Freie Universität Berlin

Essays on Time Preferences and Expectations in Dynamic Decision Making

INAUGURAL-DISSERTATION

zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft
des Fachbereichs Wirtschaftswissenschaft
der Freien Universität Berlin

vorgelegt von
Ulrich Schneider (M.Sc.)
geboren in Mainz

Berlin, Februar 2019

Gedruckt mit der Genehmigung des Fachbereichs Wirtschaftwissenschaft der Freien Universität Berlin.

Dekan:
Prof. Dr. Dr. Andreas Löffler

Erstgutachter:
Prof. Dr. Peter Haan, Freie Universität Berlin

Zweitgutachter:
Prof. Sir Richard Blundell, University College London

Datum der Disputation:
25. Januar 2019

Erklärung über Zusammenarbeit mit Koautoren und Vorveröffentlichungen

Kapitel 1
Bislang unveröffentlicht.

Kapitel 2
In Zusammenarbeit mit Peter Haan und Luke Haywood (Eigenanteil: 33.3%). Bislang unveröffentlicht.

Kapitel 3
Bislang unveröffentlicht.

Acknowledgments

This dissertation could not have been written without the great support of numerous people and institutions.
I am extremely grateful for the help, support, and constant encouragement that I received from Peter Haan, my first supervisor and co-author. Besides giving me plenty of room and patience to discuss and develop my own research ideas, he regularly provided great guidance in matters of structural modeling, estimation techniques, and academic research more generally. Without his support, both of my research stays would have been much more difficult to accomplish. I also thank my second supervisor, Richard Blundell, for the many fruitful discussions, his outstanding feedback, and his support of my research. He made my stay at the Institute for Fiscal Studies possible, which resulted in not just many helpful scientific exchanges, but also in new friendships. In addition to my two supervisors, I thank Luke Haywood, my co-author, for his great and enjoyable collaboration.
Furthermore, I am grateful to Alexander Bick for organizing my research stay at the Department of Economics at Arizona State University. He dedicated a lot of his time to discussing the early stages of various parts of this dissertation and provided crucial feedback, comments, and encouragement. A special group of people made this research stay even more productive and enjoyable, these include, among many others, Ambika Athreya, Barbara Boggiano, Luis Fernandez Intriago, Diana MacDonald, Samuel Marshall, Alberto Rivera, and Alberto Vindas Q.
Throughout my PhD, I have worked at the German Institute for Economic Research (DIW Berlin) and was a member of its Graduate Center. DIW Berlin provides one of the greatest work atmospheres I have ever experienced and I thank all the wonderful people, who work there. I am especially indebted to the members of the Public Economic Department for the many productive discussions, for their guidance, and for their support over the years. Among them Hermann Buslei, Johannes Geyer, Patricia Gallego Granados, Daniel Kemptner, Kai-Uwe Müller, Michael Neumann, Renke Schmacker, Andreas Thiemann, and Aline Zucco.
A special thanks to Songül Tolan, Verena Grass, Holger Lüthen, Annekatrin Lüdecke,

and Ronny Freier, all my office mates at various points in time: all made working even more pleasurable and interesting. A very special thanks to Georg Camehl, Sarah Dahmann, Caterina Forti Grazzini, Kristina Krell, Stefan Seifert, Rainer Siegers, Marica Valente, and Christian Zankiewicz for keeping up with me over the years, having entertaining lunch breaks, and most of all their friendships. Another special thanks to Cormac O'Dea, for his outstanding support through the later part of my PhD.

On a more technical note, I gratefully acknowledge financial support from German Research Foundation (DFG), grants CRC TRR-190 and SPP-1764, and Network of Interdisciplinary Social Policy Research funded by the Federal Ministry of Labor and Social Affairs. I also thank the HPC Service of ZEDAT, Freie Universität Berlin, for computing time. In addition, I am indebted to Nicole Haase and Juliane Metzner for their administrative support, and especially to Adam Lederer for carefully revising my use of the English language and his general mood lifting personality.

Last, but most importantly, I am deeply indebted to the joy of my life, Katha. She unselfishly supported me through all the ups and downs of pursuing a PhD and has always had my back. You made this possible.

Table of Contents

General Introduction

When designing optimal policies, it is necessary to understand human behavior. Without theory and evidence on how economic agents arrive at their decisions, it is impossible to reliably predict outcomes of any regulation or law. For most issues in public economics, it is key to understand the decision process of individual labor supply, as income tax revenues, social security, and pension spending, among many others are driven by earnings from employment. For instance, to understand if a particular tax reform will increase tax revenues, it is important to know how individuals react to changes in their net wages. Theory provides the foundation for understanding what effects drive results. The income effect tells us that reducing net wages lowers overall income, ultimately reducing both the consumption of goods and of leisure, resulting in an increase in labor supply. On the other hand, the substitution effect tells us that a lower net wage reduces the opportunity costs of leisure, resulting in a lower price relative to consumption, causing a reduction of labor supply. Therefore, theory tells us that tax increases can have an ambiguous effect on labor supply and tax revenues, a result with great meaning. It is crucial to understand that just because a rise in tax rates once resulted in an increase of labor supply, it will not always do.

When developing theories of economic relationship, economists have a long tradition of relying on the concept of utility. Based on the mathematical formulation of a utility function, all characteristics influencing a particular decision are summarized to a single value that represents preferences. Given options to choose from, an individual will select the option with the highest value of the utility function. In labor economics, as the trade-off between leisure and consumption is fundamental, preferences are often modeled as a function of both. It is astonishing how many insights can be gained by relying only on some weak assumptions of the utility function. In addition to the aforementioned income and substitution effect of changes in net wages, another early example of the usefulness of the utility function is provided by Mirrlees (1971). He derives fundamental insights into optimal taxation, which was the first crucial step in the field of optimal taxation.

Although theory builds the bedrock of understanding economic relationships, it can often lead to ambiguous results and is almost always silent on the precise size of the impact a certain mechanism has on behavior. For instance, despite knowing how income and substitution effects change behavior, it is difficult to determine which effect dominates in a particular case. In addition, Mirrlees notes that “the shape of the optimum earned-income tax schedule is rather sensitive [...] to the income-leisure preferences postulated,” (Mirrlees, 1971, p. 207) and, thus, calls for empirical evidence. Therefore, a complete economic analysis should also include the estimation for central parameters of a theory to increase the precision of a policy recommendation. Furthermore, empirical analyses can help to verify, advance, and show the bounds of a given theory. To be able to do so, the first step is to develop appropriate identification strategies that help researchers to evaluate the usefulness of a given theory to explain certain behavior.

The three chapters of this dissertation follow these lines by providing empirical insights on theoretical models to guide researchers and policy-makers to improved policy recommendations. Building on standard labor supply models, I focus on how agents account for future utilities of consumption and leisure in their current decision process. While Chapter 1 provides a strategy to estimate a standard exponential discount factor, Chapters 2 and 3 analyze elements of dynamic choice, which were previously largely ignored in the literature on life-cycle labor supply. To verify their value, all discussed models nest the theories, most dominant in the literature, while identification strategies are developed to help researchers distinguish between the different models. Chapters 2 and 3 empirically test the newly introduced elements of dynamic choice with data on observed female labor supply choices. They clearly reject the nested dominant theories of dynamic choice relying on the developed identification strategies.

With these results, the dissertation shines a spotlight on how researchers can improve their modelling and estimation strategies when economic agents have to make repeated choices over a long period. Since the vast majority of issues that economists concern themselves with are dynamic in nature, the results potentially have implications for areas of economics beyond life-cycle labor supply. In general, decisions not only influence the current state of the world, but also its future. Going to college might impact future wages, getting married might impact having children in the future, buying stocks might impact future budget constraints, smoking and exercising might impact future health, working full-time might help to advance careers, forgoing consumption and instead saving might impact income in retirement. Thus, it is central for economists to understand how individuals account for the future when making decisions.

Regarding the dynamics of female labor supply, modelling approaches have constantly improved since the first attempt to explain life-cycle patterns by Jacob Mincer in 1962

(Mincer, 1962).[1] Following his theoretical considerations, Heckman & Macurdy (1980) are the first to analyze female labor supply over the life-cycle based on a structural model. Building on the framework of per-period utility, they assume, in addition to exponential discounting, perfect foresight about preferences and wages. Although most of their results are as economic theory would predict, they find huge Frisch labor supply elasticities. Complementing their previous work, Heckman & Macurdy (1982) introduce a CRRA utility function, which is less restrictive and caused lower estimates of the Frisch elasticities, although still remaining moderately large. Responding to a critique by Cogan (1981), Kimmel & Kniesner (1998) extend Heckman & Macurdy (1980, 1982) by including fix costs of work. They find a low Frisch elasticity for women in employment, but a similar value as Heckman & Macurdy (1982) for the women out of employment. Further advances were made by Altug & Miller (1998), who introduced on-the-job human capital, state dependence utility, and aggregate shocks into their model.

The first to analyze life-cycle labor supply within a model of dynamic discrete choice are Eckstein & Wolpin (1989), who ignore the choice of hours and only looked at a working/non-working decisions. Their findings are in line with theory, as, for example, children and higher incomes of the husbands reduce women's taste for work. Van der Klaauw (1996) extends their work by endogenizing marriage and divorce decisions in addition to the employment decision. Furthermore, Francesconi (2002) start distinguishing between part-time and full-time employment decisions, while also including fertility as a choice and while modeling marriages as exogenous processes. The work of Keane & Wolpin (2010) aims to endogenized most relevant processes in addition to savings by including labor supply, marriages, fertility, and education into the choice set. This allows them to address a variety of issues, including, for example, how the marriage market has differing effects on the labor supply of white and black women. More recent studies applying dynamic discrete choice models include Blundell et al. (2016), who analysis human capital processes and welfare reforms while endogenizing the savings decision, and Adda et al. (2017), who concentrate on occupational choices. Although the analysis of life-cycle labor supply improved continually as research concentrated on the modeling of utility functions and endogenizing various processes, essential aspects of dynamic behavior were ignored. The vast majority of the literature assumes that individuals are exponentially discounting future utilities and possess rational expectations about their future, with only very few exceptions. These assumptions stand in stark contrast to findings from psychology and the behavioral economics literature. Regarding time preferences, the seminal work by Laibson (1997), who introduces time

[1]This brief overview of estimating female life-cycle labor supply is partially based on Keane et al. (2011).

preference in the form of hyperbolic discounting, is followed by a growing literature. Their findings demonstrate that individuals strongly deviate from the time-consistent exponential framework in various economic domains and that the assumption of hyperbolic discounting describes behavior more accurately. For instances, DellaVigna & Malmendier (2006) show that individuals behave time-inconsistent in their physical exercise patterns and Martinez et al. (2017) show that tax-filers tend to have a present bias. Chapter 2 builds on these findings and designs a dynamic structural life-cycle model nesting both exponential and (quasi-)hyperbolic time preferences. The results of the estimations reject the standard assumption of exponential discounting in this setting.

Chapter 3 concentrates on possible bias in the expectations of individuals when incorporating future utilities in current decisions. Since the seminal work of Tversky & Kahneman (1974), which introduces a theory for their common finding that individuals exhibit systematic biases when acting under uncertainty, the literature on social psychology and organizational behavior has intensively analyzed biased expectations. An introduction into the literature's link to economic questions is provided by Malmendier & Taylor (2015). Some examples include Svenson (1981), who finds that over 80% of individuals expect themselves to be driving saver than the average, and Ben-David et al. (2013), who find that CFOs have biased expectations about the development of the S&P500. Complementing this literature, the last chapter provides evidence that mothers overestimate their future employment opportunities. It does so by developing a strategy to identify expectations within a dynamic discrete choice model of female labor supply. Estimation results show that the estimated expectations differ significantly from rational expectations.

The negligence of these behavioral elements in the literature on female labor supply might be partly due to the issue of their identification. As Magnac & Thesmar (2002) show, neither time preferences nor expectations are generally identified from choice data in dynamic discrete choice models. Further, developing their results, Fang & Wang (2015) derive exclusion restrictions for which it is possible to identify not only exponential discounting parameters, but also hyperbolic discounting parameters. Chapter 1 complements this work by providing a different set of exclusion restrictions for models, in which economic agents are potentially restricted in their choice set. While this first chapter only focuses on the point identification of the exponential discount factor, chapter 2 expands this framework to identify (quasi-)hyperbolic discounting parameters. Chapter 3 breaks new grounds by providing an identification strategy for expectations of future employment offers based on observed employment choices.

One possible path for the identification of time preferences and expectations is to design survey questions that directly ask individuals about their preferences. An important drawback of this method lies in the typical short amount of time for each survey question and, thus, responses to more theoretical questions on preferences might not come with the necessary thoroughness and carefulness. Furthermore, a vast majority of the population seems to have trouble understanding the concept of probability (Garfield & Ahlgren, 1988), which is often required when measuring expectations. Thus, results from these survey questions should be interpreted with caution. The empirical chapters of this dissertation use only survey questions about preferences and expectations as suggestive evidence and possible verification of final results. However, the identification and estimation strategies of the proposed models rely purely on observed employment choices over time. Combining the stated and revealed preferences approach builds stronger evidence for the hypothesis that individuals deviate from rational expectations and exponential discounting.

The last two chapters of this dissertation apply the discussed models in an empirical analysis of female labor supply. The employment behavior of women is of particular interest in public economics. In most countries, the majority of benefit systems, like social security and pensions, are financed via a pay-as-you go scheme. As these systems, in the presence of the current demographic change, face increasing pressure regarding sustainability, increasing female labor supply can provide important relief. Within the OECD, the average female employment rate is, at 60.8%, around 15 percentage points lower than the male employment rate of 75.9% (OECD, 2018). Besides this difference in the extensive margin, there is also a gender difference in the intensive margin. When employed, women choose part-time employment (25.5%) significantly more often than men (9.2%). Furthermore, females retire one and a half years earlier than men, but spend an average of four and a half years longer in retirement (OECD, 2017), resulting in a longer period of receiving pension benefits compared to men.

To increase women's contributions to the tax and transfer system, a natural starting point is to analyze the role of childbirths, since the majority of mothers interrupt their working careers due to children. Chapters 2 and 3 contribute new theoretical and empirical evidence on the underlying causes of extended career breaks. While the first of these two chapters focuses on time-inconsistent behavior, which can cause women to repeatedly postpone their return to employment, the latter focuses on being too optimistic about future employment opportunities. When women overestimate their chances to re-enter employment in the future, they might not return to their previous job within maternity leave, which in most countries offers employment protection. After maternity leave ends, women face harsher labor market conditions than expected, prolonging child-related career breaks.

The rest of this section briefly summarizes the main contributions of each of the three chapters. The first chapter contributes a new set of exclusion restrictions to recover the exponential discount factor from observed choices within dynamic discrete choice models. Identification is achieved by exploiting exogenous changes in probabilistic choice restrictions. These restrictions are given when economic agents are, with a certain probability, not able to choose from all alternatives in the choice set. They can be found in various fields of economics: in the context of labor economics, individuals might not be able to choose employment when not receiving a job offer; in industrial organization, merges might not be possible due to decisions by competition regulators or political forces; in environmental economics, firms' choices of emissions might be subject to environmental regulations. Besides discussing the identification of the potentially unobservable restriction probabilities, the chapter derives a formula for identifying the exponential discount factor, relying on exogenous variation in these probabilities. Notably, this formula does not depend on the functional form assumptions of the utility (or profit) function, but only depends on observed choice and transition probabilities, as well as the change in the restriction probabilities.

Chapter 2 extends the first chapter by providing an identification strategy for time preferences of the form of hyperbolic discounting. It derives formulas for a three-period model, arguing that, although no analytic expression for the parameters can be derived for a model with more periods, identification can be achieved using the same exclusion restriction. To estimate the time preferences, the chapter exploits exogenous changes in the length of employment protection for mothers who were employed before having their child. Employment protection provides insurance against labor market frictions, therefore changing the restriction probability of having to choose non-employment when currently out of labor. A full life-cycle model of female labor supply is specified and estimated using data from the German Socio-Economic Panel. Parameter estimates are close to values typically found in the hyperbolic discounting literature and suggest that exponential discounting does not properly describe female life-cycle behavior.

In contrast to the first two chapters, chapter 3 focuses not on time preferences, but on expectations about future employment possibilities. The chapter consists of three major parts. First, a life-cycle model of female labor supply, human capital accumulation, and labor market frictions is developed. Second, a strategy to identify future employment expectations is derived and, third, the estimated model is used to quantify the career costs of biased expectations. The identification approach exploits the impact of expectations on the decision process at the end of maternity leave. The change from an employment guarantee to a situation in which individuals have to rely on job offers to leave non-employment creates a discontinuity in the future expected value of non-employment that varies with job offer expectations. To separately identify expectations, preferences, and real job offer rates, multiple maternity leave reforms, each changing the length of the employment protection, are exploited.

The model nests rational expectations, but estimations strongly indicate that women overestimate their future possibilities to return to employment. It is shown that biased beliefs prolong child-related career breaks by over 10% compared to rational expectations, and significantly increase the number of mothers never returning after having children. Labor market earnings decrease by over 12%, but individual consumption only by around 3%. This difference is due to most husbands continuing working full-time and a tax and transfer system that heavily taxes secondary earners.

Chapter 1

Identification of Time Preferences in Dynamic Discrete Choice Models – Exploiting Choice Restrictions

1.1 Introduction

Dynamic discrete choice models are applied to estimate behavior of economic agents and derive counterfactual policy analysis in various fields, including labor economics, industrial organization, environmental economics, and marketing. In these models, the discount factor is crucial for the reactions to future events as it affects how much weight agents place on their expected future when making decisions. In empirical applications, the researcher is often forced to set the discounting parameters, as they cannot generally be identified from choice probabilities and transition rates alone (see Rust (1994) and Magnac & Thesmar (2002)).

This chapter provides a new class of instruments – changes in probabilistic choice restriction – to identify the discount factor. A probabilistic choice restriction is given when economic agents are, with a certain probability, not able to choose from all alternatives of the general choice set. There are many examples of these types of restrictions. In labor economics, for example, individuals often have the choice between working a positive number of hours or not working at all. When including the demand for labor, individuals might be restricted in their choices: transitions from non-working to working are only possible if a job offer from a firm is received. Thus, an agent's choice set might be reduced to only non-employed if no job offer is received. In industrial organization, choices can be restricted when mergers are subject to approval. A firm that is considering a merger with another firm is restricted based on the probability that approval might not be granted. In marketing, the availability of products might be probabilistic restricted as products are no longer sold or are out of stock. Finally, in environmental economics, the amount of emissions a firm can produce in a given period, might be restricted with uncertainty stemming from changes in environmental regulations.

Exogenous variation in these probabilistic restrictions for a limited time can identify the discount factor within a dynamic discrete choice model. The advantage of using choice restrictions for identification is that, although, they do not enter the utility function directly, they affect the future expected utility. Therefore, changes in these restrictions exclusively change the future expected utility, while the utility of the current period stays unaffected. This allows attributing the intensity of change in current behavior due to changes in restriction probabilities to the size of the discounting parameter. Simply said, if behavior does not change even though future choice restrictions change, then economic agents are myopic. If, in contrast, behavior changes dramatically, individuals place a high value on future periods and, thus, the discount factor has to be large.

This identification strategy is comparable to Fang & Wang (2015), which exploits exogenous changes in the transition probabilities of state variables that also do not enter the utility function. Although similar in reasoning, Fang & Wang (2015) discuss a different type of variable needed for identification. Depending on the context, it might be easier to find variables that fulfill their exclusion restriction or the one presented in this chapter. Fang & Wang (2015) also concentrate on generically identifying hyperbolic discounting parameters, while this chapter focuses on point identification of the exponential discounting parameter, providing a direct formula for the parameter.

The rest of this chapter is structured as follows: Section 1.2 introduces the model framework. Section 1.3 discusses how identification can be achieved. Section 1.4 concludes.

1.2 Model

Consider a standard dynamic discrete choice model as presented, for example, in Magnac & Thesmar (2002). Time is assumed to be discrete and finite.[2] The last period is denoted by $T < \infty$. An economic agent has to choose from K alternatives in each period, however, their choice set might be restricted. The choice in period t is denoted by $d_t \in D = \{1, \ldots, K\}$ with $K \geq 2$. The agent is forward-looking and chooses the alternative with the highest expected lifetime utility. Instantaneous utility of a choice depends on the observable state in that period $x_{t,s} \in X_t = \{x_{t,1}, \ldots, x_{t,J_t}\}$ and an additional, for the econometrician unobservable, choice specific component denoted by $\varepsilon_t = \{\varepsilon_{t,1}, \ldots, \varepsilon_{t,K}\}$.[3] Note that the state set is not necessarily the same in each period. It is assumed that ε is continuously distributed over $\mathbb{R}^K$.

The order of events in each period is the following. Depending on previous period's

[2] Although this chapter concentrates on a finite horizon model, identification can also be achieved in the infinite horizon model as discussed at the end of section 1.3.

[3] Alternatively, these can be interpreted as measurement error or as the agent's optimizing errors.

state space and choice, all variables of the current state space are drawn accordingly to their distribution, including the choice specific errors. Based on these values, the agent makes a decision. However, in contrast to a standard model, the economic agents might be restricted in their choice. The following assumption about the choice restriction is made:[4]

Key Assumption 1. *The agent's choice set in a given period is restricted with probability $\pi_{\tilde{D}_{d_{t-1}}}(x_{t-1})$ to choice set $\tilde{D}_{d_{t-1}} \subset D$. With $1 - \pi_{\tilde{D}_{d_{t-1}}}(x_{t-1})$ the choice set is not restricted:*

$$d_{t-1} \in \begin{cases} \xrightarrow{1-\pi_{\tilde{D}_{d_{t-1}}}(x_{t-1})} \bullet\ D = \{1, \ldots, K\} \\ \xrightarrow[\pi_{\tilde{D}_{d_{t-1}}}(x_{t-1})]{} \bullet\ \tilde{D}_{d_{t-1}} \subset D \end{cases}$$

In the easiest set-up, the restricted choice set and the probability to be restricted is equal for all choices. Since the restricted set depends on the previous made choice, it is also possible to have as many restricted choice sets ($\tilde{D}_{d_{t-1}}$) as there are choices in the unrestricted choice set. For the rest of the chapter $\pi_{\tilde{D}_{d_{t-1}}}(\cdot)$ is referred to as the restriction probability. Note that it is necessary to assume $\pi_{\tilde{D}_{d_{t-1}}}(x_t) < 1$ for at least one element $d \in D$. After the choice is realized, a new period begins, the state space is updated and new values for ε are drawn. The following assumption is made about instantaneous utility, i.e. the direct utility of a single period:

Key Assumption 2. *The instantaneous utility for a choice d_t is given by*

$$u^\star(d_t, x_t, \boldsymbol{\varepsilon}_t) = u(d_t, x_t) + \varepsilon_{d_t,t} \tag{1.1}$$

where $u(\cdot)$ is a function of the decision and the state, and ε is the choice specific unobserved component.

With this assumption, it is possible to state the expected lifetime utility of a given period as

$$U_t(u_t, \ldots, u_T) = u^\star(d_t, x_t) + \sum_{\tau=t+1}^{T} \beta^{\tau-t}\mathbb{E}\left[u^\star(d_\tau, x_\tau)\right]. \tag{1.2}$$

The parameter β represents the discount factor. With the lifetime utility defined, the decision problem of the economic agent, which is to maximize expected lifetime

[4] Note that there are already some assumptions made. The rest of the chapter highlights the least general ones within the presented framework of a dynamic discrete choice model.

utility, can be broken down to the well-known Bellman equations (Bellman, 1957). Furthermore, key assumption 2 allows to divide the lifetime utility into a part that depends on observable parameters and a part that depends on unobservable parameters:

$$v^{\star}_{d,t}(x_t) = v_{d,t}(x_t) + \varepsilon_{d,t}, \text{ where } v(\cdot) \text{ is} \tag{1.3}$$

for t < T:

$$\begin{aligned} v_{d,t}(x_t) &= u(d_t, x_t) \\ &+ \beta \sum_{X_{t+1}} \Bigg\{ \left(1 - \pi_{\tilde{D}_{d_t}}(x_t)\right) \mathbb{E}\left[\max_{j \in D} \{v_{j,t+1}(x_{t+1}) + \varepsilon_{j,t+1}\}\right] \\ &\qquad + \pi_{\tilde{D}_{d_t}}(x_t) \mathbb{E}\left[\max_{k \in \tilde{D}_{d_t}} \{v_{k,t+1}(x_{t+1}) + \varepsilon_{k,t+1}\}\right]\Bigg\} q(x_{t+1}|d_t, x_t) \end{aligned} \tag{1.4}$$

for t = T:

$$v_{d,t}(x_t) = u(d_t, x_t) \tag{1.5}$$

Note that, in a given period, the choice set for the current period is already determined when the economic agents make their decisions. Two more assumptions are necessary to close the description of the model.

