] to estimate the time-preference parameters, the parameters of the flow utility function, of the wage process and the job offer probability. The method of simulated moments is based on minimizing the distance between moments of the simulated and the observed data. Since we are estimating a dynamic discrete choice model, the objective function is a step-function. Small changes in a parameter of our model result in changes in discrete outcomes which lead to discrete changes of the objective function. Gradient-based optimization algorithms are therefore not appropriate. By contrast, global non-gradient based search algorithm are computationally unfeasible: We estimate over 30 parameters in a model in which 85 time periods are solved in each iteration. Therefore we first approximate the global optimum using a surrogate-based optimization procedure, then complement this by a local optimization procedure in the region of the approximate solution. We thus implement this second step as an iterative procedure.

For a given set of parameter values, 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 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.

We use a global optimizer based on a surrogate function to find the approximate parameter values that minimize the distance between simulated and observed data moments. To do this, we first evaluate the objective function at a limited number, B, of points across the parameter space. Recall that the objective function is the distance between the moments of the observed data and the moments of the simulated data. We thus simulate the model at B different sets of parameter values and generate simulated moments M_k^S. To calculate the objective function $g(\Theta^b)$ we then compare the distance of the simulated moments to the moments from the observed data M_k^O for every set of parameter values Θ^b.

[19] 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.

[20] See Smith (1990), Gourieroux et al. (1993) and Gallant and Tauchen (1996).

$$g(\Theta^b) = \sum_{k=1}^{K} \left[\frac{\left(M_k^O - M_k^S(\Theta^b)\right)^2}{\mathbb{V}\mathrm{ar}(M_k^O)} \right] \tag{16}$$

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^b, and $\mathbb{V}\mathrm{ar}(M_k^0)$ the variance of the same observed moment. An overview of the moments we use for estimation is provided in Appendix A. Note that we do not use the optimal weighting matrix due to its poor small-sample properties (Altonji and Segal, 1996). Note that each $b = 1, ..., B$ of these evaluations requires solving the model according to the procedure outlined in step 2.1.

We then specify a simple (surrogate) function relating the evaluations of the objective function to the different parameter values.

$$g(\Theta^b) = \alpha_0 + \sum_{p}^{P} \alpha_{1,p} \ \theta_p^b + \sum_{p}^{P} \sum_{q \neq p}^{Q} \alpha_{2,p,q} \ \theta_p^b \ \theta_q^b \tag{17}$$

for all parameters $\theta_p \in \{\Theta\}$ and where α_1 and α_2 are thus vectors of size P and PxP, respectively. This gives a set of B equations which can be solved for α. This surrogate function can be interpreted as an interpolation of the objective function. We then minimize the surrogate function in the parameter space for all P parameters. This provides us with an approximation of the global optimum of the objective function. We therefore complement this step with a local optimizing routine.

The result of the previous step is a set of parameter values that optimizes the surrogate function. We start from this value in a search algorithm to find a precise local optimum based on direct evaluations of the objective function given by (16). To minimize the distance we use a pattern-search algorithm, i.e. testing the optimal direction in all P dimensions and making successively smaller steps. The solution is the minimum of the objective function, i.e. the set of parameter values which most closely align our simulated model moments to the observed data moments.

Table 3: Utility function

γ_c	1.880638	(0.044481)
$\gamma_{PT_{low}}$	1.803703	(33.813883)
$\gamma_{FT_{low}}$	2.033719	(27.968465)
$\gamma_{PT_{high}}$	1.675667	(2.055640)
$\gamma_{FT_{high}}$	1.994031	(1.656444)
$\gamma_{PT,AC0}$	−2.088371	(0.128414)
$\gamma_{PT,AC1}$	0.657297	(0.125392)
$\gamma_{PT,AC2}$	0.087503	(0.031092)
$\gamma_{FT,AC0}$	−2.264446	(0.121315)
$\gamma_{FT,AC1}$	0.464901	(0.130525)
$\gamma_{FT,AC2}$	0.100000	(0.031320)

Table 4: Labor market

Parameters of the wage function		
$\gamma_{w,low}$	0.313178	(0.168864)
$\gamma_{w,high}$	0.405701	(0.249426)
$\gamma_{w,e}$	0.980565	(0.173101)
Parameters of labor market frictions		
γ_{JO_low}	0.129044	(0.003975)
γ_{JO_high}	0.141328	(0.014425)
Measurement Error		
σ_{xi}	0.073118	(0.012987)

6 Results

In this section we present the parameter estimates of the model. Overall, the estimation results show the expected picture. In more detail, we find that women are risk averse (the estimated coefficient of relative risk aversion amounts to about 1.9, table 3). Moreover preferences for consumption and leisure time are non-separable and they suggest that ceteris paribus women have positive preferences for leisure time. As expected the presence of young children increase the preferences for leisure time.