Key Assumption 3. *The state space values x_t are drawn from the distribution $q(x_t|d_{t-1}, x_{t-1})$. These draws are independent of ε for any choice.*

Key assumption 3 is also known as the conditional independence assumption (see Rust, 1987; Magnac & Thesmar, 2002). Furthermore, the distribution of ε is assumed to be known, and for ease of exposition of the identification proof, it is assumed that[5]

Key Assumption 4. *ε_t is independent and identically type-I extreme value distributed with zero-mean.*

Key assumption 4 allows to rewrite the expected maximum as a closed form solution:

$$\mathbb{E}\left[\max_{j \in D} \{v_{j,t+1}(x_{t+1}) + \varepsilon_{j,t+1}\}\right] = \ln\left(\sum_{j \in D} \exp\left(v_{j,t+1}(x_{t+1})\right)\right) \tag{1.6}$$

For the rest of the chapter, it is assumed that the researcher observes data on choices d and state variables x that allows to determining $q(x_t|d_t, x_{t-1})$, and the observed choice probabilities $\Pr(d_t|d_{t-1}, x_t, x_{t-1})$.

[5] The identification proof does not depend on the specific distribution. It is only necessary that the researcher knows the distribution.

1.3 Identification

1.3.1 Identification of Restricted Choice Probabilities

A prerequisite for the identification of the discount factor is to estimate the restriction probabilities. In the best case scenario, the researcher knows these probabilities or can estimate these from another data source. However, in most cases, the restriction probabilities are unknown; as in the context of labor economics where, for example, job offers are rarely observed. In these cases, the researcher has to disentangle observed choices made due to restrictions and due to preferences. Under certain circumstances, it is possible to recover the restriction probabilities. The discussion is reduced to the following setting:

Restriction Probability Assumption 1. *All restriction sets are known. In particular, depending on the previous choice and state space, the researcher knows the possible choices in the restricted set $\tilde{D}_{d_{t-1}}$. Only the probability that forces agents to choose from $\tilde{D}_{d_{t-1}}$ is unknown.*

In addition, the following holds:

Restriction Probability Assumption 2. *For at least one restriction set $\tilde{D}_{d_{t-1}}$, the restriction probabilities $\pi_{\tilde{D}_{d_{t-1}}}(x_{t-1})$ are known.*

In general, restriction probability assumption 2 is not sufficient to identify all restriction probabilities. To insure identification, further assumptions are necessary. The following discusses three different sets of additional assumptions:

Case 1. *There exists a state, for which agents are not restricted in their choice set, Formally: $\exists j \in D$, s.t. $\pi_{\tilde{D}_j}(x_{t-1}) = 0$, $\forall x_{t-1} \in X_{t-1}$.*[6]

This case automatically fulfills restriction probability assumption 2 since there exists a choice for which the restriction probability is known to be zero. In the labor supply context, this is the case when individuals are already in employment and, thus, can freely choose to continue to work or exit employment. Let k denote previous period's decision if choices are not restricted in the subsequent period, and, d denote a previous period's decision if the subsequent choice set is restricted to $\tilde{D}_d \subset D$. Given the type-I extreme value distribution, we can derive the following observed choice probabilities:

$$\begin{aligned}
\Pr(i \notin \tilde{D}_d | k, x_t, x_{t-1}) =& \frac{\exp(v(i,x_t))}{\sum_{j \in D} \exp(v(j,x_t))} \\
\Pr(i \notin \tilde{D}_d | d, x_t, x_{t-1}) =& (1 - \pi_{\tilde{D}_d}(x_{t-1})) \frac{\exp(v(i,x_t))}{\sum_{j \in D} \exp(v(j,x_t))}
\end{aligned} \tag{1.7}$$

[6] Note that it is possible to include the necessary past values of x_t, i.e. x_{t-1}, also in X_t.

Assuming that $\Pr(i \notin \tilde{D}_d | k, x_t, x_{t-1}) > 0$, the ratio of the choice probabilities, after rearranging, results in

$$\pi_{\tilde{D}_d}(x_{t-1}) = 1 - \frac{\Pr(i \notin \tilde{D}_d | d, x_t, x_{t-1})}{\Pr(i \notin \tilde{D}_d | k, x_t, x_{t-1})}. \tag{1.8}$$

Equation (1.8) identifies the restriction probability $\pi_{\tilde{D}_d}(x_{t-1})$ relying on the observed choice probabilities. In this manner, and, with the help of the unrestricted choice set after choosing k, it is possible to recover all restriction probabilities from the data.

Case 2. *One restriction set includes only one choice and its restriction probability is known. Formally:* $\exists j \in D, s.t.\ \tilde{D}_j$ *is a singleton and* $\pi_{\tilde{D}_j}$ *is known.*

Let $i \in D$ denote the choice related to a restricted choice set with only choice d. Then the observed choice probabilities are

$$\Pr(k \notin \tilde{D}_i | i, x_t, x_{t-1}) = (1 - \pi_{\tilde{D}_i}(x_{t-1})) \frac{\exp(v_{d,t}(x_t))}{\sum_{j \in D} \exp(v_{j,t}(x_t))}. \tag{1.9}$$

Knowing $\pi_{\tilde{D}_i}$, makes it possible to recover $\frac{\exp(v_{d,t}(x_t))}{\sum_{j \in D} \exp(v_{j,t}(x_t))}\ \forall k \notin \tilde{D}_i$. Because $\tilde{D}_i$ is a singleton, it is also possible to recover the respective term for $d \in \tilde{D}_i$:

$$\frac{\exp(v_{d,t}(x_t))}{\sum_{j \in D} \exp(v_{j,t}(x_t))} = 1 - \sum_{k \notin \tilde{D}_i} \frac{\exp(v_{k,t}(x_t))}{\sum_{j \in D} \exp(v_{j,t}(x_t))} \text{ with } d \in \tilde{D}_i. \tag{1.10}$$

With the help of these fractions, all other restriction probabilities can then be identified using the observed choice probabilities on options not included in the restricted choice set:

$$\pi_{\tilde{D}_k}(x_{t-1}) = 1 - \frac{\Pr(d \notin \tilde{D}_k | k, x_t, x_{t-1})}{\exp(v_{d,t}(x_t))} \sum_{j \in D} \exp(v_{j,t}(x_t)). \tag{1.11}$$

Case 3. *Each restriction set excludes at least one choice that is also excluded in another restriction set, i.e.* $\forall \tilde{D}_j\ \exists k \notin \tilde{D}_j$ *and* $k \notin \tilde{D}_d$, *with* $d \neq j$.

In case 3, the probability to choose an option k outside the restriction set $\tilde{D}_i$, after having selected i in the previous period is

$$\Pr(k | i, x_t, x_{t-1}) = (1 - \pi_{\tilde{D}_i}(x_{t-1})) \frac{\exp(v_{k,t}(x_t))}{\sum_{j \in D} \exp(v_{j,t}(x_t))}. \tag{1.12}$$

As long as k is also excluded in another set $\tilde{D}_d$, it is possible to divide both observed choice probabilities. Since it is assumed that one restriction probability is known, it is possible to derive all other restriction probabilities.

As a final remark in this section, it is noteworthy that when all restriction probabilities are *known*, the remaining "pure"[7] choice probabilities can be identified. All observed choice probabilities are one of these two forms

$$\Pr(k|i, x_t, x_{t-1}) = (1 - \pi_{\tilde{D}_i}(x_{t-1})) \frac{\exp(v_{k,t}(x_t))}{\sum_{j \in D} \exp(v_{j,t}(x_t))} \tag{1.13}$$

$$\begin{aligned} \Pr(k|i, x_t, x_{t-1}) = &(1 - \pi_{\tilde{D}_i}(x_{t-1})) \frac{\exp(v_{k,t}(x_t))}{\sum_{j \in D} \exp(v_{j,t}(x_t))} \\ &+ \pi_{\tilde{D}_i}(x_t) \frac{\exp(v_{k,t}(x_t))}{\sum_{j \in \tilde{D}_i} \exp(v_{j,t}(x_t))}, \end{aligned} \tag{1.14}$$

depending on if k is an element of $\tilde{D}_i$ or not. With K denoting the number of possible choices from the unrestricted choice set D, there are a maximum of $K - 1$ independent unknown "pure" choice probabilities. Similarly, for each past choice, there can be $K - 1$ probabilities to choose out of a restricted choice set at maximum. This results in a maximum of $K - 2$ independent choice probabilities from a given restricted choice set since all probabilities have to sum up to one. Thus, for a given state space $\{x_t, x_{t-1}\}$, there is a maximum of $(K-1)+K(K-2) = K^2-K-1$ unknown probabilities that need to be identified. Since the choice probabilities depend on the current and last periods choice, it is always possible to observe $K(K-1) = K^2 - K$ linearly independent choice probabilities, one more than is sufficient for identification.[8]

1.3.2 Identification of Time Preferences

For illustrative purposes of the identification argument, this section starts with a simplified problem. The choice set includes two choices, a choice k, after which the economic agent is not restricted in their subsequent choice, while after choice d, agents face a certain probability to be restricted to choose d again in the subsequent period. First, the central equation for identification is developed. Afterwards, the exclusion restriction is presented and a formula for the discount factor β is discussed. Given that the choice set is not restricted after choosing k, the respective probabilities to choose k and d conditional on having previously selected k is given by

$$\Pr(k|k, x_t, x_{t-1}) = \frac{\exp(v_{k,t}(x_t))}{\sum_{j \in D} \exp(v_{j,t}(x_t))} \tag{1.15}$$

$$\Pr(d|k, x_t, x_{t-1}) = \frac{\exp(v_{d,t}(x_t))}{\sum_{j \in D} \exp(v_{j,t}(x_t))}. \tag{1.16}$$

[7] Meaning the choice probabilities without restrictions.

[8] This additional degree of freedom could in general be used to identify the restriction probabilities as the discussion of the different cases has been indicated.

Taking the logarithm of the quotient of (1.15) and (1.16) results in the difference of the value functions:

$$\begin{aligned}
\ln\left(\frac{\Pr(k|k,x_t,x_{t-1})}{\Pr(d|k,x_t,x_{t-1})}\right) = \\
& u(k,x_t) - u(d,x_t) \\
& + \beta \sum_{X_t} \left(\mathbb{E}\left[\max_{j\in D}\{v_{j,t+1}(x_{t+1}) + \varepsilon_{j,t+1}\}\right]\right) q(x_{t+1}|k,x_t) \\
& - \beta \sum_{X_t} \Big(\Big(1 - \pi_{\tilde{D}_d}(x_t)\Big) \mathbb{E}\left[\max_{j\in D}\{v_{j,t+1}(x_{t+1}) + \varepsilon_{j,t+1}\}\right] \\
& \qquad + \pi_{\tilde{D}_d}(x_t) v_{d,t+1}(x_{t+1})\Big) q(x_{t+1}|d,x_t).
\end{aligned} \tag{1.17}$$

Let $\Gamma(k,d|k,x_t,x_{t-1})$ denote $\ln\left(\frac{\Pr(k|k,x_t,x_{t-1})}{\Pr(d|k,x_t,x_{t-1})}\right)$. Collecting terms and using key assumption 4 leads to

$$\begin{aligned}
\Gamma(k,d|k,x_t,x_{t-1}) =& u(k,x_t) - u(d,x_t) \\
& + \beta \sum_{X_t} \left(\mathbb{E}\left[\max_{j\in D}\{v_{j,t+1}(x_{t+1}) + \varepsilon_{j,t+1})\}\right]\right) q(x_{t+1}|k,x_t) \\
& + \beta \sum_{X_t} \left(\Big(1 - \pi_{\tilde{D}_d}(x_t)\Big) \ln\left(\frac{\exp(v_{d,t+1}(x_{t+1}))}{\sum_{j\in D}\exp(v_{j,t+1}(x_{t+1}))}\right)\right. \\
& \qquad \left. - v_{d,t+1}(x_{t+1})\right) q(x_{t+1}|d,x_t).
\end{aligned} \tag{1.18}$$

It is possible to identify the discounting parameter β if the state space includes a variable that only influences the restriction probability for a limited time period. Assumption 5 summarizes these conditions formally:

Key Assumption 5. *There exist a state variable $x_t \in X_t$ with realizations $x_{A,t}$, $x_{B,t} \in X_t$, $x_{A,t} \neq x_{B,t}$, such that*

(1) $\forall d \in D$, $u(d,x_{A,t}) = u(d,x_{B,t})$;

(2) $\forall d \in D$, $q(x_{t+A}|d,x_{A,t}) = q(x_{t+A}|d,x_{B,t})$;

(3) $\exists k \in D$, $\pi_{\tilde{D}_k}(x_{A,t}) \neq \pi_{\tilde{D}_k}(x_{B,t})$ for at least one $t \leq \check{t}$

and $\check{t}+1$ denotes the first period of the remaining periods for which $\pi_{\tilde{D}_k}(x_{A,\check{t}}) = \pi_{\tilde{D}_k}(x_{B,\check{t}})$.

Key assumption 5 identifies the discount factor because the variable does not influence utilities directly, but only the restriction probabilities. Thus, while the current utility is unaffected by the different values of the state variable, future utilities are not. Differences in the choice probabilities corresponding to the different state values, can, consequently, be attributed to how much the expected future affects agent's decisions. Formally, this can be shown by the observable difference $\Delta_{x_{A,t},x_{B,t}} = \Gamma(k,d|k,x_{A,t},x_{A,t-1}) - \Gamma(k,d|k,x_{B,t},x_{B,t-1})$:

$$\begin{aligned}
&\Delta_{x_{A,t},x_{B,t}} = \\
&u(k,x_{A,t}) - u(d,x_{A,t}) - u(k,x_{B,t}) + u(d,x_{B,t}) \\
&+\beta \sum_{X_{A,t+1}} \left(\mathbb{E}\left[\max_{j\in D} \{v_{j,t+1}(x_{A,t+1}) + \varepsilon_{j,t+1})\} \right] \right) q(x_{A,t+1}|k,x_{A,t}) \\
&+\beta \sum_{X_{A,t+1}} \left((1-\pi_{\tilde{D}_d}(x_{A,t})) \ln\left(\frac{\exp(v_{d,t+1}(x_{A,t+1}))}{\sum_{j\in D}\exp(v_{j,t+1}(x_{A,t+1}))} \right) - v_{d,t+1}(x_{A,t+1}) \right) \\
&\qquad q(x_{A,t+1}|d,x_{A,t}) \\
&-\beta \sum_{X_{B,t+1}} \left(\mathbb{E}\left[\max_{j\in D} \{v_{j,t+1}(x_{B,t+1}) + \varepsilon_{j,t+1})\} \right] \right) q(x_{B,t+1}|k,x_{B,t}) \\
&-\beta \sum_{X_{B,t+1}} \left((1-\pi_{\tilde{D}_d}(x_{B,t})) \ln\left(\frac{\exp(v_{d,t+1}(x_{B,t+1}))}{\sum_{j\in D}\exp(v_{j,t+1}(x_{B,t+1}))} \right) - v_{d,t+1}(x_{B,t+1}) \right) \\
&\qquad q(x_{B,t+1}|d,x_{B,t}).
\end{aligned} \tag{1.19}$$

The first condition of key assumption 5 allows getting rid of the instantaneous utilities on the second line of equation (1.19). Similarly, the expected future value of choice k cancels out of the equation, due to the second condition of key assumption 5. This condition also ensures that the state space of both, A and B, is equal from $t+1$ on. Thus, equation (1.19) simplifies to

$$\begin{aligned}
&\Delta_{x_{A,t},x_{B,t}} = \\
&\quad \beta \sum_{X_{A,t+1}} \left(\left(1-\pi_{\tilde{D}_d}(x_{A,t})\right) \ln\left(\frac{\exp(v_{d,t+1}(x_{A,t+1}))}{\sum_{j\in D}\exp(v_{j,t+1}(x_{A,t+1}))} \right) - v_{d,t+1}(x_{A,t+1}) \right) \\
&\qquad q(x_{A,t+1}|d,x_{A,t}) \\
&\quad -\beta \sum_{X_{A,t+1}} \left(\left(1-\pi_{\tilde{D}_d}(x_{B,t})\right) \ln\left(\frac{\exp(v_{d,t+1}(x_{A,t+1}))}{\sum_{j\in D}\exp(v_{j,t+1}(x_{A,t+1}))} \right) - v_{d,t+1}(x_{A,t+1}) \right) \\
&\qquad q(x_{A,t+1}|d,x_{A,t}).
\end{aligned} \tag{1.20}$$

Using equation (1.16), collecting terms and rearranging results in a formula for the discount factor β:

$$\beta = \frac{\Delta_{x_{A,t},x_{B,t}}}{\left(\pi_{\tilde{D}_d}(x_{A,t}) - \pi_{\tilde{D}_d}(x_{B,t})\right) \sum_{X_{A,t+1}} \left(\ln\left(\Pr(d|k,x_{A,t+1})\right) q(x_{A,t+1}|d,x_{A,t})\right)} \tag{1.21}$$

The numerator of equation (1.21) is the difference in the logarithm of the observed choice probabilities depending on the two values of the state variable. Intuitively, if individuals are fully myopic, i.e. they do not take future values into account when making decisions, the logarithm of the ratio of choice probabilities[9] should be equal for both states of the exclusion variable. This is the case, since the variable only influences future restriction probabilities and not the current utility. If the difference in the observed choice probabilities is indeed zero, β is zero too. This corresponds to the static dynamic discrete choice model, and, thus, reflects that agents are myopic in this case.

The rest of the discussion considers only cases when the numerator of equation (1.21) is unequal zero. Without loss of generality, assume that $\pi_{\tilde{D}_d}(x_{A,t}) - \pi_{\tilde{D}_d}(x_{B,t}) > 0$, i.e. there is a higher chance of being restricted in next period's choice in state A compared to B. Then, all else being equal, the future value of choosing d, the choice with a possible future restriction, is smaller under state A than under state B.[10] With a lower future value under state A for choice d, fewer individuals will choose this option compared to state B. Thus, $\Delta_{x_{A,t},x_{B,t}} = \Gamma(k, d|k, x_{A,t}, x_{A,t-1}) - \Gamma(k, d|k, x_{B,t}, x_{B,t-1})$ is supposed to be negative. Since $\Pr(\cdot|\cdots) > 0$, $q(\cdot|\cdots) \in [0, 1]$, $\sum_{X_{A,t+1}} \left(\ln\left(\Pr(d|k, x_{A,t}, x_{A,t-1})\right) q(x_{A,t+1}|d, x_{A,t})\right) < 0$ is also negative in equation (1.21). Therefore, β is greater than zero.

Furthermore, following the argumentation, holding the difference in the restriction probabilities of states A and B constant, the greater the difference in the observed ratio of choice probabilities, the larger becomes β. Intuitively, the more behavior differs due to changes in the expected future, the more weight individuals place on future values.

In concluding this section, some insights are worth highlighting. First, although theory might predict a $\beta \in [0, 1]$, formula (1.21), does not guarantee that the empirical value lies within this interval. Second, the discussion concentrated on a case for which one choice (k) does not result in a probabilistic restriction in the next period. Equation (1.19) makes it easy to see that the value function of this choice cancels out, even when it relates to a possible restriction in the next period, as long as its restriction probability is not affected by the exclusion variable. Third, if the value function for choice d is known, it is possible to relax the third restriction in key assumption 5. Then it is no longer necessary for the restriction probabilities to be equal after $\check{t}$, since $v_{d,t+1}$ must no

[9]Remember that $\Delta_{x_{A,t},x_{B,t}} = \ln\left(\frac{\Pr(k|k,x_{A,t},x_{A,t-1})}{\Pr(d|k,x_{A,t},x_{A,t-1})}\right) - \ln\left(\frac{\Pr(k|k,x_{B,t},x_{B,t-1})}{\Pr(d|k,x_{B,t},x_{B,t-1})}\right)$

[10]This becomes apparent when regarding equation (1.4). The higher the restriction probability, the more weight is placed on the expected maximum of the restricted choice set. This expected maximum can never be greater than the expected maximum of the unrestricted choice set, since it includes at least the same possible options as in the restricted choice set. Under key assumption 4, this is formally given by $\log(\sum_{j\in\tilde{D}\subset D} \exp(v_{j,t}(x_t))) \leq \log(\sum_{j\in D} \exp(v_{j,t}(x_t)))$.

longer cancel out in equation (1.19). Finally, for a model with an infinite horizon, the central equation for the identification of β also holds, as long as the exclusion restriction is fulfilled and x_A and x_B only differ for a limited amount of time.

1.4 Conclusion

This chapter discusses a new exclusion restriction to identify the exponential discount factor in dynamic discrete choice models. It relies on temporary differences in restriction probabilities. Preceding the discussion of the discounting parameter's point identification, the identification of the restriction probabilities is examined. If these probabilities are not known and cannot be recovered from other data, it is possible to identify them if the choice set is not restricted after at least one choice. If all choices lead to restricted choice sets in the following period, the probabilities can be recovered in special cases, given that at least one of the restriction probabilities is known.
Given the restriction, transition and empirical choice probabilities, the identification of the discounting parameter β is presented. To recover β from the data, it is sufficient that a variable exists that only influences restriction probabilities temporary. In the labor market, this could be an active labor market policy that supports unemployed individuals for a given period to find employment, thus, temporarily decreasing their restriction probability to stay in unemployment. For industrial and environmental economics, the exponential discount factor might be recovered from temporarily changes in merger or emission restrictions. In marketing, the exclusion restriction can be satisfied when an alternative product is first introduced in part of a market before being offered on the whole market. This creates different groups with temporarily different choice sets. In models with restriction probabilities, depending on the specific situation, it might be far easier to find variables satisfying the presented exclusion restriction than those presented in Magnac & Thesmar (2002), Fang & Wang (2015), or Abbring & Daljord (2016). Future research might derive conditions to identify, in addition to the exponential discount factor, the parameters of hyperbolic discounting.

Chapter 2

Time Preferences and Female Labor Supply

2.1 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 & Wolpin (1989), Keane & Wolpin (1997), and more recently Blundell et al. (2016), Adda et al. (2017). 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: 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 restrictive assumption of time consistent exponential discounting with only a few exceptions (Fang & Silverman, 2009; Chan, 2017). One reason for this assumption might be that time preference parameters are generally not identified in a dynamic discrete choice framework (Magnac & Thesmar, 2002). However, as Fang & Wang (2015) show, under certain circumstances, it might be possible to recover these preference parameters from choice data. Variations in transition probabilities that do not directly affect flow utilities can be sufficient for identification. More general, exclusion restrictions that affect future transitions of individuals, but leave flow utilities unaffected, are required.

In this chapter, we exploit exogenous changes in the duration of employment protection for mothers to identify time preferences in a dynamic model of female labor supply with labor market frictions. Employment protection provides insurance against labor market frictions for mothers after parental leave, thereby influencing female labor supply possibilities. Crucially for our identification strategy, employment protection does not directly impact the flow utilities of mothers, but only future employment transitions by guaranteeing employment opportunities in future periods. We specify time preferences to be (quasi-)hyperbolic as in Laibson (1997), Fang & Silverman (2009), and Chan (2017); a specification that nests time consistent exponential discounting as a possibility.

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 observed career interruptions of mothers after childbirth, which cause large career costs (Adda et al., 2017) and, thus, are an important determinant of the female-male wage gap (Kleven et al., 2015; Ejrnaes & Kunze, 2013; Anderson et al., 2002). Similarly, the employment and welfare effects of family and labor market policies depend on the time preferences of mothers. Quasi-hyperbolic discounters are especially sensitive to current costs and benefits, resulting in possible large employment effects even for short-lived in-work benefits. In contrast, for time consistent mothers, even large subsidies for a short period should not lead to sizable employment effects as they have just a minor effect on the exponentially discounted life-cycle income. Time preferences are, thus, of prime importance for predicting reactions to different policy instruments.

The structure of time preferences is the object of numerous studies, beginning with the work of Strotz (1956), Phelps & Pollak (1968), and Pollak (1968). Following the seminal works of Laibson (1997) and O'Donoghue & Rabin (1999), applications of (quasi-)hyperbolic discounting have become popular in experimental and observational studies. For instances, Harris & Laibson (2002) rely on those preferences to explain consumption decisions, Diamond & Köszegi (2003) and Gustman & Steinmeier (2012) to explain retirement and saving decisions, and Martinez et al. (2017) to explain tax-filing behavior.

The research of Fang & Silverman (2009) and Chan (2017) is most related to our study and we extend their work in several dimensions. Fang & Silverman (2009) only rely on structural form assumptions and employment behavior over the life-cycle to identify time preferences, while Chan (2017) uses administrative data from a field experiment, which exogenously varies time limits of welfare benefits. Because his panel is rather short (three years), he actually never observes women exhausting their time limits. Our identification strategy relies on exogenous variation caused by maternity leave reforms that changed the length of employment protection for previously working mothers. Thus, we use a different kind of exclusion restriction, relying on exogenous changes in choice restrictions instead of changes in welfare eligibility. Extending Chan (2017), we use a long panel of women, which allows us to look at long-term behavior. We are also able to show that labor market frictions play an important role, an aspect ignored by both studies.

Furthermore, both Fang & Silverman (2009) and Chan (2017) only focus on welfare dependence of single mothers, not including the age of children in their models. In our study, the age of children is crucial for leisure preferences, childcare costs, and possible benefits. This might be important when examining labor supply choices, since mothers might have stronger preferences for leisure when their children are young than when they are of school age, causing variation in employment behavior depending on the age of children. Since previous studies focus exclusively on the United States of America, we are the first to provide evidence of hyperbolic preferences in the context of female labor supply for a European Country, analyzing German women. Germany, like almost all countries besides the United States of America (International Labour Organization, 2012), provides paid maternity leave to mothers who were previously employed. Our study might, therefore, provide stronger external validity for other countries than previous studies.

The remainder of the chapter is structured as follows. Section 2.2 presents the dynamic labor supply model. Section 2.3 presents the data and introduces first suggestive evidence for time-inconsistent behavior in the context of maternal employment. Section 2.4 provides background information on the institutions. Section 2.5 discusses identification of key parameters and the estimation procedure. Section 2.6 presents the estimation results, and section 2.7 concludes.

2.2 Economic Model

The life-cycle model starts after individuals have finished their education and enter the labor force. Although we differentiate between low and high education – low defined as high school or less, high as at least some college or higher education – women might enter the labor force at various ages. This entry age is assumed to be exogenous and depends on the actual, in the data, observed entry age. Once a woman enters the labor market, she makes half-yearly labor supply choices according to her discounted utility stream, while facing labor market frictions.

Flow utilities depend on leisure and consumption opportunities, with the latter depending on the hourly wage, childcare costs, as well as the tax and transfer system. The wage process allows for endogenous human capital accumulation and depreciation (at different rates in full-time, part-time, and non-employment) following Adda et al. (2017) and Blundell et al. (2016). Furthermore, we take into account incentives to work stemming from the tax and transfer system, including joint taxation, unemployment benefits, social assistance, childcare costs, and employment protection. The latter provides insurance for mothers against labor market frictions during maternity leave, which we model as stochastic job offer arrivals. Time preferences are modeled as in Laibson (1997), Fang & Silverman (2009), and Chan (2017), nesting both time-consistent exponential discounting and time-inconsistent (quasi-)hyperbolic discounting. The rest of this section discusses the functional form assumptions in more detail.

2.2.1 The Structural Model

Every half year, employed individuals (and non-employed individuals who receive a job offer) have to choose their level of labor supply $l_{i,t}$ from a choice set of non-employment ($l_{i,t} = 0$), part-time work ($l_{i,t} = 1$), and full-time work ($l_{i,t} = 2$). Non-employed women, who do not receive a job offer, are forced to choose non-employment ($l_{i,t} = 0$). Full-time workers are assumed to work twice as many hours as part-time workers.[11]

Flow Utility. The instantaneous utility is similar to Adda et al. (2017) and is given by

$$\begin{aligned} u_{i,t} =& \frac{[c_{i,t}/\bar{c}_{eq}]^{(1-\gamma_C)}}{1-\gamma_C} \\ & \times \exp\Big(\Big[\gamma_{PT_{low}} \mathbb{1}_{\{l_{i,t}=1\}} + \gamma_{FT_{low}} \mathbb{1}_{\{l_{i,t}=2\}}\Big] \times \mathbb{1}_{\{educ=low\}} \\ & \qquad + \Big[\gamma_{PT_{high}} \mathbb{1}_{\{l_{i,t}=1\}} + \gamma_{FT_{high}} \mathbb{1}_{\{l_{i,t}=2\}}\Big] \times \mathbb{1}_{\{educ=high\}}\Big) \\ & \times \left(\gamma^{PT}_{ageYC,0} + \gamma^{PT}_{ageYC,1} ageYC_{i,t} + \gamma^{PT}_{ageYC,2} ageYC^2_{i,t}\right)^{\mathbb{1}\{l_{i,t}=1, CD_{i,t}=1\}} \\ & \times \left(\gamma^{FT}_{ageYC,0} + \gamma^{FT}_{ageYC,1} ageYC_{i,t} + \gamma^{FT}_{ageYC,2} ageYC^2_{i,t}\right)^{\mathbb{1}\{l_{i,t}=2, CD_{i,t}=1\}} \end{aligned} \tag{2.1}$$

where $c_{i,t}$ denotes the consumption, $\bar{c}$ an equivalence scale,[12] $CD_{i,t}$ indicates the presence of children, and $age^{YC}_{i,t}$ the age of the youngest child. Furthermore, $\mathbb{1}_{\{condition\}}$ is an indicator function that equals 1 if the condition is true and 0 otherwise.

The first term of equation (2.1) presents the utility derived from consumption, in form of a standard CRRA function. Parameter γ_C represents women's risk aversion. The other parts of equation (2.1) determine the utility derived from leisure, where utility from non-working is used as the reference category. Our functional form captures education specific leisure preferences for part-time and full-time employment in addition to preferences depending on the age of the youngest child. The latter are modeled in a quadratic form. Both utility derived from leisure and consumption are inseparable.

Wages and Human Capital. Wages are important for the consumption opportunities of a respective labor supply choice. We model the wage process following Blundell

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

[12] We assume that $\bar{c} = 1$ for a single, and $\bar{c} = 1.4$ if children are present.

et al. (2016):[13]

$$\begin{aligned}\ln(w_{i,t}) = \ln(\gamma_{w,low})\mathbb{1}_{\{s=low\}} + \ln(\gamma_{w,high})\mathbb{1}_{\{s=high\}} \\ + \gamma_{w,s,e}\ln(e_{i,t}+1) + \xi_{i,t}\end{aligned} \tag{2.2}$$

The hourly wage rate depends on an education specific constant, $\gamma_{w,s}$, and on accumulated on-the-job human capital, e. Wages are subject to a measurement error ($\xi_{i,t}$), which follows a normal distribution with standard deviation σ_ξ. Since levels of education do not change over the life-cycle, wage trajectories are mostly driven by accumulation and depreciation of on-the-job human capital. This process is education specific and given by

$$e_{i,t} = e_{i,t-1}(1-\eta_s) + \begin{cases} 0 & \text{if } l_{i,t-1} = 0 \text{ (not employed)} \\ 0.25 & \text{if } l_{i,t-1} = 1 \text{ (part-time)} \\ 0.5 & \text{if } l_{i,t-1} = 2 \text{ (full-time)} \end{cases} \tag{2.3}$$

In each period, human capital depreciates with an education-specific rate η_s,[14] which can only be offset if the individual is employed. The possible accumulation depends on education and whether a woman works part-time or full-time. Since the decision period is semi-annual, the per-period growth in full-time employment is normalized to 0.5, and in part-time employment to 0.25.

Budget Constraint. Given the labor supply decision and the wage process, consumption is determined by

$$\begin{aligned} c_{i,t} = & w_{i,t} \times 452 \times (2 \times \mathbb{1}_{\{l_{i,t}=2\}} + \mathbb{1}_{\{l_{i,t}=1\}}) \\ & - TT(earn^W_{i,t}, ageYC_{i,t}) \\ & - cc^E \times 452 \times CD_{i,t} \times (2 \times \mathbb{1}_{\{l_{i,t}=2\}} + \mathbb{1}_{\{l_{i,t}=1\}}) \end{aligned} \tag{2.4}$$

where $earn^W_{i,t}$ stands for the gross earnings of the woman, the function $TT(\cdot)$ for the German tax and transfer system, $ageYC_{i,t}$ for the age of the youngest child, and cc^E for 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 for a working day. Therefore, the first line of equation (2.4) describes the women's half-yearly labor earnings depending on the labor supply and wage rate.

We model all relevant features of the German tax and transfer system, which depends

[13] In contrast to Blundell et al. (2016), we only model transitory shocks and do not estimate the human capital growth when working part-time.

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

on earnings, labor supply, and the presence of children. The system includes income taxation, social security contributions, and child benefits. Non-employed women are eligible for unemployment and social assistance benefits, while maternity benefits depend on the age of the youngest child and the maternity leave regime the child was born under.

In Germany, subsidized childcare slots are rationed, but we assume that employed mothers 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):

$$cc^E = \pi^S cc^S + (1 - \pi^S) cc^{NS} \tag{2.5}$$

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

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

$$\pi_{i,t}^{JO} = \gamma_{JO,low} \mathbb{1}_{\{s=low\}} + \gamma_{JO,high} \mathbb{1}_{\{s=high\}} \tag{2.6}$$

Therefore, the job offer probability only depends on education. Importantly, after giving birth, previously employed mothers benefit from employment 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 employment 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. Family dynamics are modeled as exogenous stochastic processes. The probability of having a child depends on woman's age, education and the presence of other children. It is assumed that all children live with their mother until the age of 18.

2.2.2 Intertemporal Optimization

Choice specific utilities consist of two parts, a deterministic part, depending on the choice's leisure and consumption opportunities, and a stochastic part, given by a preference shock ($\varepsilon_{i,l,t}$). As standard in the literature of dynamic discrete choice models, choice specific shocks are assumed to follow a type-1 extreme value distribution and are independently distributed over choices and time. Thus, the instantaneous utilities can be written as

$$v_{i,l,t} = u_{i,t}(l_t) + \varepsilon_{i,l,t} \tag{2.7}$$

These notations allow to denote the expected lifetime utility U_t for a given periods as

$$\max_{\{l_t,l_{t+1},\ldots,l_T\}} U_t(l_t, l_{t+1}, \ldots, l_T, \Omega_t) = v(l_t, \Omega_t) + \beta\mathbb{E}\left[\sum_{\tau=t+1}^{T} \delta^{\tau-t} v(l_\tau, \Omega_\tau) \middle| \Omega_t\right] \tag{2.8}$$

where we drop the index i for ease of notation, use $\mathbb{E}[\cdot]$ as the expectations operator and Ω_t to denote the state space at time t:

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

Employment protection is denoted by the indicator jp_t, which equals one if a woman is in employment protection, and zero otherwise.

All future period's utilities are discounted when added to the perceived lifetime utility, depending on how further into the future they are. In addition to an exponential discount factor δ, individuals discount any future values by β. Figure 2.1 illustrates this process. The difference between the discount factor of any two future periods (e.g. period 2 and 3 in figure 2.1), is captured by the exponential parameters δ. While, for the current perceived lifetime utility, all future periods are further discounted by the parameter β. For $\beta = 1$, this additional discounting cancels out, and time preferences are purely exponential. For $\beta < 1$ individuals display a present bias (O'Donoghue & Rabin, 1999), indicating the impulse for immediate gratification.

These time-preferences are also known as (β, δ)-preferences (O'Donoghue & Rabin, 1999) and are often applied when individuals are prone to time-inconsistent behavior. In our framework, this behavior might arises, once a woman progresses in time. Consider, for example, the situation depicted in figure 2.1. When women make their choices in period t_0, their consequences for period t_1 and t_2 enter the perceived lifetime utility only with a difference in weight of δ. Once the woman progressed to period t_1, the weight difference changes to $\beta\delta$, potentially changing behavior to a choice that leads to more utility in period t_1 and less utility in period t_2 compared to time-consistent behavior. In our model, this could lead women to repeatedly postpone their return to employment after having a child. For instance, women might plan to return to employment at a certain age of their child, but once they reach this period, they favor current utility from leisure over future career opportunities and, thus, plan to stay out

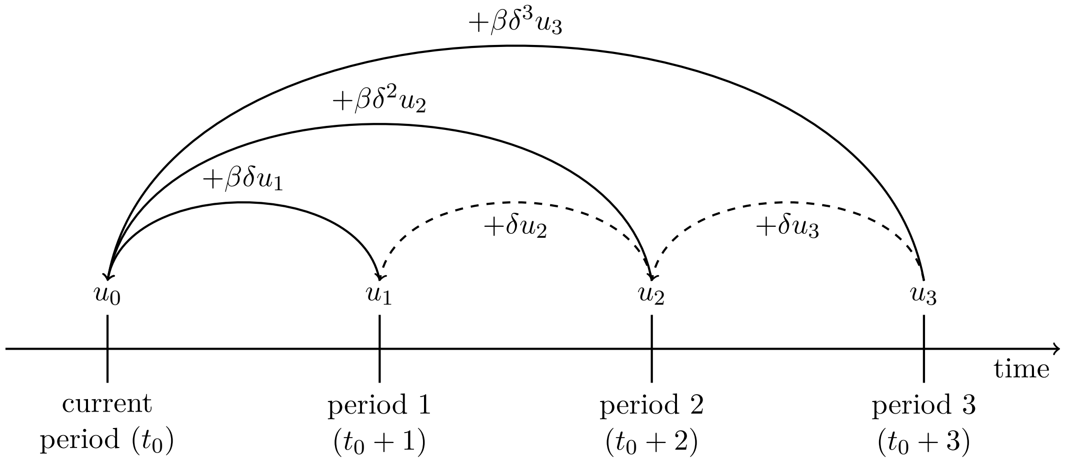

Figure 2.1: Quasi-hyperbolic discounting

of employment for one more period. In the next period, they face the same situation, again postponing their return. Therefore, this behavior can lead to longer career breaks than under exponential discounting.

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 contrast, individuals who are unaware of their time-inconsistencies are called *naïve*. We follow Fang & Wang (2015) and assume that agents are fully naïve. This is also in line with evidence of Chan (2017), who cannot reject a specification in which individuals are fully naïve.

2.2.3 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, comparing utility streams arising in $t+j$ and $t+j+1$ (with $j \geq 1$) from the point of a decision period t. Making use of the assumption that agents are fully naïve, we can rewrite the value function in recursive form (see Fang & Silverman, 2009; Chan, 2017):

$$
\begin{aligned}
&\text{for } t+j \neq T: \\
&\qquad V_{t+j}(\Omega_{t+j}) = \max_{l_{t+j}\in\{0,1,2\}} v(l_{t+j}, \Omega_t) + \delta \mathbb{E}\left[V_{t+j+1}(\Omega_{t+j+1}) \middle| \Omega_{t+j} \right] \\
&\text{for } t+j = T: \\
&\qquad V_{t+j}(\Omega_{t+j}) = \max_{l_{t+j}\in\{0,1,2\}} v(l_{t+j}, \Omega_{t+j})
\end{aligned}
\tag{2.9}
$$

Note that since $j \geq 1$, we are only considering future periods, and consequently β is not included in equation (2.9).
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 $t+j$. We use the notation $\mathbb{E}\max_{t+j}(\widetilde{\Omega}_{t+j+1})$ to denote a specific realization of the state space in $t+j+1$. The assumption of a finite horizon allows solving 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 given by:

$$\begin{aligned}
&\text{for } t+j < T-1: \\
&\qquad \mathbb{E}\max_{t+j}^{\text{non-emp}}(\widetilde{\Omega}_{t+j+1}) = \gamma + u(l_{t+j+1}=0, \widetilde{\Omega}_{t+j+1}) \\
&\qquad\qquad + \delta \mathbb{E}\max_{t+j+1}(\widetilde{\Omega}_{t+j+1}) \\
&\text{for } t+j = T-1: \\
&\qquad \mathbb{E}\max_{t+j}^{\text{non-emp}}(\widetilde{\Omega}_{t+j+1}) = \gamma + u(l_{t+j+1}=0, \widetilde{\Omega}_{t+j+1})
\end{aligned} \tag{2.10}$$

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 $\mathbb{E}\max_{t+j}$ is defined by:

$$\begin{aligned}
&\text{for } t+j < T-1: \\
&\qquad \mathbb{E}\max_{t+j}^{\text{emp}}(\widetilde{\Omega}_{t+1}) = \gamma + \log\left[\sum_{l_{t+j+1}=0}^{2} \exp\Big(u(l_{t+j+1}, \widetilde{\Omega}_{t+j+1})\Big)\right] \\
&\qquad\qquad + \delta \mathbb{E}\max_{t+j+1}(\widetilde{\Omega}_{t+j+1}) \\
&\text{for } t+j = T-1: \\
&\qquad \mathbb{E}\max_{t+j}^{\text{emp}}(\widetilde{\Omega}_{t+j+1}) = \gamma + \log\left[\sum_{l_{t+j+1}=0}^{2} \exp\Big(u(l_{t+j+1}, \widetilde{\Omega}_{t+j+1})\Big)\right]
\end{aligned} \tag{2.11}$$

Building on equations (2.10) and (2.11), 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 $\mathbb{E}\max$:

$$\begin{aligned}
\mathbb{E}\max{}_{t+j} \mid (l_{t+j} = 0, \Omega_{t+j}) = \sum_{\widetilde{\Omega}_{t+j+1}} & \Pr\left(\widetilde{\Omega}_{t+j+1}\middle|\,\Omega_{t+j}\right) \Big[\\
& \pi^{JO}_{i,t+j} \times \mathbb{E}\max{}^{\text{emp}}_{t+j}(\widetilde{\Omega}_{i,t+j+1}) \\
& + (1 - \pi^{JO}_{i,t+j}) \times \mathbb{E}\max{}^{\text{non-emp}}_{t+j}(\widetilde{\Omega}_{i,t+j+1})\Big] \\
\mathbb{E}\max{}_{t+j} \mid (l_{t+j} \in \{1,2\}, \Omega_{t+j}) = \sum_{\widetilde{\Omega}_{t+j+1}} & \Pr\left(\widetilde{\Omega}_{t+j+1}\middle|\,\Omega_{t+j}\right) \Big[\\
& (1 - \pi^{JL}_{i,t+j}) \times \mathbb{E}\max{}^{\text{emp}}_{t+j}(\widetilde{\Omega}_{i,t+j+1}) \\
& + \pi^{JL}_{i,t+j} \times \mathbb{E}\max{}^{\text{non-emp}}_{t+j}(\widetilde{\Omega}_{i,t+j+1})\Big]
\end{aligned} \tag{2.12}$$

With equation (2.12), we can rewrite equation (2.8) as

$$\max_{\{l_t, l_{t+1}, \ldots, l_T\}} U_t(l_t, l_{t+1}, \ldots, l_T, \Omega_t) = u(l_t, \Omega_t) + \beta\delta\mathbb{E}\max{}_t \tag{2.13}$$

For our identification strategy, it is worth pointing 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 (2.13).

2.3 Data and Descriptive Evidence

2.3.1 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).[15] During annual SOEP interviews, participating individuals are asked to fill out a monthly calendar for 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 50. As the reforms exploited here do not apply to them, we exclude women who have worked as civil servants or were self-employed at some point. The final data set is, therefore, an unbalanced panel that individuals enter and leave at various points in time. We observe over $6,200$ women, on average for five and a half years. Additionally, we observe $1,375$ births and a total of $3,861$ children aged between 0 and 18. In total we have $419,855$

[15] We use some additional answers from wave 2007, since they are retrospective for wave 2006.

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.[16]

2.3.2 Suggestive Evidence of Time-inconsistent Choices

Naïve hyperbolic agents are prone to making systematic errors when predicting their own behavior in the future, since plans about future decisions are continuously revised. Following this argument, table 2.1 gives a first indication that mothers might indeed be time-inconsistent on average. The data is based on a question of the SOEP survey that asks non-employed individuals when they plan to return to employment. Focusing only on mothers, who have a guarantee to return to their previous position for three years, the table compares stated preferences and actual realizations.

The "1st Year" column of table 2.1 presents the preferred return time to the labor

Table 2.1: Returns 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: 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 actual observed duration of career break or a career break of at least five years.

market, stated at the first interview after the last, in the data, observed 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 after five years. Tracking the career breaks of all

[16] These are based on 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.

these women, reveals a strong shift toward longer career breaks than initially stated in this first year. The fraction of mothers who return to the labor market within one year is only around half the fraction who wanted to return within that time span. Additionally, the ratio of realized career breaks that last five years or longer (25.5%) is ten percentage points higher than the previously stated ones.

The previous trend continues in the second and third years, which suggests that one-time errors, like unexpected childcare availability or lacking support of the partner, are unlikely to explain the gap completely. Mothers should have updated their beliefs in the first year and adapt their expectations for the second and third years. In addition, at least for mothers, who stated wanting to return within one year, it is possible to exclude factors stemming from the labor demand side. This results from all mothers, reported in table 2.1, having the right to return to their previous position within the first three years after the birth of their child. For models based on rational expectations and exponential discounting, it is hard to explain the observed gap between preferences and realizations. By introducing the possibility of time-inconsistent behavior into a model, it is possible to capture this pattern in a natural way.

2.4 Institutions

In this analysis, we focus on Germany for two reasons. First, several policy reforms changed the length of maternal employment protection during our observation period, which we exploit for identification of the discount parameters of our model. Second, one of the policy regimes has a very generous employment protection period, which allows mothers to return to their previous job within three years. This helps to identify leisure preferences depending on the early ages of children, since returning behavior is not driven by labor market frictions during this period.

For the identification of time-preferences, we concentrate on six major expansions of maternity leave coverage between 1986 and 1993,[17] which we summarize in three major policy regimes. One objective of these reforms was to allow mothers to spend more time with their children during their very early development. Another objective was to strengthen mothers' labor market attachment, since a longer employment protection period was seen to ease the return to the labor market after maternity leave. In the following, we briefly discuss the policies and reforms necessary for our identification strategy.

Since the late-1960s, mothers are entitled to 14 weeks of paid leave around childbirth.

[17]The summary of the parental leave reforms is 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.

In general, the time is divided into six weeks before the expected birth and eight weeks after, and women are not allowed to work during these weeks. For the 14 weeks, employees cannot be dismissed and are guaranteed a comparable job to their previous position upon their return. Women receive the average income of the three months before entering maternity leave, replacing their would-be earnings by 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 employment protection period was extended to six months after childbirth, while a new maternity leave payment for the time between the eighth week and the end of the sixth month after childbirth was introduced. For this period, women received DM750 per month. All the newly introduced maternity benefits were only paid if the mother was employed before childbirth.
The reforms we focus on started in January 1986. An overview of these are shown in table 2.2. The first reform expanded the employment protection and maternity benefit period from six to 10 months at the beginning of 1986, and then further to 12 months in January 1988.[18] Maternity payments from weeks six to eight remained at an income replacement of 100% or DM600[19] if the mother was non-employed before. From month three to six, maternity benefits declined from DM750 to DM600[19] 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 & Ludsteck, 2014).

A further increase in the length of employment protection and maternity benefit from

Table 2.2: 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 DM600,[19] 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		

[18] Additionally, parental leave for fathers was introduced. However, on average only around 1% of fathers took parental leave between 1987 and 1994 (Vaskovics & Rost, 1999).
[19] This is equivalent to $585 in 2016.

12 to 15 months took place in July 1989, and another increase to 18 months in July 1990. In January 1992, the employment 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.[20] Some minor clarifications in family policy were introduced in 2001, but the core regime of 1993 still continued.

Table 2.2 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 six months, we cannot take into account changes in employment protection from 10-12 or 15-18 months. Therefore, we approximate the duration of employment 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 is 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 employment protection periods, help us to identify the parameters of our structural model as we explain in more detail in the next section.