Further we find the expected differences by education in wages and the job offer rates (table 4).

Finally we turn to the estimated coefficient of the time preferences (table 5): we find evidence for time inconsistent behavior. In more detail, we show that the β coefficient is

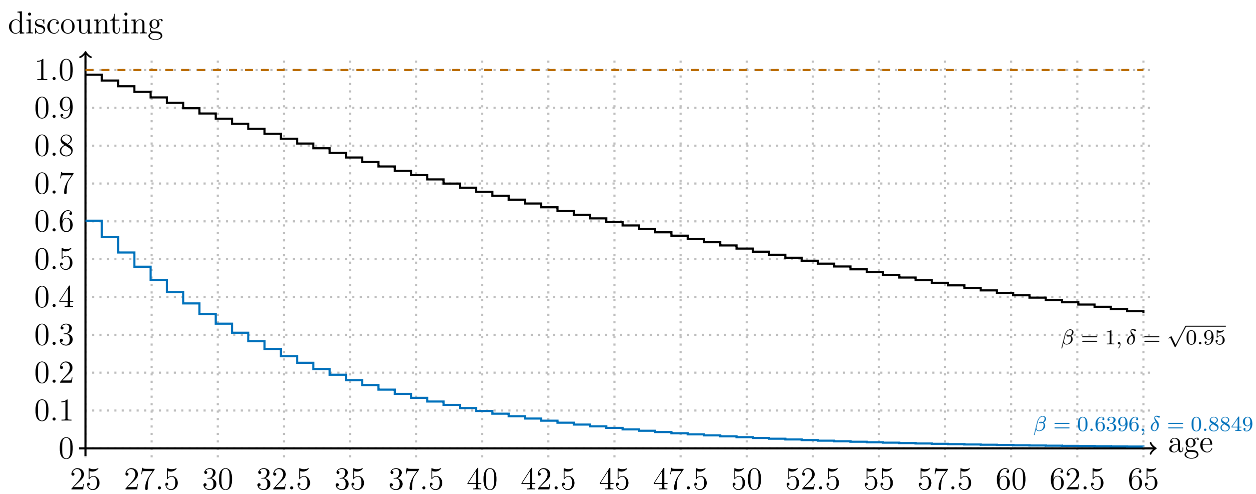
Table 5: Discounting

β	0.639588	(0.134307)
δ	0.884862	(0.067039)

about 0.6 and significantly different from unity. This suggest that women are present-biased. The point estimate for the long-run time discount factor δ is about 0.89. These values imply considerably larger discounting of future utility streams than standard models assume.

Figure 2 shows the weight in decision-making that future utility streams have for a 25-year old women according to our model estimates (lower line) and according to the standard assumption of exponential discounting with a yearly factor of 0.95. Utility streams accruing after the age of 50 enter at only a fraction of the weight compared to immediate utility (under 5%). This is consistent with Thiemann (2015) who reports no reaction to changes in the pension system that reduced incentives for mothers to return swiftly to employment after childbirth. He concludes that the finding are not in line with "perfectly rational" mothers but that high discount rates could explain his findings. While figure 2 confirms that our estimates are consistent with this finding, the standard discount factors would suggest pension benefits (earnings at age 60) should be discounted by less than 70% compared to current earnings.

Figure 2: Discounting over the life-span



6.1 In sample fit

TO BE COMPLETED

7 Policy simulations

TO BE COMPLETED

8 Conclusion

In this paper we test for time-inconsistent behavior. We exploit exogenous changes in the duration of job protection for mothers which was increased in several steps from one to three years 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 *directly* affect mothers' flow utility, i.e. it influences choices only by guaranteeing *future* employment opportunities.

Estimates based on German panel data show strong evidence for time inconsistent preferences. The large estimated present bias leads us to reject exponential discounting, an assumption common to most models of dynamic discrete choice. Our results are highly relevant for the correct specification of dynamic models used to evaluate the labor supply effects of tax policies, child-care support or pension benefits.

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.

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.

Augenblick, N., M. Niederle, and C. Sprenger (2015): "Working Over Time: Dynamic Inconsitency in Real Effort Tasks," *Quarterly Journal of Economics*, 130, 1067–1115.

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," .

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.

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.

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.

Martinez, S.-k., S. Meier, and C. Sprenger (2017): "Procrastination in the Field: Evidence from Tax Filing," .

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.

Thiemann, A. (2015): "Pension wealth and maternal employment: Evidence from a reform of the German child care pension benefit," .