2.5 Identification and Estimation

2.5.1 Identification

2.5.1.1 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 discounting parameters when only relying on observed choice probabilities. Magnac & Thesmar (2002) derive specific conditions for identification of the exponential discounting parameter. Building on these conditions, Fang & Wang (2015) develop specific exclusion restrictions that 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

[20] 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 DM100,000, for singles DM75,000, to receive the full benefits for the first six months.

only differ in the values of such a variable. Although these two do not differ in their flow utilities, they do differ in their expected future utility streams, since the state space develops differently. Observed differences in choice probabilities are then only caused by their expected futures and, thus, can inform the researcher how individuals weight these expected utility streams.

Our framework provides a natural instrument that is based on the same intuition as the exclusion restrictions in Fang & Wang (2015): the change in the length of employment protection. Although variation in the length of the protection does not change exogenous transition probabilities, it does change the probability of being able to choose employment in future periods. Thus, the length of employment protection does not enter the utility function directly (see equation (2.1)), but might change lifetime utility, because of its effect on discounted future expected utilities. A comparison of two groups of individuals, which differ only by the length of employment protection, permits identifying how individuals value future expected utilities.

Assume that one group has employment protection for 3 years after childbirth, while another group is protected for only 1.5 years. Despite the difference in their probabilities of being able to enter employment when their child is two years old, their instantaneous utility does not differ on average in the first period after child birth.[21] 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 different expected futures, individuals must be myopic.

Comparing groups of individuals that only differ in their employment probability in the very next period,[22] informs about the one-period ahead discount factor, which in our model is given by the product $\beta\delta$. Comparing groups that differ in employment opportunities further in the future,[23] can then help to identify the additional needed discounting compared 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 & 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 2.2 shows that women behaved differently across different regimes, which is a strong indicator of the salience of the policy.

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

[22] For instance, compare groups in regime I and regime II, when the youngest child is 0.5.

[23] For instance, compare groups in regime II and regime III, when the youngest child is 0.5.

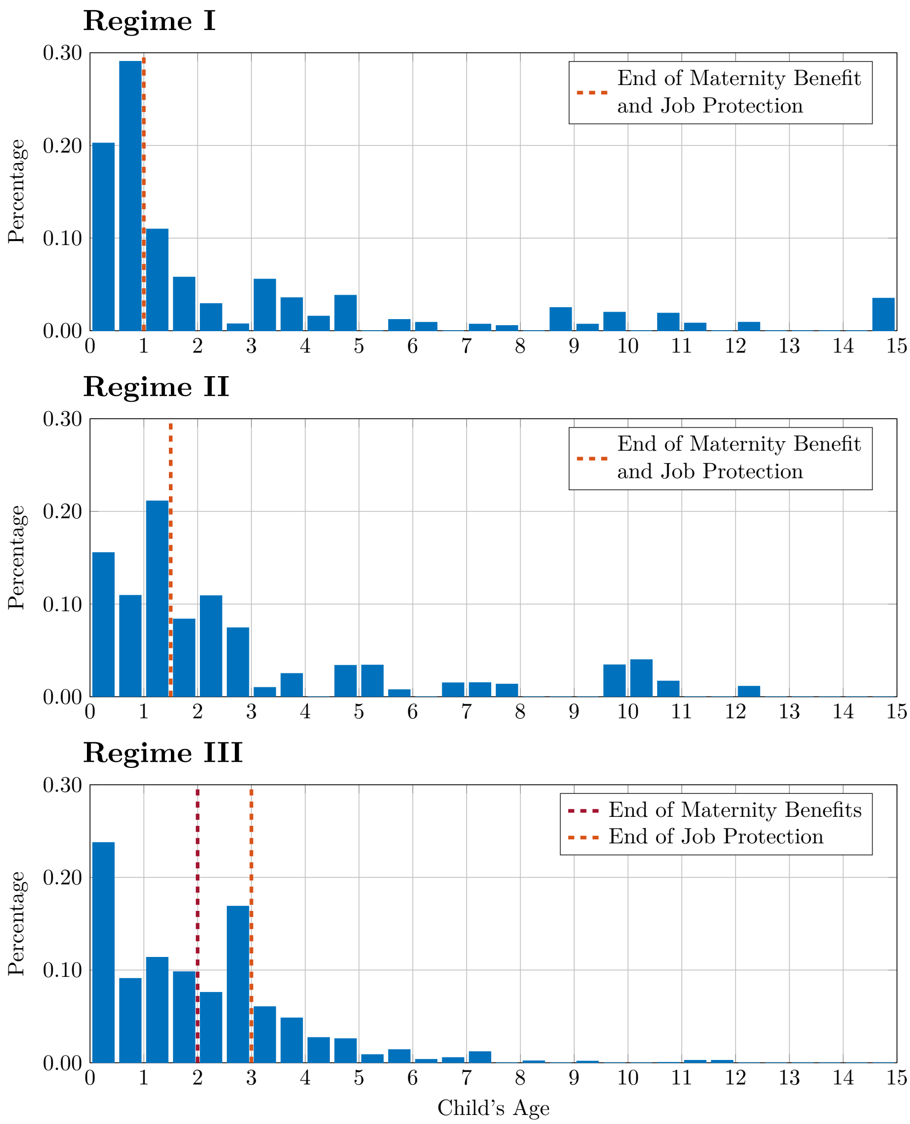

Figure 2.2: Length of career breaks

Notes: Histogram of the length of career breaks after the birth of a child 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 has another child. Hence, only mothers for whom we observe the employment status from the birth of a child until they are employed or have another child are included. Observations are weighted with SOEP sampling weights.

Figure 2.2 reports the distribution of career breaks of mothers for different lengths of employment protection. Most mothers return just in time to not lose their guarantee to be able to return to their former employment. That the behavior is strongly affected by the length of employment protection is 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 point, 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.

2.5.1.2 Formal Illustration

We illustrate our identification argument in a three period model with two groups of individuals. Assume the two groups, A and B, only differ with respect to the length of employment protection, i.e. the time within they are guaranteed to re-enter employment by law. Group A benefits from three periods of employment protection after childbirth, while group B only from two periods. We denote employment protection in period t by jb_t, and no employment protection by $\overline{jb}_t$.[24]

In this three period model, it is possible to use choice probabilities of the last period to identify all utility levels of the third period. Similar to the utility function of the discussed model, assume that the utility of non-employment is normalized to zero, since only differences in utility can be identified. Only considering individuals who enjoy employment protection in the last period and given the distributional assumption about the error terms, the choice probabilities for a particular state space are

$$\begin{aligned}
\Pr(l_3 = 0|A, \widetilde{\Omega}_3) =& \frac{1}{\sum_{j=0}^{2} \exp(u(j, \widetilde{\Omega}_3))} \\
\Pr(l_3 = 1|A, \widetilde{\Omega}_3) =& \frac{\exp(u(1, \widetilde{\Omega}_3))}{\sum_{j=0}^{2} \exp(u(j, \widetilde{\Omega}_3))} \\
\Pr(l_3 = 2|A, \widetilde{\Omega}_3) =& \frac{\exp(u(2, \widetilde{\Omega}_3))}{\sum_{j=0}^{2} \exp(u(j, \widetilde{\Omega}_3))}
\end{aligned} \tag{2.14}$$

[24]Note that the state-space of the model presented in section 2.2 includes only the past status of job protection, because last period's employment status and the policy regime jointly determine the current period's employment protection status. In this subsection, we explicitly indicate the potential employment protection status of future periods.

Dividing the choice probability for part-time and full-time employment by the choice probability of non-working, allows for identifying all utility levels for the last period:[25]

$$
\begin{aligned}
u(1, \widetilde{\Omega}_3) &= \ln\Big(\Pr(l_3 = 1|A, \widetilde{\Omega}_3)\Big) - \ln\Big(\Pr(l_3 = 0|A, \widetilde{\Omega}_3)\Big) \\
u(2, \widetilde{\Omega}_3) &= \ln\Big(\Pr(l_3 = 2|A, \widetilde{\Omega}_3)\Big) - \ln\Big(\Pr(l_3 = 0|A, \widetilde{\Omega}_3)\Big)
\end{aligned} \tag{2.15}
$$

Having identified the flow utilities of period three and observing the transition probabilities from period two to three, all $\mathbb{E}\max(\cdot)$ values can be computed, since these are only functions of the two. Note that not only the expected future maxima of individuals of group A are identified, but also those for group B. Given the utility values, it is possible to use the observed transition probability from non-employment to full-time and part-time employment to identify the job offer probabilities. These are the only additional unknowns for the computations of the $\mathbb{E}\max(\cdot)$ values also for individuals of group B. For identification of the product $\beta\delta$, consider the second period's decision of individuals, who have not started to work after having their last child. Their respective lifetime utilities are given by

$$
\begin{aligned}
U_2^A &= v(l_2; \Omega_2) + \beta\delta\mathbb{E}\max\left(l_2, \Omega_2, jp_3\right) \\
U_2^B &= v(l_2; \Omega_2) + \beta\delta\mathbb{E}\max\left(l_2, \Omega_2, \overline{jp}_3\right)
\end{aligned} \tag{2.16}
$$

Similar to equation (2.15), it is possible to extract a combination of utilities and future expected maxima by using the logarithm of observed choice probabilities:

$$
\begin{aligned}
\ln\left(\frac{\Pr(l_2 = 1|A, \widetilde{\Omega}_2)}{\Pr(l_2 = 0|A, \widetilde{\Omega}_2)}\right) &= u(1, \widetilde{\Omega}_2) + \beta\delta\Big(\mathbb{E}\max(1, A, \widetilde{\Omega}_2) - \mathbb{E}\max(0, A, \widetilde{\Omega}_2)\Big) \\
\ln\left(\frac{\Pr(l_2 = 1|B, \widetilde{\Omega}_2)}{\Pr(l_2 = 0|B, \widetilde{\Omega}_2)}\right) &= u(1, \widetilde{\Omega}_2) + \beta\delta\Big(\mathbb{E}\max(1, B, \widetilde{\Omega}_2) - \mathbb{E}\max(0, B, \widetilde{\Omega}_2)\Big)
\end{aligned} \tag{2.17}
$$

Since the utility function does not differ for individuals of the two groups, subtracting both logarithms identifies the product $\beta\delta$:

$$
\beta\delta = \frac{\ln\left(\frac{\Pr(l_2=1|B,\widetilde{\Omega}_2)}{\Pr(l_2=0|B,\widetilde{\Omega}_2)}\right) - \ln\left(\frac{\Pr(l_2=1|A,\widetilde{\Omega}_2)}{\Pr(l_2=0|A,\widetilde{\Omega}_2)}\right)}{\Big(\mathbb{E}\max(1, B, \widetilde{\Omega}_2) - \mathbb{E}\max(1, A, \widetilde{\Omega}_2)\Big) - \Big(\mathbb{E}\max(0, B, \widetilde{\Omega}_2) - \mathbb{E}\max(0, A, \widetilde{\Omega}_2)\Big)} \tag{2.18}
$$

[25] In a model with more than three periods, the length of employment protection also does not directly influence the flow utilities. Thus, following the statement of Fang & Wang (2015), utility functions can also be recovered, using observed choice probabilities. 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; thus it is orthogonal to these preferences.

Note that the numerator of equation (2.18) is equal to zero if the choice probabilities of both groups are identical, meaning the difference in the expected utilities of the two groups does not affect current choices. In other words, individuals are myopic, i.e. β or δ have to be zero in the discussed model, which is also the only case when the left-hand-side of equation (2.18) becomes zero. Overall, the numerator is expected to be positive, since we observe more individuals choosing employment at the end of the employment protection period than in its middle (see figure 2.2).
The first difference in the denominator is zero, since no job offer is needed when previously employed and, thus, there is no value of an additional period of employment protection. The second difference in the denominator is negative, since the expected maximum, when being non-employed, is smaller for individuals without employment protection. To be able to choose from all possibilities in the choice set, individuals of group B have to rely on a job offer, while individuals of group A do not. The better employment opportunities for group A result in a higher value for the $\mathbb{E}\max$. With the positive value of the numerator, the right-hand-side of equation (2.18) should be positive. Furthermore, the stronger the reaction to the end of the employment protection, i.e. the more mothers of group B return to employment in period 2, the greater the product $\beta\delta$. This is intuitively sound, since a stronger reaction to the future ability to choose from all alternatives in the choice set implies a higher overall discount factor.
Having identified all values for $\mathbb{E}\max(l_2, \Omega_2)$ and the product of $\beta\delta$, it is possible to identify all flow utilities for period two with the same approach, utilities in period three are identified. For the isolated identification of δ, choices in the first period are crucial:

$$
\begin{aligned}
U_1^A =& v(l_1; \Omega_1) + \beta\delta\mathbb{E}\left[\ln\left(\sum_{l_2} \exp(u(l_2, \Omega_2) + \delta\,\mathbb{E}\max(l_2, \Omega_2, jp_3))\right)\middle| l_1, \Omega_1\right] \\
U_1^B =& v(l_1; \Omega_1) + \beta\delta\mathbb{E}\left[\ln\left(\sum_{l_2} \exp(u(l_2, \Omega_2) + \delta\,\mathbb{E}\max(l_2, \Omega_2, \overline{jp}_3))\right)\middle| l_1, \Omega_1\right]
\end{aligned}
\tag{2.19}
$$

Like before, a difference in a logarithm of observed choice probabilities can be derived:

$$\begin{aligned}
&\frac{1}{\beta\delta}\left(\ln\left(\frac{\Pr(l_1=1|B,\widetilde{\Omega}_1)}{\Pr(l_1=0|B,\widetilde{\Omega}_1)}\right)-\ln\left(\frac{\Pr(l_1=1|A,\widetilde{\Omega}_1)}{\Pr(l_1=0|A,\widetilde{\Omega}_1)}\right)\right)= \\
&\quad=\mathbb{E}\left[\ln\left(\sum_{l_2}\exp(u(l_2,\Omega_2)+\delta\,\mathbb{E}\max(l_2,A,\Omega_2))\right)\middle|\,0,A,\widetilde{\Omega}_1\right] \\
&\quad-\mathbb{E}\left[\ln\left(\sum_{l_2}\exp(u(l_2,\Omega_2)+\delta\,\mathbb{E}\max(l_2,B,\Omega_2))\right)\middle|\,0,B,\widetilde{\Omega}_1\right] \\
&\quad-\mathbb{E}\left[\ln\left(\sum_{l_2}\exp(u(l_2,\Omega_2)+\delta\,\mathbb{E}\max(l_2,A,\Omega_2))\right)\middle|\,1,A,\widetilde{\Omega}_1\right] \\
&\quad+\mathbb{E}\left[\ln\left(\sum_{l_2}\exp(u(l_2,\Omega_2)+\delta\,\mathbb{E}\max(l_2,B,\Omega_2))\right)\middle|\,1,B,\widetilde{\Omega}_1\right]
\end{aligned} \tag{2.20}$$

The left-hand-side elements are all identified in equation (2.20). As we previously argue, the utility values and the future expected maxima for period two are also identified, leaving parameter δ as the lone value not identified on the right-hand-side of the equation. Therefore, it can be recovered from the data using the observed choice probabilities from the first period. In combination with the separate identification for the exponential discount factor δ, the identified product $\beta\delta$ also identifies the present-bias parameter β.

2.5.2 Estimation

We follow a two-step procedure to estimate the parameters of our model. In a first step, we estimate the parameters of the arrival of children, childcare costs and the job destruction rate.[26] Appendix A.1 discusses the estimation of various parts of the first step in more detail.

In a second step, we use the method of simulated moments[27] to estimate the parameters for the time-preferences, the flow utility function, the wage process, and the job offer probabilities. 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 discrete changes in outcomes, which lead to discrete changes of the objective function. Thus, gradient-based optimization algorithms are not appropriate. Instead, a pattern search method is employed, which is a derivative-

[26] 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.

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

free optimization routine. It is implemented here using the Dakota toolkit (see Adams et al., 2013), which allows for parallelization of the estimation process.
For a given set of the 27 parameter values, we solve the model as described in section 2.2.3. Afterwards, we 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 from the periods during 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. The simulated data set is used to compute the moments. To calculate the objective function $g(\Theta^b)$, the squared distances between the simulated and observed data moments is divided by the sample variance of the respective moments and summed up:

$$g(\Theta^b) = \sum_{k=1}^{K} \left[\frac{\left(M_k^O - M_k^S(\Theta^b)\right)^2}{\mathbb{V}\mathrm{ar}(M_k^O)} \right] \tag{2.21}$$

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 AppendixA.2. Note that we do not use the optimal weighting matrix due to its poor small-sample properties (Altonji & Segal, 1996). Standard errors of Θ are estimated following Gourieroux et al. (1993).

2.6 Results

In this section, we discuss the estimated parameters of the model. Overall, the estimation results show the expected picture.[28] Table 2.3 reports the estimates for the utility function. With utility from leisure for non-employment normalized to one and with the CRRA part of the utility function being negative, the positive values for the general part-time and full-time parameters indicate that low educated women treat leisure as a normal good. Both parameters are very imprecisely estimated for women with at least some college education, which might be caused by their rather small overall number of observations. Somewhat surprisingly, working preferences fade as children become older. Table 2.4 reports the estimates for the wage function, the human capital process, and the job offer rate. Without any on-the-job human capital, low educated

[28] We set the parameters of relative risk aversion to 2.0 for the reported results.

Table 2.3: Utility function parameters

	Low education		High education	
γ_{PT}	2.022152	(1.573594)	−0.938837	(535.851131)
γ_{FT}	2.278973	(1.227480)	2.239022	(18.263409)
$\gamma_{PT,AC0}$	−1.088764	(0.073510)	−1.184388	(0.136607)
$\gamma_{PT,AC1}$	0.786549	(0.073608)	0.514872	(0.114318)
$\gamma_{PT,AC2}$	0.023861	(0.018758)	0.024042	(0.031928)
$\gamma_{FT,AC0}$	−1.240039	(0.086890)	0.446021	(0.146263)
$\gamma_{FT,AC1}$	0.513382	(0.084350)	0.058413	(0.144747)
$\gamma_{FT,AC2}$	0.041831	(0.019271)	0.048157	(0.027279)

women receive €6.37 per hour on average, which corresponds to $9.71 in 2016. Women with at least some college education receive an average hourly wage of €9.79 ($14.92), when having no on-the-job human capital. At the beginning of the working life, an additional year of full-time employment results in a wage increase of 18.34% for low educated, and 17.64% for high educated women. These numbers drop to 1.57% and 0.69% respectively, after ten consecutive years of full-time employment. With a yearly rate of 15.6%, human capital depreciates much stronger for women with at least some college education, compared to the yearly rate of 5.0% for lower educated women. The probability of receiving a job offer within a year is around 23.2% for low educated, and 21.7% for high educated women. These numbers are comparable to Haan & Prowse (2015), who estimate a retirement model with SOEP data, which also includes men. On average women need about three and a half year to enter employment, which might contribute to longer career breaks even in the absence of present-bias. Previous studies overlook this aspect.

Finally, table 2.5 reports the estimates for the discounting parameters. We find ev-

Table 2.4: Employment related parameters

	Wage and human capital			
	Low education		High education	
γ_w	6.371972	(0.110681)	9.791208	(0.408616)
$\gamma_{w,e}$	0.245181	(0.018509)	0.241587	(0.065249)
η	0.025749	(0.014658)	0.081735	(0.058509)
σ_{xi}		0.020323	(0.003070)	
	Labor market frictions			
	Low education		High education	
π^{JO}	0.123742	(0.004484)	0.114880	(0.013150)

idence for time-inconsistent behavior. While the yearly exponential discount factor of 0.9173 is close to typical values found in the literature, we find evidence for present-bias with an estimate of 0.7707 for β. Although the present-bias parameter is not estimated with high precision, we can clearly reject the null hypothesis of $\beta = 1$ at confidence level of 5%. Both values imply a considerable larger discounting of future utility streams than usually assumed in the literature. Comparing our estimates with Fang & Silverman (2009), who estimate β around 0.34 and δ around 0.87, we find larger values. The authors note that their findings are lower than most of the results found in the literature. We think that because we include labor market frictions, our estimations are expected to be higher, since these can result in lower employment rates after the birth of children, one of the identifying correlations in Fang & Silverman (2009).[29]

Figure 2.3 compares the weight of future utilities in the decision process of a 20-year

Table 2.5: Half-yearly discounting parameters

β	0.770732	(0.110038)
δ	0.957770	(0.051144)

old women, between our model estimates (blue line) and an exponential discount factor of 0.95 (black line). Since our estimated parameter for the exponential discounting is close to the square root of 0.95, differences in the discounting of two future periods are similar in both models. Thus, the blue line is mostly shifted downwards, with the size of the shift determined by the estimated present bias of 0.77. This finding is interesting, because it clearly shows that those women are not low discounters per se, such that their behavior can be modeled as exponential discounting with a low discounting parameter. Rather, utilities in the far future are still important for the current period's choice, although, overall, a stronger weight is placed on instantaneous utility. Thus, our hypothesis of time-inconsistent behavior from subsection 2.3.2 is verified by the structural estimation.

2.7 Conclusion

In this chapter, we test for time-inconsistent behavior. We exploit exogenous variations in the duration of employment protection for mothers to identify time preferences in a dynamic model of female labor supply with labor market frictions. Employment protection provides insurance against labor market frictions for mothers after parental

[29]Chan (2017) only estimates fractions of the population who exhibit present-bias behavior. He finds that around 25% do so. Without knowing the exact distribution, it is not possible to compare his results with our findings.

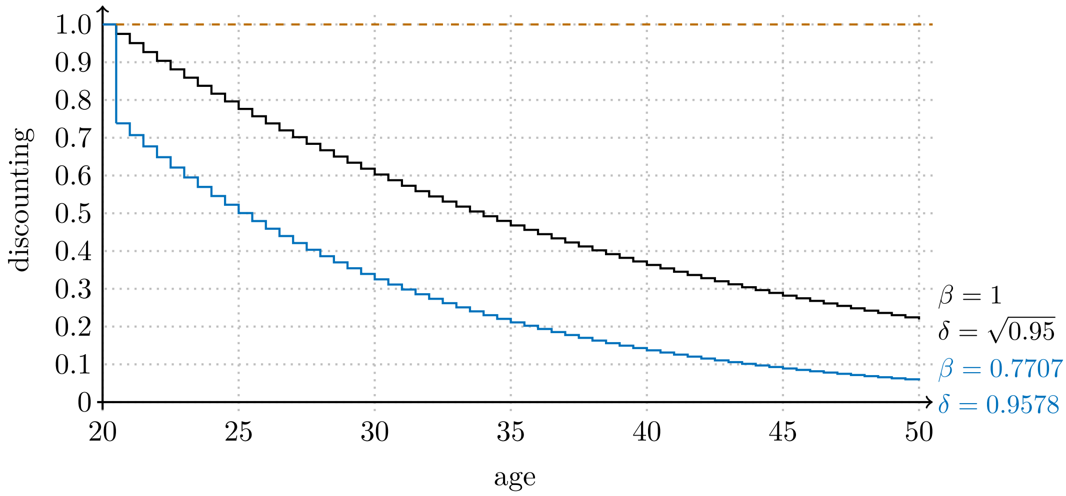

Figure 2.3: Discounting over the life-cycle

leave, thereby influencing the labor supply choices of women. Crucially for our identification strategy, employment 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, childcare support or pension benefits.

Chapter 3

Identifying and Estimating Beliefs from Choice Data – An Application to Female Labor Supply

3.1 Introduction

It is well known that beliefs are often systematically biased. Drivers have upwards biased beliefs of how safe they drive, students have biased expectations of their humor, grammar, and logic skills with respect to others, while finance professionals overestimate the precision of their stock market predictions.[30] In the labor market, individuals might have biased expectations of their future employment prospects. Career costs caused by overestimating future employment prospects can be especially high for mothers who interrupt their working careers after giving birth to children. If mothers are too optimistic about their employment prospects, they might not return during maternity leave, a period when a return to their previous job is guaranteed, although optimal under unbiased expectations. After maternity leave, mothers have to rely on new employment offers that arrive with a lower probability as anticipated. Therefore, overestimating future employment prospects prolongs child-related employment interruptions, increasing the career costs of having children.[31] In contrast, if mothers underestimate their future employment opportunities on average, more women return to employment during maternity leave than under rational expectations, reducing the costs of motherhood. Previous literature acknowledges the importance of potentially biased expectations, but does not identify these within the context of child-related

[30]See Svenson (1981), Kruger & Dunning (1999) and Ben-David et al. (2013). (De Bondt & Thaler, 1995, p. 389) even conclude, “perhaps the most robust finding in the psychology of judgment is that people are overconfident.”

[31]Children are one important factor for the career dynamics of women. This is reflected by the average employment rates across OECD countries, which are 11 percent lower for women with at least one child (aged 0-14) than for women without a child in 2014. The career costs of having children can be high, for instance, Adda et al. (2017) estimate that fertility reduces the net present value of income by 35%, of which they attribute 76% to the lower employment of mothers.

career breaks or quantifies their consequences for the working careers of mothers.

In this chapter, I develop a life-cycle model of female labor supply and human capital accumulation, derive a strategy to identify job offer expectations within this model, and quantify the career costs of biased expectations of future employment prospects. To estimate the model, I use survey data from the German Socio-Economic Panel Study (SOEP), since German maternity leave regulations provide an ideal environment to identify the key parameters of the model. The identification approach exploits the impact of expectations on the decision process at the end of maternity leave, which provides mothers with a guarantee to return to their previous job. This change from an employment guarantee to a situation in which individuals have to rely on job offers to leave non-employment creates a discontinuity in the future expected value of non-employment that varies with job offer expectations. To separately identify expectations, preferences, and real job offer rates, several maternity leave reforms are exploited, which change the length of the employment protection.

Expectations of future employment opportunities are modeled, such that rational expectations are nested, which is the predominant assumption when estimating life-cycle models of female labor supply. This allows for directly testing if beliefs about future employment opportunities are biased. In a second step, I quantify the life-cycle costs of these biased expectations. Holding the preference parameters constant, but restricting expectations to be rational, I simulate life-cycle choices and compare them with actual observed choices. I can, thus, examine the welfare costs of wrongly estimating future employment possibilities.

In the first part of this chapter, I develop a life-cycle model of female labor supply and human capital accumulation (see for example Keane et al., 2011).[32] In this model, women choose their labor supply in half-yearly intervals, facing labor market frictions. To enter employment after being non-employed, women need to receive a job offer, which arrives with a probability related to their age and human capital. In contrast to the standard life-cycle framework, which assumes that individuals have rational expectations, I explicitly model expectations of the future job offer arrival probability. These expectations are especially crucial when deciding whether or not to return within the employment-protected maternity leave period. If an individual overestimates her future employment opportunities, she might not return during the employment protection. Having overestimated her chances to receive a job offer, the career break is, on average, longer than she expected. Because on-the-job human capital depreciates when

[32]Research employing a life-cycle model of female labor supply include, for example, Blundell et al. (2016) who focus on how welfare reforms within a life-cycle model of labor supply and human capital, Adda et al. (2017) who evaluate how fertility influences occupational and employment choices over the life-cycle, and Low et al. (2010) who investigate the influence of different types of risks that individuals face over their working life-cycle.

non-employed, longer non-employment spells are not only more costly because of the lost income from employment, but also because of the losses in human capital.

In the second part of this chapter, I present a novel approach to identify the job offer expectations within the discussed life-cycle model using only choice data. The approach relies on the impact of expectations on the decision process of returning to employment at the end of maternal leave. If mothers expect the arrival rate of future job offers to be very high, the future expected cost of not returning to their protected job are low. If, in contrast, the expectations of the future job offer probability are rather low, the future expected costs, in terms of lost income and human capital, are relatively high. Essentially, the lower the expected job offer rate, the higher the expected costs of not returning to the guaranteed job and the higher the probability of women returning at the end of the employment protection. Therefore, the mass of mothers returning directly at the end of the employment protection is at least partly a result of the mothers' expectations of their future employment opportunities.[33]

Although expectations influence the returning behavior of mothers at the end of their employment protection period, there are other factors that might drive returning behavior. To control for other potential influences, I exploit several reforms of the German maternity leave regulations. These reforms first extend the employment protection period from 1 year to 1.5 years, and then to 3 years. The three policy regimes create three groups of individuals facing different lengths of employment protection when the youngest child is between 1 and 1.5 years old. Employment rates of mothers in the regime with the longest lasting employment protection aid the identification of leisure preferences regarding the age of the youngest child. Their returns to employment around 1.5 years after childbirth are not influenced by the expectations of future employment possibilities, but are solely driven by leisure preferences. Comparing the mass of returning mothers shortly before the youngest child turns 1.5 between the regime that grants mothers 3 years of employment protection and the regime that grants mothers 1.5 years, identifies the excess mass due to expectations of future job offers. Having identified preferences and expectations, it is possible to use the non-employment to employment transitions of mothers with children older than 1.0 of the regime with the shortest employment protection to identify the real job offer rate.

[33] The identification approach has some similarity to the literature using observed bunching for the identification of elasticities (see Saez (2010) and Kleven & Waseem (2013)). The classical bunching approach would use a kink or notch in the tax schedule to recover underlying labor supply elasticities. In contrast, this chapter uses a discontinuity over time in the guarantee of returning to a previous job. Additionally, the counterfactual situation of not having this discontinuity is available, since policy reforms prolonged the length of employment protection over the years. Another paper using bunching to identify welfare costs of behavioral elements instead of elasticities is Rees-Jones (2018). He quantifies the tax evasion costs introduced by loss-aversion when individuals owes taxes at the end of a tax year instead of receiving a refund.

In the third part of the chapter, I estimate the model and quantify the career costs of overconfidence. Individuals receive full-time employment offers with around 50% in a year, and part-time employment offers with around 17%, but these probabilities decrease with the time spend in non-employment. Individuals display a strong bias in their expectations, and anticipate the job arrival rate to be 66% higher compared to the real rates, on average. These findings are in line with the suggestive evidence constructed from several questions of the SOEP questionnaire about future employment expectations. Simulating the model once with the estimated expectations and once with rational expectations allows for quantifying the costs of overconfidence. Under biased expectations, child related career breaks are, on average, 6 months longer. Women lose between 12% and 18% of the net present value of earnings from employment. The net present value in household consumption is much lower, lying between 3% and 4%. There are two main reasons for this difference. First, partners contribute the larger share to the overall household income, since they are mostly working full-time (and do not interrupt their career due to childbirth), while mothers re-entering the labor market typically work part-time. Second, the German tax system, with its joint taxation system, heavily taxes second earner's income. The simulations also show that the costs of overconfidence decrease with the length of the employment protection.
The life-cycle loss in earnings from employment due to biased expectations are meaningful from a public economics perspective. They resemble losses in income taxation in addition to the possible social security provided to mothers who have not returned to employment due to their biased expectations. In addition, the consequences for the individual are substantial: The lost lifetime earnings translate into lower pension benefits making them more vulnerable to poverty in retirement. The consequences might justify interventions by policy makers. Potential policies might provide more information about employment prospects after child-related career breaks, for instances by introducing mandatory consulting meetings with an employment agency. Other measures might include financially incentivizing returning to work within the employment protection, for example by providing in-work benefits toward the end of the employment protection.

Contribution to the Literature

The two major contributions of this chapter are, first, identifying beliefs from actual employment behavior over time and, second, estimating the life-cycle costs of biased beliefs. These two contributions connect the behavioral literature and the literature on life-cycle labor supply. I extend the behavioral literature, which predominantly derives its empirical evidence from specially designed experiments, by finding evidence of biased beliefs in the context of labor supply choices over time. Furthermore, I can estimate the long-term consequences of the bias, which is only possible to a limited extent in an experimental setting. The contribution to the literature on life-cycle employment behavior comes from allowing non-rational expectations. Deviation from ration expectations are generally ignored in this literature, often because an identification strategy is not available.

In general, there is little evidence of biased beliefs and its consequences using a revealed preference approach outside of laboratory experiments. Since the work of Tversky & Kahneman (1974), which introduces a theory for their common finding that individuals exhibit systematic biases when acting under uncertainty, the literature on social psychology and organizational behavior[34] intensively analyzes overconfidence. An introduction into the literature's link to economic questions is provided by Malmendier & Taylor (2015).[35] The majority of the findings stem from experiments, since most surveys capture expectations too broadly to provide convincing evidence on overconfidence. Although laboratory experiments are ideal for exploring behavior and testing possible theories about the decision making process, these might not be well suited to quantify real economic consequences. This chapter closes this gap by identifying expectations from observed choice data within a life-cycle model of labor supply.

The labor economics literature investigating expectations outside experiments can be divided into two parts.[36] One part uses subjective data in reduced form analysis to determine the impact of expectations on labor outcomes. Most of this research investigates how future earnings and labor market attachment expectations influence education and other investment in human capital decisions. For example, Sandell & Shapiro (1980) and Shaw & Shapiro (1987) show that individuals who do not expect strong future labor attachment, invest less in human capital than individuals with stronger expected attachment. This is further underlined by Gronau (1988) and Blau & Ferber (1991). The other part of the literature concentrates on testing more directly if expectations are unbiased by comparing surveyed expectations with actual behavior. For instance, Hamermesh (1985), Bernheim (1988), and Hurd et al. (2004) find individuals

[34] Moore & Healy (2008) survey this literature.

[35] For seminal work, see, for example, Svenson (1981) who finds that 83% of participants in a laboratory experiment stated that they are in the top 30% regarding driving safety, Kruger & Dunning (1999) who find that students who scored in the bottom quartile (and thus find themselves in the 12*th* percentile, on average) in tests regarding humor, grammar, and logic skills, believe themselves to be in the 63*rd* percentile of the distribution, and Ben-David et al. (2013) who show that only 36.3% of the time the S&P500 falls into the 80% confidence interval provided by CFOs of mid-size and large U.S. corporations. Further examples include Weinstein (1980) and Slovic (2000). The literature mainly uses three definitions of overconfidence: (1) the overestimation of the probability of positive events; (2) the overestimation of one's performance compared to others; and (3) the overestimation of the precision of one's information. The model and identification approach of this chapter correspond to the first definition.

[36] An exception is the work by Attanasio et al. (2017), who estimate Euler equations for consumption using subjective expectation data.

are mostly able to predict their retirement age.

The majority of these studies use questions only allowing for yes-no answers to elicit expectations. Manski (1990) shows that even in the absence of aggregate shocks, binary expectations questions are ill-equipped for investigating the hypothesis of rational expectations.[37] In addition, nearly all these questions mix pure expectations of exogenous events with preferences that prevent a clear distinction between these two factors. In contrast, the model and identification strategy presented in this chapter do not rely on questions to elicit expectations and, therefore, does not suffer from these problems. It also allows clearly differentiating between biased expectations of exogenous future events and changes in preferences.

A stronger focus on biased beliefs of future employment prospects represents the work of Spinnewijn (2015). He examines the optimal unemployment insurance design when job seekers overestimate their chances of finding employment. In addition to a theoretical analysis of how to adjust the Baily formula for optimal unemployment insurance (Baily, 1978; Chetty, 2006) in the presence of overconfidence, Spinnewijn (2015) calibrates a job search model with various degrees of biased expectations. He finds that overconfident agents are less responsive to future incentives and shows that it can be optimal for unemployment benefits to increase over time. Complementing this work, this chapter concentrates on the individual career costs of mothers in a life-cycle framework. Since the majority of female career breaks are family-related, adjusting maternity-leave policies might be more effective for women than adjusting unemployment insurance.

Another empirical investigation of overconfidence and its consequences in labor supply contexts is Hoffman & Burks (2017). They investigate the overconfidence in productivity by truck drivers, finding it contributes to fewer employees quitting. Overall, this causes welfare to increase, since the companies face large initial training costs when hiring new drivers. I extend this research by discussing the effect of overconfidence on the career development of mothers. In contrast to Hoffman & Burks (2017), my results indicate that there can be substantial costs when individuals are too optimistic about their future employment possibilities.

This chapter also contributes to the literature focusing on employment and maternal welfare in a life-cycle context. Adda et al. (2017), using a life-cycle model of occupational choice, find that family-oriented women already choose occupations that are family friendly but not necessarily well paid. They estimate the cost of having children

[37] A short example should illustrate this statement. Assume a single event A occurs with the probability of 51%. If the event is realized, a subject will work the next period, otherwise she will spend time in home production. If asked if they will expect to be working next period, all subjects will answer with “yes,” since “no” is more unlikely. On average, this results in a discrepancy between the stated expectations and realizations of 49 percentage points. For a more general discussion of the importance of expectations in economics and their measurement see Manski (2004).

to be about 35% of lifetime income. Some of these costs also stem from lost earnings and depreciation of human capital during career breaks. Blundell et al. (2016) estimate a model of human capital accumulation and depreciation that points to very low human capital accumulation in part-time employment and, therefore, stagnating careers for mothers who tend to work less than full-time. Employing a similar model of life-cycle labor supply, I extend their findings by dividing career costs into expected and unexpected ones. While anticipated career costs do not necessarily justify policy interventions when markets are close to perfect, biased expectations can be regarded as market imperfections and, thus, make a stronger case for additional regulations. An example of a more harmless intervention is the direct provision of information, for example, in the form of letters. These seem to work well in some fields of public economics (see for example Bhargava & Manoli (2015), who investigate welfare take-up, Duflo & Saez (2003) who investigate retirement decisions in the US, and Dolls et al. (2016), who investigate retirement decisions in Germany).

Finally, this chapter adds to the growing literature of behavioral public economics. Because optimal policy design depends on the behavior of individuals, ignoring behavioral insights may lead to ineffective policy recommendations. Some behavioral insights can also lead to more efficient policies, such as providing additional information or commitment devices, which might otherwise have been ignored. An example in the context of labor supply is DellaVigna et al. (2017), who exploit a reform of the unemployment benefit system in Hungary, showing that job seekers have reference-dependent preferences. They argue that in this case a multi-step unemployment insurance is optimal. Other examples are DellaVigna & Paserman (2005) and Chan (2017), who investigate time-inconsistent preferences in the form of hyperbolic-discounting. The former find that measurements of the impatience of job seekers and their respective unemployment lengths are in line with the hyperbolic-discounting model. Chan (2017) identifies discounting parameters using data from a field experiment. He finds evidence for a welfare-trap: individuals, who are not currently employed, postpone their decision to start working due to time-inconsistent behavior. I extend this literature by determining how expectations might contribute to the length of non-employment durations.

The chapter proceeds as follows. Section 3.2 discusses the institutional framework. Section 3.3 describes the data. Section 3.4 presents some descriptive characteristic of the data and provides suggestive evidence for the biased expectations of future employment prospects. Section 3.5 develops the structural life-cycle model. Section 3.6 discusses the identification and estimation of the model parameters, in particular the identification of beliefs. Section 3.7 presents the results and discusses their implications. Section 3.8 concludes.

3.2 Maternity Leave Policy in Germany

German maternity leave regulations provide an ideal setting for the identification of expectations of future employment possibilities. Several policy reforms extended the period granting mothers the right to return to their previous work positions, which provides exogenous variation for identification. In total, I exploit multiple major expansions of maternity leave coverage between 1986 and 1993, which I summarize as three major policy regimes.[38] The objective of these reforms was twofold. First, they intended to encourage mothers to spend more time with their children during their early development. Second, they aimed to increase maternal labor market attachment, since longer employment protection was viewed as an instrument to ease returning to the labor market. Since the identification approach relies on the exogenous variation created by these reforms, a more detailed discussion of the maternity leave system and its changes is discussed in the following.

Starting in the late-1960s and through 1986, mothers were entitled to 14 weeks of paid leave around childbirth, during which women were generally not allowed to work. While on leave, employers could not dismiss mothers and had to provide a comparable job to the previously held position for women returning within leave. During the 14 weeks, mothers received their average income of the three months before entering maternity leave, resulting in an income replacement rate of 100%. The core of this law is still effective in 2017,[39] with later reforms mainly changing the regulations after 14 weeks. In the late-1970s, the first major reform extended maternity leave coverage, lengthening the employment protection period to six months after childbirth and introducing a new maternity leave payment for the time between the end of the 14$^{\text{th}}$ week and the end of the 6$^{\text{th}}$ month. In this period, women, who were employed before having a child, received DM 750[40] per month.

Reforms used for identification started in 1986, with table 3.1 providing an overview. The first reform expanded the employment protection and maternity benefit period from six to ten months at the beginning of 1986, then extending it to 12 months in January 1988.[41] Maternity payments from week six to week eight remained at an income replacement of 100% or DM 600[42] if the mother was previously unemployed.

[38]The summary of the parental leave reforms through 1985 are mainly based on Zmarzlik et al. (1999). For later reforms, see Bundeserziehungsgeldgesetz [BErzGG] [Federal Child-Raising Benefit Act], Dec. 6, 1985, BGBl.I at 2154 (F.R.G.) and its changes through its abolition in 2007. This chapter concentrates on West Germany.

[39]Minor reforms specified more precisely the conditions under which mothers are allowed to work during this period.

[40]This is equivalent to $ 758 in 2017.

[41]Additionally, paternity leave was introduced. However, between 1987 and 1994 only about 1% of fathers took parental leave (Vaskovics & Rost, 1999).

[42]This is equivalent to $ 606 in 2017.

Between three and six months, maternity benefits declined from DM 750[40] to DM 600[42] 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 during the two years prior to childbirth. Around 84% of individuals were eligible for the full benefits (Schoenberg & Ludsteck, 2014). I summarize these conditions as forming policy regime I, which provides one full year of employment protection and maternity benefits.

Table 3.1: Parental leave reforms from 1986 until 2006

	Month, Year (1)	Emp. Prot. (2)	Maternity Benefits (3)
Regime I	January, 1986	10 months	3-6 months DM 600,[42] 7-10 months means tested
	January, 1988	12 months	up to 12 months, means tested
Regime II	July, 1989	15 months	up to 15 months, means tested
	July, 1990	18 months	up to 18 months, means tested
Regime III	January, 1992	36 months	up to 18 months, means tested
	January, 1993		up to 24 months, means tested
	January, 2007	maternal benefits are related to previous earnings	

A further expansion of employment protection and the maternity benefit period from 12 months to 15 months took effect in July 1989, followed by another extension to 18 months in July 1990. These reforms are summarized in regime II, providing 1.5 years of employment protection and maternity benefits. In January 1992, the employment protection period was further extended to a total of three years. In contrast, the maximum maternity payment period initially remained constant at 18 months, before being extended to two years a year later. Minor changes in family policy were introduced in 2001, but the core regime of 1993 still continued.[43] This forms regime III, which provides 3 years of employment protection and 2 years of maternity benefits. These policies did not noticeably change, before a major reform of maternity benefits in 2007, which changed benefits from a lump-sum payment to an income replacement based on pre-birth earnings.

[43]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 receiving the full benefits during the first six months.

3.3 Data

For estimations, I use longitudinal data from the German Socio-Economic Panel Study (SOEP) covering 1986 through 2006 (see Wagner et al., 2007, for a description of the SOEP).[44] Since 1984, the SOEP interviews private households and persons on an annual basis in Germany. Each year, all household members older than 16 are interviewed, conditioned on their voluntary collaboration. New additions to a household, including partners and children are added and remain in the sample, even if they leave their first registered household. The base SOEP sample was expanded with several booster samples over the years and questionnaires cover a wide range of topics, including details on demographics, education, labor market dynamics, earnings, and other income, among others.

While the SOEP interviews individuals every year, it asks participants to fill out a monthly calendar for the previous year. In particular, individuals are asked about last year's employment history, which allows me to construct a semi-annual data set by combining the current year's questionnaire with information from the questionnaire of the following year. A particularly useful feature is that the SOEP has a possible option "maternity leave" for the monthly employment state. Emphasizing that employment protection is a well-known and well-understood policy, almost all previously employed mothers consistently make use of this option. The SOEP also asks newly surveyed individuals, older than 16, to fill out a specific questionnaire collecting information about their life before they were included in the sample. This enables me to collect information on each individual's age when finishing their education and starting their work life.

I restrict the sample to women and, when applicable, their partners between the age of 18 and 50,[45] having no university degree after finishing education. The education limitation is mainly due to the small number of highly educated women in the sample having children around the time of the maternity leave reforms used for identification. Because some reforms took place before the reunification of Germany, I exclude individuals living in East Germany. Additionally, I exclude self-employed women and women working in the public sector. The reforms do not impact self-employed individuals, because they do not have an employer who has to guarantee them their job for a given period. In contrast, individuals working in the public sector might enjoy more generous maternity leave conditions, especially longer periods of employment protection.[46] For

[44] I also use the wave corresponding to 2007, since it includes responses addressing 2006.

[45] Some exogenous processes estimations, include women until the age of 70 in order to have more robust estimates for the later years.

[46] Note that to qualify for the more generous maternity leave conditions in the public sectors, individuals must to have been employed regularly in this sector directly before having their child. Therefore,

similar reasons, women who are subject to least one of the following criteria are also not included in the sample: living outside Germany, being severely disabled, or having at least one multiple birth. Missing information on the age of their children, their labor market entry age, or their labor market experience also leads to exclusion.

Some further data cleaning and labeling is worth highlighting. The labor market experience for a given year is constructed by combining the answers of a working history questionnaire and the recorded employment state of follow-up interviews. Part-time and full-time experiences are separately measured. Wages are defined as gross monthly earnings divided by actual working hours during the same period. Since the model does not include any macro economic processes, all money values are deflated using the Consumer Price Index and the base year 2000 (Federal Reserve Bank of St. Louis, 2016). To reduce the importance of measurement error in wages, the wage distribution is trimmed at the 4^{th} and 98^{th} percentiles, from below and above, respectively.[47]

The resulting data set is an unbalanced panel in which individuals enter and leave the panel at various ages. In total there are $3,944$ women. Approximately 55% of these are observed for more than 5 years, about 20% for more than 10 years. Additionally, I observe $1,510$ births and a total of $2,934$ children aged 18 or younger. In total, the sample has $49,924$ semi-annual observations. Table 3.2 shows the distribution of family types for three age groups. Women tend to get married before turning 30. This seems to be an important factor for having children, since the number of single mothers is rather low, with a peak of 8% at age 35. At this age, 83% of women have had a child, potentially interrupting their working career. Ninety percent of the mothers are married and, thus, might be able to rely on their husband's income.

Table 3.2: Distribution of family types at different ages

	Mothers		Non Mothers		Number of
	Singles	Marriage	Singles	Marriage	Observations
	(1)	(2)	(3)	(4)	(5)
women aged 25	0.04	0.38	0.24	0.33	937
women aged 30	0.06	0.65	0.09	0.20	925
women aged 35	0.08	0.75	0.06	0.11	923

Finally, since the identification exploits the different maternity leave regimes, an overview of the number of observations in the three regimes is helpful. It is provided in table 3.3. Although the SOEP is a survey and regime II was only in place for 2.5 years,

mothers are not able to receive these more generous conditions when switching sectors *after* having had their child.

[47] After trimming, the lowest hourly gross wage is €4.21 and the highest gross wage is €25.72.

I observe 110 distinct women with a child under the age of 3 for the second regime. For these women, 999 labor supply decisions are recorded. For regimes I and III, I observe 359 and 854 women with young children, respectively. These make $1,395$ and 6607 labor supply choices.

Table 3.3: Observations per regime

	women (1)	decisions (2)
Regime I	359	1395
Regime II	110	999
Regime III	854	6607

Notes: Column 1 represents the number of women observed in the respective regime who have a child under the age of 3. Column 2 represents the number of decisions observed for these women.

3.4 Suggestive evidence for overconfidence

The SOEP questionnaire allows for a first test of biased beliefs of future employment opportunities. Since 1999, all non-employed subjects who indicated that they might seek employment in the future are requested to state probabilities of future life events, including the likelihood of being employed within the next two years.[48] Knowing the month of the interview, it is possible to track individuals who stated a probability to enter employment and determine if they found employment within two years. Comparing the average stated probability and the average of individuals who have found employment provides a first estimate of a bias in expectations. As table 3.4 shows, individuals systematically overestimate their future attachment to the labor market.[49] It is important to note that due to the limited number of mothers answering these questions, the table includes also women without children.

The first column shows the average stated likelihood and the average actual realization of the whole sample. Columns two to four list the values for women who stated a likelihood of being in employment within the next two years of at least 30%, 50%, and 80%, respectively. To list values of these groups seems important, since Kassenboehmer & Schatz (2017) show that low stated probabilities are primarily driven by individuals who are long-term unemployed and lost faith in their ability finding employment.

In table 3.4, the average stated probability is provided in row 1, the actual realization of

[48] For the exact questions, see appendix B.1.1.