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: Overview of Moments

Table 9: Transition Rate into Employment

	Low Education				High Education			
Moment	Data	Sim	SE Data	SE Diff	Data	Sim	SE Data	SE Diff
Rate	0.0950	0.1185	0.0030	7.7814	0.0987	0.1303	0.0083	3.8010

Table 10: Log Wages at Entrance in Working Life

	Low Education				High Education			
Moment	Data	Sim	SE Data	SE Diff	Data	Sim	SE Data	SE Diff
Mean	1.9178	1.8919	0.0164	1.5783	2.4855	2.3890	0.0399	2.4186
Variance	0.1860	0.0151	0.0100	17.1182	0.1231	0.0250	0.0277	3.5474

Table 11: Log wage of Full-Time Workers

	Low Education				High Education			
Moment	Data	Sim	SE Data	SE Diff	Data	Sim	SE Data	SE Diff
Mean	2.3140	2.3229	0.0094	0.9479	2.6189	2.6585	0.0262	1.5081

Table 6: Employment Rates by Education

Moment	Low Education				High Education			
	Data	Sim	SE Data	SE Diff	Data	Sim	SE Data	SE Diff
Full-time Employment Rates								
All	0.3440	0.3085	0.0080	4.4597	0.3994	0.3575	0.0215	1.9446
No children	0.6407	0.5632	0.0103	7.4900	0.7498	0.6635	0.0246	3.5072
All mothers	0.1527	0.1517	0.0083	0.1202	0.1495	0.1623	0.0203	0.6292
Youngest child [0, 3[	0.0525	0.0421	0.0056	1.8671	0.0567	0.0444	0.0159	0.7701
Youngest child [3, 6[	0.1217	0.1088	0.0100	1.2862	0.1143	0.1214	0.0242	0.2939
Youngest child [6, 11[	0.1539	0.1801	0.0111	2.3505	0.1833	0.1847	0.0327	0.0430
Part-time Employment Rates								
All	0.2416	0.2211	0.0068	3.0284	0.2324	0.2215	0.0154	0.7126
No children	0.1316	0.0942	0.0070	5.3505	0.1034	0.0665	0.0155	2.3901
All mothers	0.3125	0.2992	0.0096	1.3778	0.3244	0.3203	0.0235	0.1747
Youngest child [0, 3[	0.1049	0.0928	0.0068	1.7925	0.1311	0.1050	0.0206	1.2681
Youngest child [3, 6[	0.3315	0.2591	0.0138	5.2570	0.3514	0.2687	0.0361	2.2900
Youngest child [6, 11[	0.4118	0.4878	0.0153	4.9852	0.4317	0.4942	0.0382	1.6384

Table 7: Employment Rates by Regime

Age Youngest Child	Regime I				Regime II				Regime III			
	Data	Sim	SE Data	SE Diff	Data	Sim	SE Data	SE Diff	Data	Sim	SE Data	SE Diff
Full-time Employment Rates												
0-6 months	0.0570	0.0292	0.0171	1.6306	0.0075	0.0196	0.0074	1.6318	0.0228	0.0182	0.0048	0.9568
6-12 months	0.0670	0.0312	0.0162	2.2080	0.0286	0.0323	0.0145	0.2586	0.0241	0.0233	0.0053	0.1554
12-18 months	0.0919	0.0225	0.0225	3.0794	0.0976	0.0496	0.0215	2.2353	0.0280	0.0311	0.0055	0.5609
Part-time Employment Rates												
0-6 months	0.0104	0.0295	0.0071	2.6971	0.0522	0.0322	0.0245	0.8190	0.0342	0.0399	0.0059	0.9688
6-12 months	0.0491	0.0512	0.0155	0.1343	0.0286	0.0526	0.0179	1.3394	0.0649	0.0592	0.0080	0.7208
12-18 months	0.0919	0.0284	0.0219	2.8957	0.0732	0.0684	0.0188	0.2545	0.0951	0.0856	0.0100	0.9566

Table 8: Log Wage Regressions on Accumulated Experience and Lagged Wages

	Low Education				High Education			
Moment	Data	Sim	SE Data	SE Diff	Data	Sim	SE Data	SE Diff
Constant	0.4464	0.4036	0.0415	1.0325	0.4343	0.6003	0.1152	1.4401
$\ln(w_{t-1})$	0.7846	0.7939	0.0121	0.7650	0.8534	0.7479	0.0290	3.6422
Log accumulated working years	0.1261	0.0631	0.0935	0.6742	-0.0788	0.0731	0.2492	0.6096
Lagged log accumulated working years	-0.1058	-0.0138	0.0832	1.1070	0.0646	-0.0295	0.2226	0.4226
Variance of residuals	0.0484	0.0094	0.0719	0.5414	0.0515	0.0090	0.3056	0.1389