[49] See also Kassenboehmer & Schatz (2017) who find similar results in a sample including men.

the same individuals in row 2. Row 3 states by how much individuals overestimate their chances to find employment within two years. Row 4 provides the p-value of testing the hypothesis that the difference between row 1 and 2 is zero, while the last row reports the number of observations for each respective group. The critique of Manski (1990) that standard expectation data is not well equipped to investigate the question of rational expectations does not apply in this case, since individuals are explicitly asked to state the probability using a Likert-type scale with a range of 11 values. Manski's critique is mostly applicable to questions with binary answers, since these are more challenging for estimating underlying probabilities.[50]

Table 3.4: Employment expectations and realizations

	Average (1)	Stated >= 30 (2)	Stated >= 50 (3)	Stated >= 80 (4)
Stated	45.44 %	66.55 %	72.89 %	92.75 %
Actual	35.95 %	45.93 %	48.75 %	58.78 %
Overestimation	26.40 %	44.89 %	49.52 %	57.79 %
p-value	0.0000	0.0000	0.0000	0.0000
Observations	1079	690	569	279

Notes: Row 1 represents the stated probability to be in employment within the next two years, row 2 states the real percentage of individuals having found employment within two years, row 3 shows by how much individuals overestimated the probability on average, row 4 denotes the p-value of the hypothesis that there is no difference between row 1 and row 2 and row 5 shows the number of observations. The original question reads as follows: "How likely is it that you will start paid work within the next two years?" Only subjects who stated that they might want to work in the future are asked this question. The answers are recorded on an 11-point Likert-type scale from 0 to 100 percent. Individuals for whom I can neither observe the length of their unemployment spell nor that they were unemployed for more than two years are excluded. Survey weights are used.

In addition, the data is collected between 1999 and 2006, a period during which Germany's unemployment rate did not fluctuate much, staying between 9% and 12% percent. Two years in which the question about future employment expectations was asked were followed by a decline in the unemployment rate, while the other two years were followed by a recession.[51] Thus, it is plausible to assume that the differences in

[50]The Likert-type scale has the range 0% to 100% in 10 percentage points steps. In the worst case, under rational expectations, individuals would estimate their likelihood only slightly above the median between two points on the scale and, thus, always choose the higher value. Hence, a deviation below 5% does not necessarily provide enough evidence to reject the null hypothesis of rational expectations. Even when subtracting 5% off all stated likelihoods, the differences stated in table 3.4 are still significant at the 1% level.

[51]For a detailed overview of the unemployment rate during the time of the interviews see appendix

expectations and realizations are not driven by aggregate shocks.
The table unveils two important aspects. First, it shows that only a minority (35.95%) of interviewed women enter employment within two years and second, individuals overestimate the likelihood of finding employment by 26%, on average. This gap between the stated likelihood and the observed outcomes widens monotonically with higher stated probabilities. For individuals who stated a probability over 50%, the average prediction is about 73%, but in reality only 49% find employment within two years, causing an overestimation of about 48%. A second aspect is that expectations are not random, as there is a positive correlation between stated expectations and realizations. Individuals who state a higher likelihood to find employment, also have a higher probability to enter employment within two years. It seems that individuals can, to a certain degree, predict their likelihood in relation to others, but systematically overestimate the likelihood of finding employment on average.

3.5 Model

Although, there is some suggestive evidence showing that women indeed substantially overestimate their future employment possibilities, it is unclear how accurate stated probabilities are, since even college students often have trouble grasping the concept of probability (Garfield & Ahlgren, 1988). Another strategy to test the hypothesis of rational expectations is to rely on a revealed preferences approach, identifying expectations from actual employment choices. To do so, a standard dynamic discrete choice model of labor supply is developed, enhanced by explicitly modelling future employment expectations. The presented model relates strongly to the model in Blundell et al. (2016), who also investigate female labor supply over the life-cycle.[52] In addition to permit identification of expectations from choice data, the model can be used to analyze the costs of biased beliefs; an exercise, challenging to perform based on the descriptive analysis presented in the previous section.

3.5.1 Outline of the model

Figure 3.1 provides an overview of the model's general life-cycle process. The main focus lies on the working life of females between the ages of 18 and 50. Since the age at

B.1.2.

[52]Besides explicitly modelling expectations, the major difference to the following discussed model is that labor market frictions in the form of job offers and layoffs are not explicitly modeled in Blundell et al. (2016). In contrast, the authors integrate a savings decision in their model. Besides Blundell et al. (2016), there is a long history for dynamic life-cycle models of labor supply, see for example Heckman & Macurdy (1980), Eckstein & Wolpin (1989), Van der Klaauw (1996), Attanasio et al. (2008), and Adda et al. (2017).

which individuals leave education is considerably heterogeneous in Germany, the model begins after the respective individual has finished education and enters the working life.

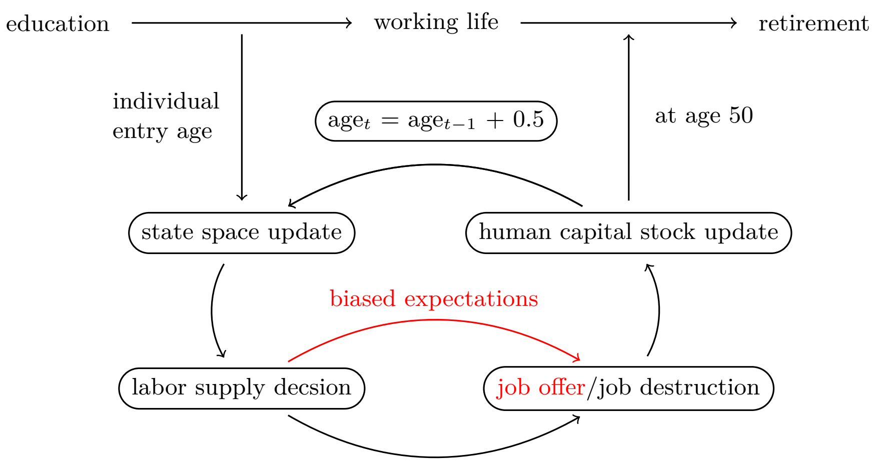

Figure 3.1: Outline of the model

As figure 3.1 shows, all women enter the model's life-cycle element after finishing their education. Each decision period lasts for six months and begins with the determination of the state space of the current period. This includes the forming and termination of marriages, the birth of children, and the realization of wage and taste shocks. Similar to Blundell et al. (2016), partnerships and children are not modeled as explicit choices, but as exogenous stochastic processes. These processes depend on the characteristics of the women, including age and current family characteristics. When individuals maximize their expected lifetime utility, they account for the possibilities of partnerships and children.[53]

After the state space for the current period is set, each woman chooses the number of hours she wants to supply for the current period. The possible supplied hours are discretized into three categories, non-employment (0 hours per week), part-time employment (20 hours per week), and full-time employment (40 hours per week). The realization of the hours choice depends on the labor market and the woman's previous employment state. If a woman was not employed in the previous period, she needs to receive a job offer if she desires to work in the current period. If a woman was employed in the previous period, she might lose her employment involuntary due to plant closure or other external factors. In this case, she can work neither full-time nor part-time

[53] Not modeling marriages and the arrival of children as choices is mainly due to the additional computational burden when doing so. An important limitation resulting from this modelling choice is that counterfactual simulations are assumed to not impact these processes. There is also a long history of modeling partners and children in this way, see, among others, Van der Klaauw (1996), Sheran (2007) and Blundell et al. (2016).

in the current period and must await the next period for possible reemployment. An important feature of the model is that women have expectations of the probability of receiving a job offer in a future period. Although these expectations can align with the true job offer probability, individuals are allowed to systematically under- or overestimate these probabilities. In contrast, all other probabilities, for instances about partnerships and children, are restricted to align with the true probabilities.

At the end of each period, the on-the-job human capital stock is updated. It is a crucial factor for determining the wage that an individual woman can realize on the labor market. All individuals start with an on-the-job human capital stock of zero since it can only be acquired through employment. Similar to Blundell et al. (2016), human capital might grow with different rates for full-time and part-time employment. Furthermore, human capital depreciates with every period spend outside of employment. Each period ends with the update of the on-the-job human capital. The outlined process repeats until the age of 50, which is the last period of the model.

3.5.2 The structural model

This subsection provides further details about the functional form assumptions. Individuals must make a labor supply decision (l_t) each period, depending on their characteristics after entering the working life phase of the model. The characteristics include the age (t), on the job human capital (e_t), the employment state of the last period (l_{t-1}), the presence of a partner (p_t), the presence of children (cd_t), age of the youngest child (ac_t), current policy regime (r_t), and the employment protection state (jp_t). In principle, they can choose between non-employment ($l_t = NE$), part-time employment ($l_t = PT$), and full-time employment ($l_t = FT$).[54]

Flow utility. The instantaneous utility of a choice depends on its consumption opportunities and its leisure time. Consumption and the utility from leisure are allowed to vary with the presence of a partner, the presence of children, and the age of the youngest child. I assume the instantaneous utility is non-separable between consumption and leisure, but utility is separable over time. The functional form is given by

$$u_{i,t} = \frac{(c_{i,t}/\bar{c}_{eq} - 1)^{(1-\gamma_c)}}{1-\gamma_C} \times \exp\left(U^L\left(l_{i,t}, p_{i,t}, cd_{i,t}, ac_{i,t}\right)\right) + \varepsilon_{i,t}, \tag{3.1}$$

where $c_{i,t}$ denotes the consumption and $\bar{c}$ an equivalence scale[55] that controls for the

[54] I assume 260 paid working days in a given year, which equals 130 working days in a half-year. Part-time employment is standardized to be 20 working hours in a week (520 hours a half-year), full-time employment to be 40 working hours in a week (1040 working hours a half-year). Both hour values are the median hours worked in the sample, when subjects stated that they are working part-time or full-time, respectively.

[55] I assume that $\bar{c} = 1$ for single women without children, $\bar{c} = 1.4$ for single mothers, $\bar{c} = 1.6$ for

members of the household. The CRRA-parameter γ_c represents the risk aversion. $U^L(\cdot)$ represents the utility from leisure, and $\varepsilon_{i,t}$ the choice-specific shock, which is independently and identically distributed over time and choices with a type-1 extreme value distribution with zero mean. The utility of leisure is normalized to 1 if the individual is not working in the current period. If working, leisure preferences vary with working hours, education, and the presence of a partner and children, as shown in equation (3.2).

$$
\begin{aligned}
U^L\left(l_{i,t}, p_{i,t}, cd_{i,t}, ac_{i,t}\right) = & \sum_{l' \in \{PT,FT\}} \left(\gamma_{l'} \mathbb{1}_{[l_{i,t}=l', cd_{i,t}=0]} + \gamma_{l',p} \mathbb{1}_{[l_{i,t}=l', p_{i,t}=1]}\right) \\
& + \sum_{l' \in \{PT,FT\}} \mathbb{1}_{[l_{i,t}=l', cd_{i,t}=1]} \left(\gamma_{l',ac_0} + \gamma_{l',ac_1} ac_{i,t} + \gamma_{l',ac_1} ac_{i,t}^2\right) \\
& + \sum_{l' \in \{PT,FT\}} \sum_{ac' \in \{0,0.5,\ldots,3.5\}} \mathbb{1}_{[l_{i,t}=l', ac_{i,t}=ac', cd_{i,t}=1]} \gamma_{l',ac'} \\
& + \gamma_{PT,state} \mathbb{1}_{[l_{i,t}=l_{i,t-1}=PT]} + \gamma_{FT,state} \mathbb{1}_{[l_{i,t}=l_{i,t-1}=FT]}
\end{aligned}
\tag{3.2}
$$

The leisure preferences, depending on the age of the youngest child, are modeled in a quadratic manner, but allowed to deviate for very young ages. This allows capturing irregularities in preferences that cause mothers to return during their employment protection and are important to control for, since the identification strategy relies on the excess mass of mothers returning due to their low expectations of future employment opportunities, when controlling for preferences. The last two terms capture a possible state dependency since, in the data, women rarely switch directly from part-time to full-time employment or vice versa.

Wages and Human Capital. The decision to work and the resulting hours choice depend on the consumption opportunities of the respective choice. Income from employment is driven by the following wage process:

$$
\ln(w_{i,t}) = \ln(\gamma_{w,\,const.}) + \gamma_{w,e} \ln(e_{i,t} + 1) + \xi_{i,t} \tag{3.3}
$$

The hourly wage rate depends on a constant and accumulated on-the-job human capital. $\xi_{i,t}$ is to be assumed a measurement error that follows a normal distribution with standard deviation σ_ξ. Wage differences over time are driven by on-the-job human

couples without children, and $\bar{c} = 2$ for couples with children.

capital. It evolves in the following manner:

$$e_{i,t} = \begin{cases} e_{i,t-1}(1-\eta) & \text{if } l_{i,t-1} = NE \\ e_{i,t-1}(1-\eta) + \lambda & \text{if } l_{i,t-1} = PT \\ e_{i,t-1}(1-\eta) + 0.5 & \text{if } l_{i,t-1} = FT \end{cases} \tag{3.4}$$

Human capital at the end of each period depends on the previous period's human capital and the employment state of the current period. In each period, the on-the-job human capital depreciates with the rate $(1-\eta)$.[56] Accumulation depends on the working hours with potentially different gains for part-time and full-time employment. Being in a model with a semi-annual decision period, the gain of full-time employment is normalized to be 0.5. The gain from part-time employment is estimated in order to not restrict the model 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 determined by:

$$\begin{aligned} c_{i,t}(l_{i,t}, p_{i,t}, ac_{i,t}) = 130 \times \Big(& w_{i,t} - \mathbb{1}_{\{cd_{i,t}\}} cc(ac_{i,t}) \Big) \\ & \times (4 \times \mathbb{1}_{\{l_{i,t}=PT\}} + 8 \times \mathbb{1}_{\{l_{i,t}=FT\}}) \\ & + \mathbb{1}_{\{p_{i,t}\}} earn^p_{i,t} - TT(earn^w_{i,t}, earn^p_{i,t}, cd_{i,t}, ac_{i,t}) \end{aligned} \tag{3.5}$$

where $earn^w_{i,t}$ and $earn^p_{i,t}$ stand for the gross labor earnings of the woman and her partner, respectively. The function $cc(\cdot)$ stands for the childcare costs, which are taken from the data and depend on the age of the youngest child.[57] For each hour that the woman works, she needs childcare for children under the age of 6. $TT(\cdot)$ represents the German tax and transfer system. I model all key features of the German tax and transfer system. In particular, joint taxation, unemployment benefits, social assistance, and childcare benefits are modeled carefully, since they might strongly affect the financial incentives to work.

Job offers and expectations. Women face labor market frictions when seeking employment. To switch from non-employment to part-time or full-time work, women must receive a job offer for the respective number of hours. An offer is also necessary to switch from part-time to full-time employment or vice versa.[58] The arrival probability of these offers depend on the lifetime under- or non-employment periods $(\rho_{i,t})$, which

[56] At the start of the working life, every individual is assumed to have zero on-the-job human capital.

[57] I follow the approach of Wrohlich (2011) by including individuals without positive childcare costs when computing the average expected childcare costs. One hour of care costs €1.82 for children under the age of 3 and €1.15 for children between the age of 3 and 6.

[58] No offer is needed to keep working part-time or full-time.

are given by the age minus the labor force entry age and the current human capital stock:

$$\rho_{i,t} = t - t_{\text{labor force entry}} - e_t. \tag{3.6}$$

Given $\rho_{i,t}$, the job offer probabilities are given by

$$\pi^{PT}(l_{t-1}, \rho_t) = \begin{cases} \dfrac{\exp\left(\gamma_{PT} + \gamma_{PT,\rho,1}\rho_{i,t} + \gamma_{PT,\rho,2}\rho_{i,t}^2\right)}{1 + \exp\left(\gamma_{PT} + \gamma_{PT,\rho,1}\rho_{i,t} + \gamma_{PT,\rho,2}\rho_{i,t}^2\right)} & \text{if } l_{t-1} \neq PT \text{ and } jp_{i,t} = 0 \\ 1 & \text{if } l_{t-1} = PT \text{ or } jp_{i,t} = 1 \end{cases} \tag{3.7}$$

$$\pi^{FT}(l_{t-1}, \rho_t) = \begin{cases} \dfrac{\exp\left(\gamma_{FT} + \gamma_{FT,\rho,1}\rho_{i,t} + \gamma_{FT,\rho,2}\rho_{i,t}^2\right)}{1 + \exp\left(\gamma_{FT} + \gamma_{FT,\rho,1}\rho_{i,t} + \gamma_{FT,\rho,2}\rho_{i,t}^2\right)} & \text{if } l_{t-1} \neq FT \text{ and } jp_{i,t} = 0 \\ 1 & \text{if } l_{t-1} = FT \text{ or } jp_{i,t} = 1 \end{cases} \tag{3.8}$$

Equations (3.7) and (3.8) highlight that no job offer is required to continue working part-time or full-time. In addition, maternity leave entitles women with the right to return to employment, independent of the hours. It is possible that women systematically over- or underestimate these offer probabilities and it is assumed that individuals do not update their expected job offer rates over time.[59] With $\widetilde{\pi}^{\cdot}(l_{t-1,\rho_t})$ standing for the expected job offer rate, the following relation between the expected and the true job offer rate is given:

$$\begin{aligned} \widetilde{\pi}^{PT}(l_{t-1}, \rho_t) &= \begin{cases} \alpha\, \pi^{PT}(l_{t-1}, \rho_t) & \text{if } \pi^{PT}(l_{t-1}, \rho_t) < 1 \\ 1 & \text{if } \pi^{PT} = 1 \end{cases} \\ \widetilde{\pi}^{FT}(l_{t-1}, \rho_t) &= \begin{cases} \alpha\, \pi^{FT}(l_{t-1}, \rho_t) & \text{if } \pi^{FT}(l_{t-1}, \rho_t) < 1 \\ 1 & \text{if } \pi^{FT} = 1 \end{cases} \end{aligned} \tag{3.9}$$

$$\text{where} \qquad \alpha \in \left[0, \frac{1}{\max\{\pi^{PT}(l_{t-1}, \rho_t), \pi^{FT}(l_{t-1}, \rho_t)\}}\right]$$

The parameter α determines the degree of deviation from the true job offer rate. It can never fall below zero since this would result in a negative expected job offer rate. Similarly, α must not exceed the inverse of the true job offer arrival rate since individuals are restricted to expect the job offer arrival rate not to be greater than one. Individuals do understand the concept of maternity leave and know that they have the right to

[59] Due to the rare event of being non-employed and then re-entering employment, it seems plausible that individuals do not have many opportunities to learn about the real job offer rate over the life-cycle.

return to their previous position. Depending on the size of α, the individuals might have rational expectations, underestimate, or overestimate the true job offer rate:

$$\begin{aligned} \text{rational expectations if } \alpha &= 1 \\ \text{underestimation if } \alpha &< 1 \\ \text{overestimation if } \alpha &> 1 \end{aligned} \tag{3.10}$$

The nesting of rational expectations in the model allows for straightforward testing of the hypothesis of non-biased expectations, by testing the hypothesis of $\alpha = 1$.

Job loss. When employed in the previous period, a woman can also involuntary lose her employment. Provided the woman worked in the previous period, there is an exogenous probability that the plant closes, denoted in the model as $\pi^L(l_{t-1})$. In this case, she is not able to choose employment in the current period and must await a job offer in the next period if she wants to re-enter employment. This probability is estimated outside the model using information provided in the SOEP sample. The questionnaire asks participants for the particular reason when a transition from employment to non-employment occurs. Among the answer options, only involuntary reasons like layoffs and plant closures are used to estimate the job loss probability.

Family dynamics. The birth of children, along with the formation and termination of partnerships, are modeled as exogenous stochastic processes depending on the woman's age and current family demographics. The probability of having a first child differs from the probability of having additional children.[60] In the model, and in line with Blundell et al. (2016), only the age of the youngest child is important, thus whenever a new child is born, the age of the youngest child is reset to zero. Children live in the household until they turn 18. Beginning a new partnership depends only on age, while separations also depend on the presence of a child and their age. Partners contribute to the household consumption and affect the women's leisure preferences. To keep the computational burden manageable, the partners' earnings are modeled to depend on the characteristics of the woman, including her age and family characteristics.[61] Agents in the model know and accommodate for these probabilities when forming expectations of future periods.

3.5.3 Maximizing expected lifetime utility

Given the preferences, the labor market frictions, and the external processes, women maximize their expected lifetime utilities each period. In a given period t, the opti-

[60]Since the model's decision period is a half-year, women are not able to have an additional child if the youngest child has not reached the age of one.

[61]This approach is similar to Van der Klaauw (1996), Sheran (2007) and Adda et al. (2017).

mization problem is formally given by

$$\max_{\{l_t,l_{t+1},\ldots,l_T\}} V_t(l_t,l_{t+1},\ldots,l_T,\omega_t) = u(l_t,\omega_t) + \mathbb{E}\left[\sum_{\tau=t+1}^{T}\beta^{\tau-t}u(l_\tau,\omega_\tau)\middle|\,\omega_t\right] \tag{3.11}$$

where the index of i is dropped for the ease of notation. The parameter β represents the discount factor, $\mathbb{E}[\cdot]$ the expectation operator, and ω_t a realization of the state space Ω_t in period t. The state space is defined as

$$\Omega_t = \{e_t, l_{t-1}, cd_t, ac_t, p_t, jp_{t-1}, r_t\}$$

Having specified the lifetime utility, and assuming the separability between the choice-specific error term and the rest of the utility function, the model can be represented in a two period decision process characterized by the Bellman (1957) equations (3.12) and (3.13).

$$\begin{aligned}
&v_t(l_t,\omega_t)\\
&\quad=u^\star(l_t,\omega_t)+\ \varepsilon_{l_t,t}\\
&\quad\quad+\beta\sum_{\substack{\omega_{t+1}\\ \in\Omega_{t+1}}}\Bigg\{\pi^L(l_t)\mathbb{E}\left[v^\star_{t+1}(NE,\omega_{t+1})+\varepsilon_{NE,t+1}\right]+(1-\pi^L(l_t))\Bigg(\\
&\qquad\qquad+\widetilde{\pi}^{PT}(l_t,\rho_t)(1-\widetilde{\pi}^{FT}(l_t,\rho_t))\mathbb{E}\left[\max_{\substack{j\in\\ \{NE,PT\}}}\{v^\star_{t+1}(j,\omega_{t+1})+\varepsilon_{j,t+1}\}\right]\\
&\qquad\qquad+\widetilde{\pi}^{FT}(l_t,\rho_t)(1-\widetilde{\pi}^{PT}(l_t,\rho_t))\mathbb{E}\left[\max_{\substack{j\in\\ \{NE,FT\}}}\{v^\star_{t+1}(j,\omega_{t+1})+\varepsilon_{j,t+1}\}\right]\\
&\qquad\qquad+\widetilde{\pi}^{PT}(l_t,\rho_t)\widetilde{\pi}^{FT}(l_{t-1},\rho_t)\mathbb{E}\left[\max_{\substack{j\in\{NE,\\ PT,FT\}}}\{v^\star_{t+1}(j,\omega_{t+1})+\varepsilon_{j,t+1}\}\right]\\
&\qquad\qquad+(1-\widetilde{\pi}^{PT}(l_t,\rho_t))(1-\widetilde{\pi}^{FT}(l_t,\rho_t))\mathbb{E}\left[v^\star_{t+1}(NE,\omega_{t+1})+\varepsilon_{NE,t+1}\right]\Bigg)\\
&\qquad\qquad\Bigg\}q(\omega_{t+1}|l_t,\omega_t)
\end{aligned} \tag{3.12}$$

$$\begin{aligned}
&\text{for } t=T:\\
&\quad v_t(l_T,\omega_T)=\ u^\star(l_T,\omega_T)+\varepsilon_{l_T,T}
\end{aligned} \tag{3.13}$$

where $q(\omega_{t+1}|l_t,\omega_t)$ denotes the probability of arriving at state space ω_{t+1} given choice l_t and state space ω_t, and $u^\star_{i,t}$ the utility function without the choice specific error term, i.e. $u^\star_{i,t} \equiv u_{i,t} - \varepsilon_{i,t}$. Similarly, $v(\cdot)^\star \equiv v(l_t,\omega_t) - \varepsilon_{l_t,t}$ denotes the value function without the choice specific error term. Furthermore, if the current choice is non-employment ($l_t = NE$), the job loss probability is zero ($\pi^L(l_t) = 0$), since women can only lose their job if they are employed. The biased expectations of the future job arrival rates enter only the value function, since they represent what the individual beliefs about future possibilities.

Two assumptions help with the formulation of the stated value functions. First, I assume that individuals do not know that their expected job offer probability might differ from the actual offer rate and, second, they do not update their expected job offer probability over the life-cycle. This causes individuals to treat the expected employment probability as given when maximizing their expected lifetime utility. As a result, there is no correlation between the expected job offer rate and the expected choice specific error component. Additionally, I assume that mothers fully understand the institutional settings and know that they receive employment protection when they qualify for it. Thus, mothers in employment protection do not have a bias about their possible choice restriction in the next period. They correctly assume they can return to employment, less the probability of a plant closure, during that time. Since the model has a finite horizon, it can be solved by backwards induction using equations (3.12) and (3.13) for a given set of parameters.

3.6 Identification and Structural Estimation

Although section 3.4 provides some suggestive evidence that individuals might overestimate their probability of finding employment, the evidence is based on a stated preferences approach. In contrast, the revealed preferences approach presented in this section relies only on actual choices. Both approaches have their advantages and disadvantages,[62] but employing both creates a stronger case when collecting evidence for overconfidence and its resulting costs.

For the ease of discussion of the identification, some simplifications are useful. First, since the focus lies on career interruptions and subsequent returns to employment, the following discussion summarizes full-time and part-time employment into a single state

[62] In this chapter, the major difference between both methods is that the stated preference approach relies on a single question for which individuals do not face any consequences when answering it carelessly or even non truthfully, while the revealed preference approach is based on actual choices in the real world. There are several reasons why the suggestive evidence should be interpreted with care. First, individuals might answer strategically, thinking their answers might influence policy choices. Second, the framing and the precise wording of the question might influence subjects differently, resulting in a wide variety of possible interpretations of their answers. Third, the concept of probabilities might be challenging for a significant number of subjects, thus preventing some individuals from answering the question correctly (see for example Garfield & Ahlgren, 1988; Tversky & Kahneman, 1973). Although the labor market state is recorded via the same survey and, thus, also potentially prone to some of the described problems, the question refers to already made choices and is much easier to interpret.

E.[63] Consequently, the job offer rate is for general employment and is denoted by π^E. In addition, assuming that $\varepsilon_{j,t}$ is type-I extreme value distributed with zero mean, equation (3.12) can be rewritten as

$$\begin{aligned} v_t(l_t = NE, \omega_t) &= u(l_t, \omega_t) \\ &+ \beta \sum_{\substack{\omega_{t+1} \\ \in \Omega_{t+1}}} \Bigg\{ \widetilde{\pi}^E(\rho_t) \ln \left(\sum_{j \in \{NE,E\}} \exp(v^\star_{t+1}(j, \omega_{t+1})) \right) \\ &\qquad + \Big(1 - \widetilde{\pi}^E(\rho_t)\Big)\, v^\star_{t+1}(NE, \omega_{t+1}) \Bigg\} q(\omega_{t+1}|l_t, \omega_t) \\ v_t(l_t = E, \omega_t) &= u(l_t, \omega_t) \\ &+ \beta \sum_{\substack{\omega_{t+1} \\ \in \Omega_{t+1}}} \Bigg\{ \Big(1 - \pi^L\Big) \ln \left(\sum_{j \in \{NE,E\}} \exp(v^\star_{t+1}(j, \omega_{t+1})) \right) \\ &\qquad + \pi^L v^\star_{t+1}(NE, \omega_{t+1}) \Bigg\} q(\omega_{t+1}|l_t, \omega_t). \end{aligned} \tag{3.14}$$

To reduce the space of some equations, I additionally apply the following notation:

$$\begin{aligned} LS(E, \omega_{t+1}) &= \ln \left(\sum_{j \in \{NE,E\}} \exp(v^\star_{t+1}(j, \omega_{t+1})) \right) \\ LS(NE, \omega_{t+1}) &= v^\star_{t+1}(NE, \omega_{t+1})) \\ \Delta LS(E - NE, \omega_{t+1}) &= LS(E, \omega_{t+1}) - LS(NE, \omega_{t+1})) \end{aligned} \tag{3.15}$$

3.6.1 Identification

The identification of the model combines a bunching related approach, exploiting a discontinuity in the future expected value of non-employment, with exogenous variation from three major maternity leave reforms. The bunching caused by the discontinuity primarily identifies the expectations of the future job offer rate. The reforms create counterfactual-like situations, such that it is possible to evaluate behavior in the absence of the discontinuity. This approach allows for separate identification of the real job offer arrival rates, the expectation of these rates, and individual preferences. The section starts with an overview of the identification strategy to provide intuition on how observed choices identify the model's key parameters. Afterwards, a more formal discussion is presented.

[63] The distinction between both states is mostly important for estimating possible career costs, since income from part-time and full-time employment differ.

3.6.1.1 Overview of Identification Strategy

Identifying Expectations. The end of employment protection introduces a discontinuity in the probability to be able to choose employment in subsequent periods. During the protection, mothers can freely decide if they want to return to employment or if they want to remain non-employed, since they are guaranteed their previous position by law. The only risk they face is that their plant might close, resulting in a loss of their employment guarantee. If their career break lasts beyond the end of the protection period, the likelihood to be restricted in future choices changes, since they need to rely on a job offer to choose employment in this situation. The change from an employment guarantee to an uncertain situation at the end of maternity leave causes mothers who prefer to stay longer at home with their child to consider returning at the end of their employment protection. Stated differently, some mothers return within the employment protection period because they expect it to be hard to find employment once the protection ends.

Figures 3.2 and 3.3 describe this situation graphically. Figures 3.2 visualizes the differences in the underlying process while in employment, in non-employment or in maternity leave. Only when an individual is currently non-employed but wants to work, the job offer rate directly affects the outcome. The switch from the underlying process at the end of the employment protection is visualized in Figure 3.3. At this point, if a mother does not return to employment, she has to rely on a job offer in the future.

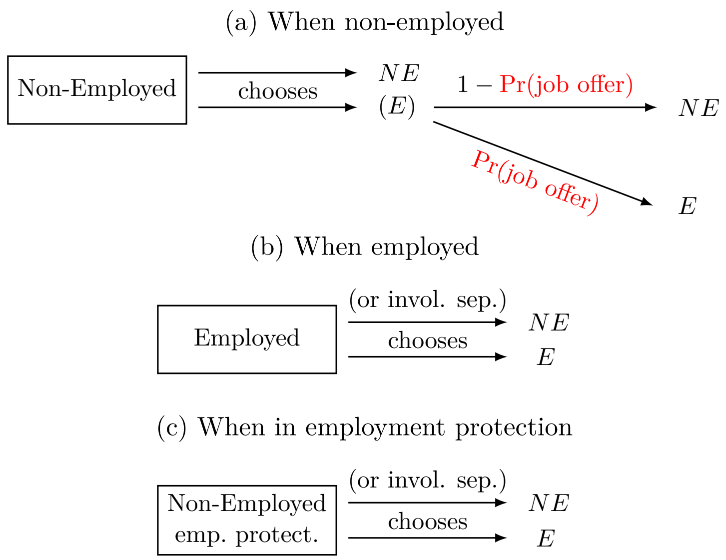

Figure 3.2: Underlying processes

The higher a mother expects her job arrival rate to be, the lower she expects to be restricted in her future choices and the less likely she is to return at the end of her

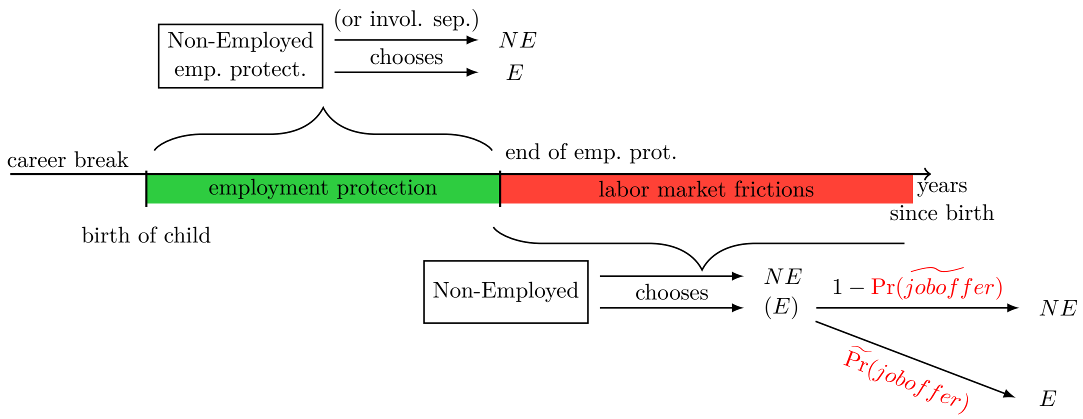

Figure 3.3: Change in underlying processes at the end of maternity leave

maternity leave. In the extreme, if she expects to always find employment, the end of employment protection does not present a discontinuity for her, since a return to employment is always possible. In contrast, if she believes that there are no future employment opportunities, then returning within the maternity leave period provides the only option for her to return to employment. Thus, the number of mothers returning right before the end of their employment protection informs about the average expectations of the future job offer probability. An important detail of this process is that mothers must decide if they want to return within the employment protection, before they experience the real labor market conditions, thus they have to rely on their expectations regarding future employment possibilities.

Equations (3.16) and (3.17) formalize this change in the future expected value function in the period before the employment protection ends. Both equations show the trade-off between the future value of employment and non-employment conditioned on the mother not having re-entered employment since giving birth. Equation (3.16) depicts the situation of an individual enjoying employment protection for at least the next period, while equation (3.17) depicts the situation in which there is no future employment protection. In equation (3.16), the correct $\pi^E(\rho)$ is known to the individual and equals one minus the probability of a job loss, since the mother enjoys employment protection and is fully aware of it. Note that the probability to be able to choose from the entire choice set is the same, independent of choosing re-entering employment or staying non-employed. In equation (3.17) and, thus, in the last period of employment protection, this changes. When choosing to be non-employed in the current period, but desiring to work in the next period, an individual has to rely on a job offer. In contrast, choosing to be employed in the current period, only restricts the choice set with the probability of a plant closure.

Employment protection in the next period:

$$
\begin{aligned}
&[v_t(NE_t,\omega_t)-u(NE_t,\omega_t)] - [v_t(E_t,\omega_t) - u(E_t,\omega_t)] = \\
&\qquad \beta \sum_{\substack{\omega_{t+1} \\ \in\Omega_{t+1}}} \left\{\left(1-\pi^L\right)\Delta LS(E-NE,\omega_{t+1}) + LS(NE,\omega_{t+1})\right\} q(\omega_{t+1}|NE_t,\omega_t) \\
&\qquad - \beta \sum_{\substack{\omega_{t+1} \\ \in\Omega_{t+1}}} \left\{\left(1-\pi^L\right)\Delta LS(E-NE,\omega_{t+1}) + LS(NE,\omega_{t+1})\right\} q(\omega_{t+1}|E_t,\omega_t)
\end{aligned}
\tag{3.16}
$$

No employment protection in the next period:

$$
\begin{aligned}
&[v_t(NE_t,\omega_t)-u(NE_t,\omega_t)] - [v_t(E_t,\omega_t) - u(E_t,\omega_t)] = \\
&\qquad \beta \sum_{\substack{\omega_{t+1} \\ \in\Omega_{t+1}}} \left\{\widetilde{\pi}^E(\rho_t)\Delta LS(E-NE,\omega_{t+1}) + LS(NE,\omega_{t+1})\right\} q(\omega_{t+1}|NE_t,\omega_t) \\
&\qquad - \beta \sum_{\substack{\omega_{t+1} \\ \in\Omega_{t+1}}} \left\{\left(1-\pi^L\right)\Delta LS(E-NE,\omega_{t+1}) + LS(NE,\omega_{t+1})\right\} q(\omega_{t+1}|E_t,\omega_t)
\end{aligned}
\tag{3.17}
$$

Therefore, even if the current utility of being non-employed exceeds the current utility of being employed by far, the woman might still choose employment due to its higher future value.[64] The higher the difference in future values are, the more likely the choice of re-entering employment in the last period of the employment protection and, thus, the more individuals return to employment in this last period. The difference in the future values depends strongly on the expected job offer rate. The higher this expected job offer rate is, the less severe is the difference in the future values, and the less bunching should be observed. In fact, if individuals expect the job offer rate as high as $(1-\pi^L)$, there is no discontinuity at all and, thus, there should be no observable difference between the returning rate shortly before the end of employment protection and shortly after the end of employment protection.
Figure 3.4 visualizes the identification of the job offer expectations. For ease of the discussion, the left panel of the figure illustrates the situation when time is assumed to be continuous. The right panel depicts the difference in the value functions and the observed outcomes when time is assumed to be discrete. The side-by-side placement provides a better understanding of the translation from the underlying preferences (top left panel) to the observed outcomes in the data (bottom right panel). All x-axes denote the time since the birth of the youngest child. The employment protection is assumed to end with period three. The top graphs plot the difference in the value functions between non-employment and employment, while the bottom graphs show the density

[64] Note that as long as there is some utility in choosing employment, the expected maximum of being able to choose between employment and non-employment is always greater than the expected maximum of only being able to choose non-employment.

of mothers returning to employment depending on the age of the youngest child.[65]
Three scenarios are illustrated. The first scenario, indicated by the dotted black line, represents the counterfactual situation when the employment protection does not end with period 3. It equals the scenario in which individuals expect the job offer rate to be exactly as high as the probability of not losing employment due to a plant closure. The second scenario, indicated by the solid blue line, graphs a situation in which individuals have rational expectations of their future employment possibilities. The third scenario depicted by the dashed gray line represents individuals with biased expectations. These are overestimating the probability of finding employment after the end of the employment protection. However, they anticipate that the probability is lower than in the protection period and therefore differ from the individuals in scenario I.
To identify the expected job offer rate, the discontinuity in the value function is critical. As depicted by the three scenarios, the more optimistic individuals are about their future employment prospects, the lower is the bunching of returnees at the end of the employment protection. This illustrates how the mass of returning mothers is linked to their future job offer expectations. In the underlying process, the majority of women would return to employment just before the end of the employment protection, while in the discrete data, the majority returns in the last period of the employment protection, spanning over more time. It is important that the decision to return within the employment protection is made before actually facing the labor market frictions. Therefore, the individual can only have expectations of the job offer probability. This reinforces that the excess mass at the end of the employment protection is strongly influenced by future expectations of these frictions.
Separating Expectations from Preferences and Real Job Offer Rates. In principle, the bunching of returnees at the end of employment protection can be caused by other discontinuities in the value function. For example, the bottom graphs of figure 3.4 would look similar if leisure preferences depending on the age of the youngest child discontinuously change at the end of period three. To separate behavior due to discontinuities in preferences from behavior due to job offer expectations, it is beneficial to know the counterfactual scenario in which the employment protection goes beyond the third period. Furthermore, the identification of the real job offer rate can be determined when behavior can be compared to a counterfactual situation, for which

[65] All graphs assume that if employment protection never ends, the difference between the value of non-employment and employment shrinks over time. This is mainly due to the decreasing utility of leisure as the child gets older. Additionally, the more the human capital depreciates while not working, the smaller the difference between the future values of non-employment and employment gets. In the graph, it is assumed that the first effect dominates the second. That the density is not monotonically increasing in the bottom figures is due to the assumption that the majority of individuals are assumed to have returned before the end of period 3.

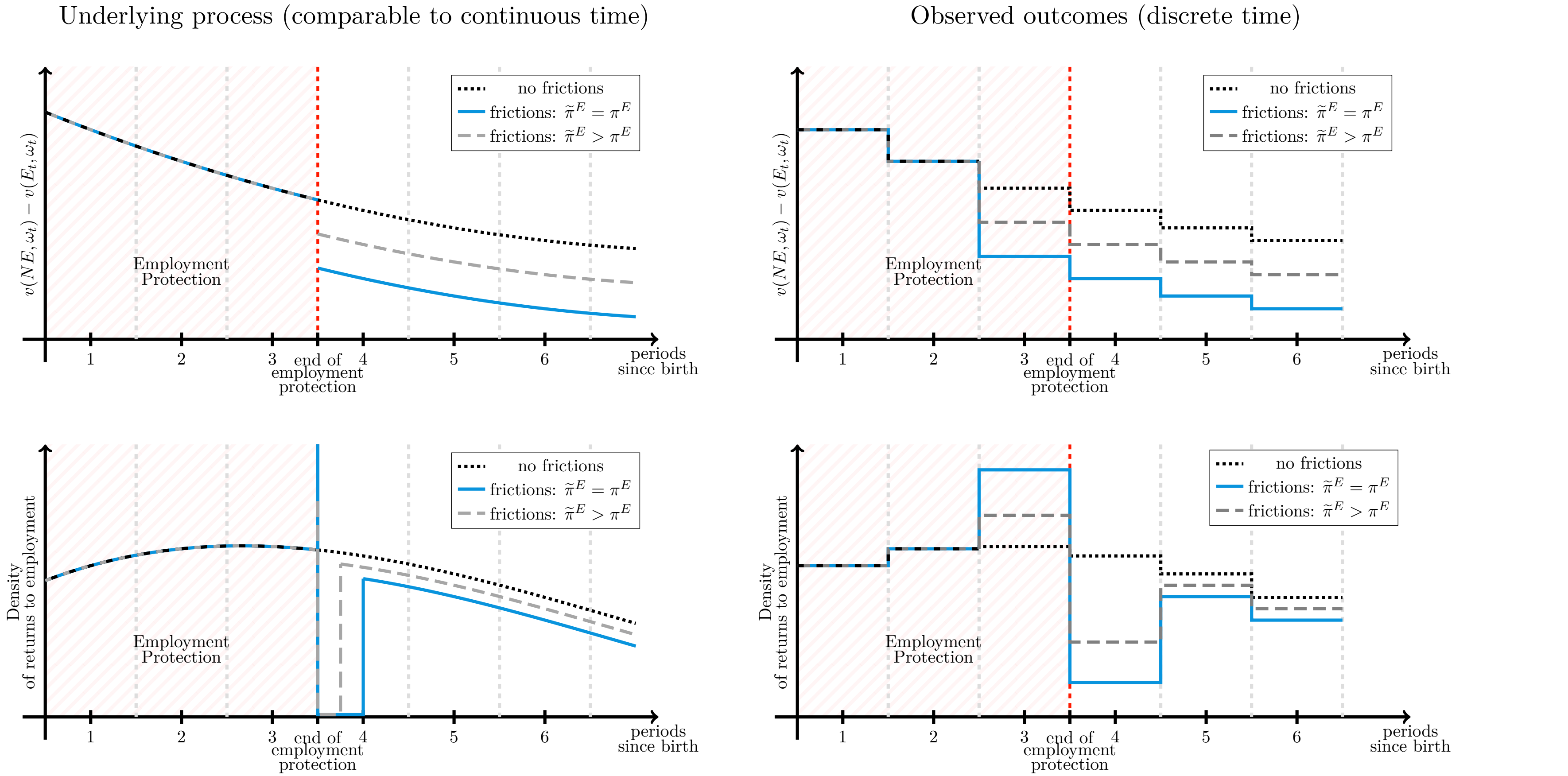

Figure 3.4: Identification of expectations

Notes: The figure visualizes the identification of the job offer expectations. The left panel of the figure illustrates the situation when time is assumed to be continuous, the right panel depicts the difference in the value functions and the observed outcomes when time is assumed to be discrete. All x-axes denote the time since the birth of the youngest child. The employment protection is assumed to end with period three. The top graphs plot the difference in the value functions between non-employment and employment, while the bottom graphs show the density of mothers returning to employment depending on the age of the youngest child.

the employment protection ends before the third period. The maternity leave regimes presented in section 3.2 provide such counterfactual-like scenarios. Figure 3.5 visualizes the time lines of these regimes. It shows all three regimes and their respective lengths of employment protection depending on the years since childbirth. The green bars on top of a regime's timeline indicate that the individual's jobs are still protected, while in contrast the red bars below indicate that the individual has to receive a job offer if she desires to re-enter employment.

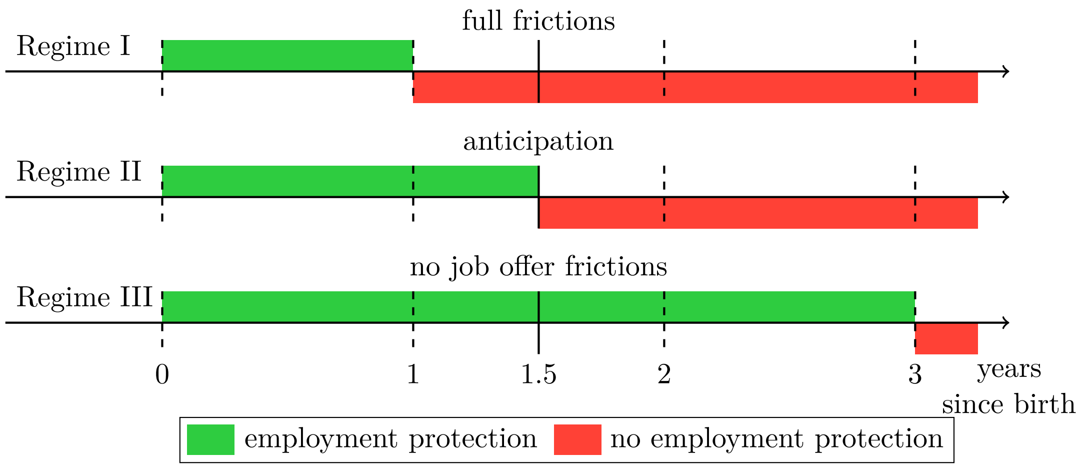

Figure 3.5: Identification: policy reforms

The different durations of employment protection by the three regimes exogenously creates three groups, each facing different environments when the youngest child reaches the age of 1.5 years. In regime II, mothers face the end of the employment protection. Thus, the number of people returning to employment in the period before their youngest child turns 1.5 in regime II contributes to the identification of job offer expectations. Returning mothers in regime III, whose youngest child reaches the age of 1.5 are not influenced by their expectations of future job offer rates, since their employment protection continues until the child reaches the age of 3. Therefore, the number of returning mothers shortly before their child exceeds the age of 1.5, can be used to control for discontinuities in preferences at this age. Finally, the transition rates of mothers in regime I, when their youngest child reaches the age of 1.5, can be used for identification of the real job offer rate.[66]

More generally, transition probabilities from non-employment to employment of women in regime III can be used to identify leisure preferences depending on the age of the youngest child. As equation (3.16) shows, the difference in future expected value

[66]Note that regime III and regime II also function as counterfactual scenarios for regime I, and regime I functions to regime II as a counterfactual scenario for regime III.

functions is only generated by the different transition probabilities of the state space ($q(\omega_{t+1}|NE_t, \omega_t) \neq q(\omega_{t+1}|E_t, \omega_t)$) when in employment protection. The only systematic difference for these transition probabilities is the human capital stock growth, which results in higher future wages when in employment. Once consumption preferences are identified, it is possible to quantify the differences in the future expected values of non-employment and employment, since these differences are only related to future consumption. This allows for controlling for future expected values, when comparing observed choice probabilities of women in employment protection. The differences in choice probabilities are, thus, only driven by preferences of the instantaneous utility of working or non-working, ultimately identifying leisure preferences.

Having identified job offer expectations and leisure preferences, the true job offer rate can be recovered by comparing the transition rates of non-employed individuals in regime 1 with the ones from regime 3 for the periods in which individuals face different employment protection states in the regimes. The transitions of the individuals in regime 3 for this time period are not affected by the labor market frictions, while, in contrast, individuals from regime 1 are. The difference in the transition rates informs about how restricted individuals are in their choice set when they are not in employment protection.

3.6.1.2 Formal Identification

After having provided intuition for the identification approach, this subsection shows how the key parameters are formally identified, while for the ease of discussion only a choice between non-employment and employment is considered.[67] Occasionally, the state space is split into $\Omega_t^- = \Omega_t \setminus x_{t-1}$ and x_{t-1}, denoting that x_{t-1} is excluded from the state space and listed separately. For example, I denote observed choice probabilities as $\Pr(l_t|l_{t-1}, \omega_t^-, jp_t)$, describing the probability of choice l_t, conditioned on last period's choice l_{t-1}, this period's state space ω_t^- and employment protection state jp_t.[68] Similarly, $v(l_t, \omega_t^-, jp_t)$ denotes the value function of choice l_t, given the state space ω_t^- and the employment protection state jp_t.

Since the conditional observed choice probabilities are crucial for the identification of the model, the discussion starts with the formulas of the general observed choice probabilities, followed by the choice probabilities when individuals are in employment

[67]Note that all parameters are estimated together using the method of simulated moments. For this method, all moments can influence all parameter estimates, such that there is no *exclusive* matching of individual moments to respective parameters. It is, however, necessary to discuss the formal identification of key parameters to better understand which moments predominantly drive results and what assumptions are necessary for identification.

[68]Being in employment protection is then indicated by 1, while not being in employment protection by 0.

protection. Given the model, the probability to be employed, conditioned on also being employed in the previous period, is the product of the unconditional choice probability and the probability of not being laid off:

$$\begin{aligned}\Pr\left(E|E,\omega_t^-, jp_t = 0\right) = \\ &\left(1-\pi^L\right)\Pr\left(\varepsilon_{NE,t+1} < v^\star(E,\omega_t) + \varepsilon_{E,t+1} - v^\star(NE,\omega_t)\right) \\ =&\left(1-\pi^L\right)\frac{\exp\left(v^\star(E,\omega_t)\right)}{\sum\limits_{j\in\{NE,E\}}\exp\left(v^\star(j,\omega_t)\right)}\end{aligned} \tag{3.18}$$

The probability to be employed conditioned on being non-employed in the previous period is similar, but depends on the job offer instead of the job loss probability:

$$\begin{aligned}\Pr\left(E|NE,\omega_t^-, jp_t = 0\right) = \\ &\pi^E(\omega_t)\Pr\left(\varepsilon_{NE,t+1} < v^\star(E,\omega_t) + \varepsilon_{E,t+1} - v^\star(NE,\omega_t)\right) \\ =&\pi^E(\omega_t)\frac{\exp\left(v^\star(E,\omega_t)\right)}{\sum\limits_{j\in\{NE,E\}}\exp\left(v^\star(j,\omega_t)\right)}\end{aligned} \tag{3.19}$$

The observed choice probabilities are based on the true job offer probability, since these ultimately determine if a woman has received a real job offer and, thus, can be observed in employment. The biased job offer expectations only occur indirectly in the value functions of equations (3.18) and (3.19).

The main difference between a mother in employment protection and a mother who is not, is the need to receive a job offer when transitioning from non-employment to employment. Given the likelihood of a plant closure, the observed choice probability of employment of a mother in maternity leave is

$$\begin{aligned}\Pr\left(E|NE,\omega_t^-, jp_t = 1\right) = \\ &(1-\pi^L)\Pr\left(\varepsilon_{E,t+1} < v^\star(E,\omega_t) + \varepsilon_{E,t+1} - v^\star(NE,\omega_t)\right) \\ =&(1-\pi^L)\frac{\exp\left(v^\star(E,\omega_t)\right)}{\sum\limits_{j\in\{NE,E\}}\exp\left(v^\star(j,\omega_t)\right)}.\end{aligned} \tag{3.20}$$

The choice probabilities for non-employment are not being discussed, since these are given by the complementary probabilities and do not add to the identification.

Identification of the True Job Offer Probability. Having established the observed choice probabilities, real job offer rates can be identified by dividing the ones of women who only differ in last period's choice:

$$\frac{\Pr\left(E|NE,\omega_t^-,jp_t=0\right)}{\Pr(E|E,\omega_t^-,jp_t=0)}=\frac{\exp\left(v^\star(E,\omega_t)\right)}{\sum\limits_{j\in\{NE,E\}}\exp\left(v^\star(j,\omega_t)\right)}\pi^E(\omega_t)\frac{\sum\limits_{j\in\{NE,E\}}\exp\left(v^\star(j,\omega_t)\right)}{(1-\pi^L)\exp\left(v^\star(E,\omega_t)\right)} \tag{3.21}$$

As discussed, the involuntary separation rate is directly estimated from the data and can be treated as known. Rearranging terms leads to

$$\pi^E(\omega_t)=\frac{\Pr\left(E|NE,\omega_t\right)\left(1-\pi^L\right)}{\Pr\left(E|E,\omega_t\right)}. \tag{3.22}$$

Equation (3.22) identifies the job offer probability, which is assumed to be known for the rest of the identification discussion.[69]

Identification of Expected Job Offer Probability. The formal identification of expectations of the job arrival rate uses two employment protection regimes and does not rely on the functional forms of the utility function and wage process. The crucial element for the identification of expectations is the end of the employment protection and its resulting discontinuity in future expected values. A starting point is to derive the logarithm of observed choice probabilities for a group of individuals from regime I,

[69]Note that, in a similar manner, it is possible to identify the part-time and full-time job offer probability for a given state space. These can be recovered, since there are six independent conditional choice probability equations, including six unknowns: the two job offer probabilities and a maximum of four "pure" choice probabilities from various choice sets, where choice probabilities in the absence of labor market frictions are referred to as pure. The remaining three pure choice probabilities can then be constructed from the complementary events.

who have a child of age one, but are no longer fertile:[70]

$$\begin{aligned}
&\mathbb{E}\left[\ln\left(\frac{\Pr(E|NE,\omega_t^-,ac=1,jp_t=1,r=I)}{\Pr(NE|NE,\omega_t^-,ac=1,jp_t=1,r=I)-\pi^L}\right)\right]\\
&\quad=\mathbb{E}\Bigg[u(E,\omega_t^-,ac=1)-u(NE,\omega_t^-,ac=1)\\
&\qquad\quad+\beta\sum_{\omega_{t+1}}\left\{(1-\pi^L)LS(E,\omega_{t+1}^-,ac=1.5)+\pi^L LS(NE,\omega_{t+1}^-,ac=1.5)\right\}\\
&\qquad\qquad\quad q(\omega_{t+1}|E,\omega_t^-,ac=1,jp_t=1,r=I)\\
&\qquad\quad-\beta\sum_{\omega_{t+1}}\left\{\widetilde{\pi}^E(\omega_t)LS(E,\omega_{t+1}^-,ac=1.5)+(1-\widetilde{\pi}^E(\omega_t))LS(NE,\omega_{t+1}^-,ac=1.5)\right\}\\
&\qquad\qquad\quad q(\omega_{t+1}|NE,\omega_t^-,ac=1,jp_t=1,r=I)\Bigg]\\
&\quad=\mathbb{E}LR_{RI}(\omega_t^-,ac_t=1,jp_t=1,r=I)
\end{aligned}\tag{3.23}$$

where $\mathbb{E}LR_{RI}(\omega_t^-,ac=1)$ is introduced in order to simplify notation by denoting the expected logarithm of the choice probabilities. Since the women are in their last period of employment protection, staying non-employed in the current period causes them to rely on a future job offer if they desire to re-enter employment later. In contrast, women from regime II can remain non-employed for another period, before they have to rely on job arrivals. Their difference in the expected future value functions is:

$$\begin{aligned}
&\mathbb{E}LR_{RII}(\omega_t^-,ac_t=1,jp_t=1,r_t=II)\\
&\quad=\mathbb{E}\Bigg[u(E,\omega_t^-,ac=1)-u(NE,\omega_t^-,ac=1)\\
&\qquad\quad+\beta\sum_{\omega_{t+1}}\left\{(1-\pi^L)LS(E,\omega_{t+1}^-,ac=1.5)+\pi^L LS(NE,\omega_{t+1}^-,ac=1.5)\right\}\\
&\qquad\qquad\quad q(\omega_{t+1}^-|E,\omega_t^-,ac=1,jp_t=1,r=II)\\
&\qquad\quad-\beta\sum_{\omega_{t+1}}\left\{(1-\pi^L)LS(E,\omega_{t+1}^-,ac=1.5)+\pi^L LS(NE,\omega_{t+1}^-,ac=1.5)\right\}\\
&\qquad\qquad\quad q(\omega_{t+1}^-|NE,\omega_t^-,ac=1,jp_t=1,r=II)\Bigg]
\end{aligned}\tag{3.24}$$

By subtracting (3.23) from (3.24), it is possible to eliminate the instantaneous utilities, such that only the difference in the future expected values is left. Furthermore, expected value functions and transition probabilities are the same for both regimes, when

[70] Assuming that there are no future children simplifies the formal discussion, but is not necessary to identify expectations.

individuals choose employment in t and, thus, also cancel out.[71] It allows writing the difference as

$$
\begin{aligned}
&\mathbb{E}LR_{RI}(\omega_t^-, ac_t = 1, jp_t = 1, r = I) - \mathbb{E}LR_{RII}(\omega_t^-, ac_t = 1, jp_t = 1, r_t = II) \\
&\quad = -\beta \sum_{\omega_{t+1}} \left\{ \widetilde{\pi}^E(\omega_t) \Delta LS(E - NE, \omega_{t+1}^-, ac = 1.5) \right\} q(\omega_{t+1}^- | NE, \omega_t^-, ac = 1.0) \\
&\qquad + \beta \sum_{\omega_{t+1}} \left\{ (1 - \pi^L) \Delta LS(E - NE, \omega_{t+1}^-, ac = 1.5) \right\} q(\omega_{t+1}^- | NE, \omega_t^-, ac = 1.0) \\
&\quad = \beta \sum_{\omega_{t+1}} \left\{ \left[(1 - \pi^L) - \widetilde{\pi}^E(\omega_t) \right] \Delta LS(E - NE, \omega_{t+1}^-, ac = 1.5) \right\} \\
&\qquad q(\omega_{t+1}^- | NE, \omega_t^-, ac = 1.0)
\end{aligned} \tag{3.25}
$$

The left hand side of this equation and the transition probabilities can be directly computed from the data. As stated above, the rate of involuntary job separations π^L is also known. In addition, $\Delta LS(E - NE, \omega_{t+1})$ can be computed relying only on observed choice probabilities and the known job separation rate:

$$
\begin{aligned}
\Delta LS(E - NE, \omega_{t+1}) &= \ln \left(\sum_j \exp(v(j, \omega_{t+1})) \right) - \ln \left(\exp(v(NE, \omega_{t+1})) \right) \\
&= \ln \left(\frac{\sum_j \exp(v(j, \omega_{t+1}))}{\exp(v(NE, \omega_{t+1}))} \right) \\
&= \frac{(1 - \pi^L)}{\Pr(NE | E, \omega_t) - \pi^L}
\end{aligned} \tag{3.26}
$$

Plugin equation (3.26) into equation (3.25) results in

$$
\begin{aligned}
&\mathbb{E}LR_{RI}(\omega_t^-, ac = 1) - \mathbb{E}LR_{RII}(\omega_t^-, ac = 1) \\
&\quad = \left[(1 - \pi^L) - \widetilde{\pi}^E(\omega_t) \right] \beta \sum_{\omega_{t+1}} \left\{ \frac{(1 - \pi^L)}{\Pr(NE | E, \omega_t) - \pi^L} \right\} \\
&\qquad q(\omega_{t+1}^-, ac = 1.5 | E, \omega_t^-, ac = 1)
\end{aligned} \tag{3.27}
$$

Since all terms, beside $\widetilde{\pi}^E(\omega_t)$ are known in this equation, it identifies the job offer expectations. The identification only depends on observed choices and transition probabilities, the lay-off rate and the discount factor. It is not driven by the functional forms

[71] Additionally, it is assumed that there are no financial differences between being non-employed with employment protection and being non-employed without the protection. This simplifies the formal identification argumentation. Since maternity benefits are means tested in the regimes, this refers to groups that do not qualify for these benefits. Identification can also come from individuals receiving maternity benefits since the structural model and estimation can account for differences in incomes, once consumption preferences are identified.

of the utility function, of the wage process, or the real job offer probabilities. The equation also reflects the intuitive arguments made in section 3.6.1.1. The left-hand side consists of the differences in choice probabilities between regimes I and II, when the youngest child is one year old. The more that mothers return to employment in regime I compared to regime II, the greater the left-hand side. On the right-hand side, the discount factor and the transition probabilities are both positive. The same holds for the term from equation (3.25), since it is a logarithm of an inverse of a probability, leaving the term $(1 - \pi) - \widetilde{\pi}^E(\omega_t)$ to determine the sign of the right-hand side.

When the observed choice probabilities in both regimes are equal, the left-hand side is equal to zero. This means that there is not an excessive number of returning mothers at the end of the employment protection due to low job offer expectations. In this case, the job offer expectations have to match the probability of not being laid off, which aligns with the only case of the right-hand side of equation (3.26) becoming zero. Since, the probability of involuntary job separations is rather low, job offer expectations must be rather high in this case. When more mothers return in regime I than in regime II, the left-hand side is positive. It reflects a higher mass of returning mothers at the end of maternity leave than can be explained solely with instantaneous leisure preferences. In this case, the expected job offer rate has to be lower than the probability of not being laid off to cause the right-hand side to be positive. Furthermore, the higher the difference between the regimes, the lower the expectations of being able to return to employment in the future, exactly linking the excess mass of returning at the end of maternity leave to their job offer expectations. In the unlikely scenario that more mothers return to employment in regime II than in regime I, the left-hand side becomes negative. The job offer expectations must be higher than the probability of not being laid off, in this case, since it is only then that the right-hand side is negative, too.

Because the expected and the real job offer rate are connect, the parameter α is also identified:

$$\alpha = \frac{\pi^E(\omega_t)}{\widetilde{\pi}^E(\omega_t)} \tag{3.28}$$

In a model, in which job offers differ between part-time and full-time employment, the identification approach is similar. As noted before, the real job offer rates and choice probabilities in the absence of labor market frictions can be identified in such a model, independent of expectations. This allows for an analogous comparison of observed choice probabilities to recover differences in the future expected value functions between having an additional period of employment protection and having reached the end of protection. Comparing the different choices for part- and full-time in these two states similarly identifies job offer expectations.

3.6.2 Estimation Procedure

The estimation procedure is divided into two parts. In a first step, the discount factor is set, and the exogenous parameters and processes are estimated. In a second step, the parameters of the structural model are estimated. The semi-annual discount factor β is set to $\sqrt{0.98}$, the square root of the annual discount factor found, for example, in Attanasio et al. (2008) and Blundell et al. (2016). Both studies employ similar utility functions, which allow for non-separability of leisure and consumption in the framework of female labor supply over the life-cycle.[72] Job separations are estimated via a linear regression quadratic in age and occur with an average probability of 4.8%, decreasing over the life-cycle. Childcare costs are estimated as averages for children under three (€1.82 per hour) and for children between three and six (€1.15 per hour). The exogenous processes of marrying and divorcing partners, of the partner's income, and of the arrival of children are estimated with the method of simulated, relying on averages over the life-cycle. Appendix B.2 provides further details on these estimations. In the second step, relying on the parameters estimated in the first step, a method of simulated moments estimation is carried out to recover the structural parameters of the model.[73] To estimate the 42 parameters of the model, a total of 490 moments are used. These consist mostly of conditional choice probabilities and transition rates. To identify the job offer expectations, regime specific employment choices are targeted. An overview of the distribution of moments is provided in table 3.5.

[72]Haan & Prowse (2015) use the same discount factor in an estimation based on SOEP data.

[73]Due to computational limitations, the second step does not incorporate that parameters of the first step are estimated.

Table 3.5: Overview of moments

Moments	Number	Structural parameters primarily identified
(1)	(2)	(3)
Average full-time and part-time employment rates, unconditional, and conditional on partnerships, presence of a child, age of the youngest child	21	Utility function parameters
Transition rates from non-employment to employment, from employment to non-employment, and from employment to employment	30	Job offer probabilities
life-cycle employment rates, overall, full-time and part-time	192	Utility function parameters and job offer probabilities
Part-time and full-time employment rates depending on the age of the youngest child	111	Leisure preferences depending on the age of the youngest child
Log wages at beginning of the life-cycle, log wage distribution for full-time and part-time, change of wages after non-employment spell	19	Human capital and wage process parameters
Log wages of the life-cycle	64	Human capital and wage process parameters
Regime specific employment rates depending on the age of the youngest child	48	Job offer expectations

Given the set of moments, parameters are estimated by the method of simulated moments. This method tries to maximize the similarity between the simulated data and the observed data, where similarity refers to the chosen moments. The estimation procedure is as follows:

1. For a given set of parameters (Θ), the described model is solved via backwards induction.

2. Given the choice-specific value functions, all life-cycle decisions for all observed women are simulated. For each woman in the sample, ten life-cycles are simulated.

3. For a given woman, all periods that are not observed in the SOEP data are deleted in the simulated data.[74] Steps 1, 2 and 3 result in a simulated data set with ten times as many observations as in the observed SOEP data.

4. For the simulated and the observed sample all moments are computed and the value of the following objective function is computed:

$$f(\Theta) = \left\{ \sum_{k=1}^{K} \left[\left(M_k^d - \frac{1}{s} \sum_{s=1}^{10} M_k^s(\Theta) \right)^2 / \, \mathbb{V}\mathrm{ar}\left(M_k^d \right) \right] \right\} \tag{3.29}$$

   where K is the number of moments, M_k^d denotes the k-th data moment, and M_k^s the k-th simulated data moment using data from replication s.

5. Given the value of the objective function, the optimization algorithm then chooses new parameters.

6. Steps 1 - 5 are repeated until $\hat{\Theta} = \arg\min_{\Theta} f(\Theta)$ is found.

Note that equation (3.29) does not use the asymptotically optimal weighting matrix, because of its poor small sample properties (see Altonji & Segal, 1996). Instead, I use a diagonal matrix with sample variances of the respective moments as its elements. These variances are estimated using bootstrapping with clustering at the individual level.[75] Since the simulated choices are discrete outcomes, the objective function is a step function and does not possess valid derivatives at all points. Therefore, a pattern search method is employed, which is a derivative-free optimization routine. It is implemented here using the Dakota toolkit (see Adams et al., 2013) that allows for parallelization. Standard errors of Θ are estimated following Gourieroux et al. (1993).

[74] Furthermore, wages are only recorded when the simulated individual is employed and the original SOEP interview took place at the given period. To account for non-random missing wages, a linear probability model is used to fit the probability of not observing a wage given the state space variables. Simulated wages are then deleted according to this probability.

[75] I use 1001 replications following Davidson & MacKinnon (2000).

3.7 Empirical Results

3.7.1 Goodness of fit

Figures 3.6, 3.7, and 3.8 illustrate the fit of the model. While the observed data is presented in solid blue lines, the estimated model is shown in dashed magenta lines. The model nicely reflects the employment behavior over the life-cycle. While part-time work mostly increases over the life-cycle, full-time work decreases until the age of 35. Although the utility function does not include age related terms, the model can reproduce these trends. The model has some issues fitting the very last periods, which might be related to the model not including a retirement decision, but rather ending at age 50. Thus, toward the end of the model, dynamics introduced by human capital accumulation are slowly eliminated and choices only depends on the instantaneous utility-leisure trade-off. In real life, individuals might continue working, to accumulate more capital towards retirement.

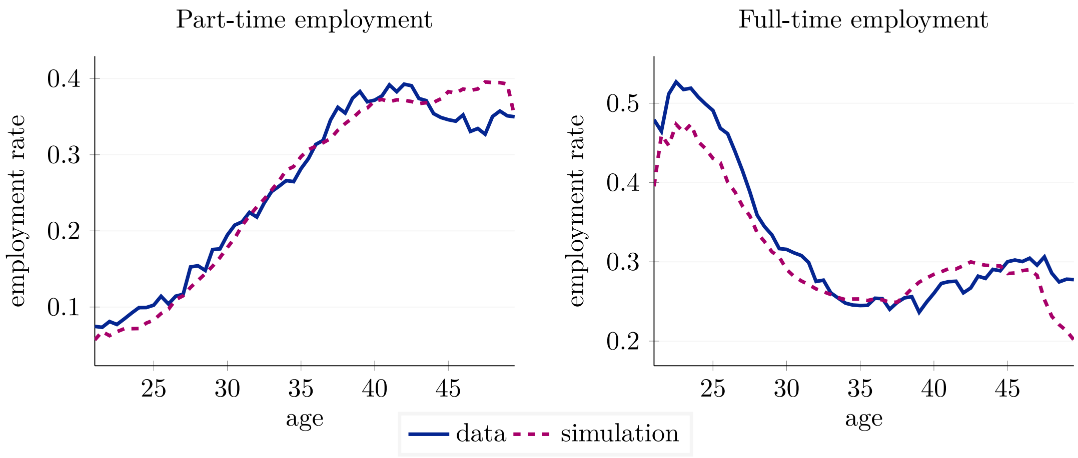

Figure 3.6: Employment rates by age

Notes: Comparison of the observed and simulated female employment rates over the life-cycle. Observed rates are in solid blue lines and are based on the SOEP data. Simulated rates are in dashed magenta lines and are based on the estimated model.

An important part for the estimation of career costs caused by child related employment interruptions is a close fit of employment behavior around childbirth. As figure 3.7 illustrates, the model can replicate these choices accurately. Before having a child, women are equally likely to be employed in the data and in the simulations. This is important, since only employed individuals have the right to return to employment during maternity leave. In particular, the model closely fits the first three years after the birth of a child. These are the most important moments for identification of the job offer beliefs. Overall, the model can replicate the larger return to part-time employment within the first five years, while also fitting the overall trend in full-time employment.

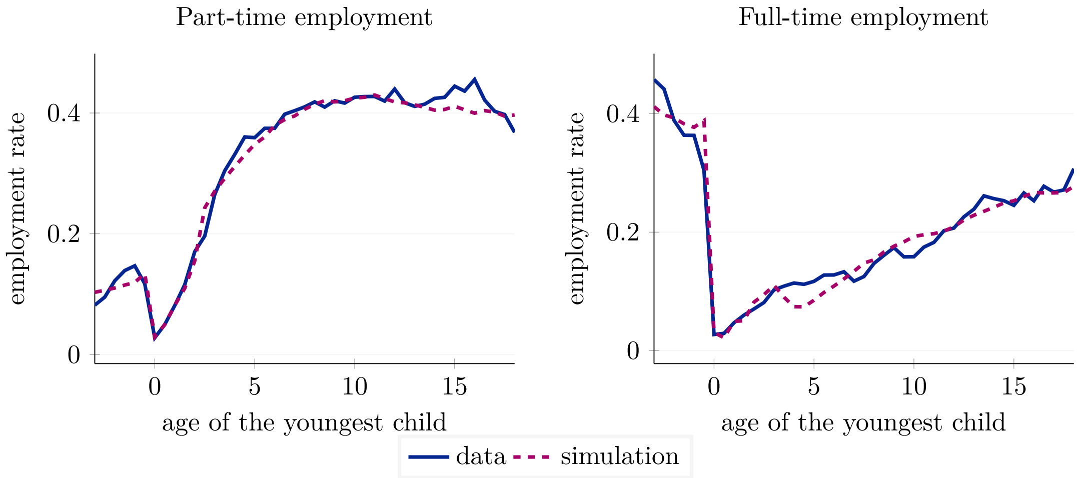

Figure 3.7: Employment rates by time to/since childbirth

Notes: Comparison of the observed and simulated female employment rates by time to/since childbirth. Observed rates based on SOEP data are in solid blue lines. Simulated rates estimated based on the estimated model are in dashed magenta lines.

In addition to the employment rates, the model is able to replicate average wages over the life-cycle. The average wage increase at the beginning of the working life until the age of 30 is especially well fitted. Afterwards, wages in both the simulation and the data stagnate. At the end of the life-cycle, average wages only slightly increase, which is replicated by the model. Overall, the good fit of wages insure that the costs of biased beliefs can be accurately estimated.

3.7.2 Parameter Estimates

Before discussing the parameters of the job offer probabilities and their expectations, it is helpful to look at some of the estimated parameters of the utility function and wage process that are reported in tables 3.6 and 3.7, respectively. The risk preference parameters γ_c is slightly higher than values usually found in the literature,[76] indicating a higher risk aversion of the sampled women compared to other studies. Because the parameters is greater than one, the utility function becomes negative for all values, thus the higher a parameter in the exponential, the lower the total level of utility. Considering the values of the estimated parameters, $\gamma_{part\text{–}time}$ is smaller than $\gamma_{full\text{–}time}$, indicating that having less leisure time, in general, decreases utility. The age of the youngest child has a different effect on utility from part-time and full-time leisure.

[76]For example, Haan & Prowse (2014) estimate a CRRA risk aversion parameter of about 2.54 in a study regarding labor supply and retirement decisions using SOEP data.

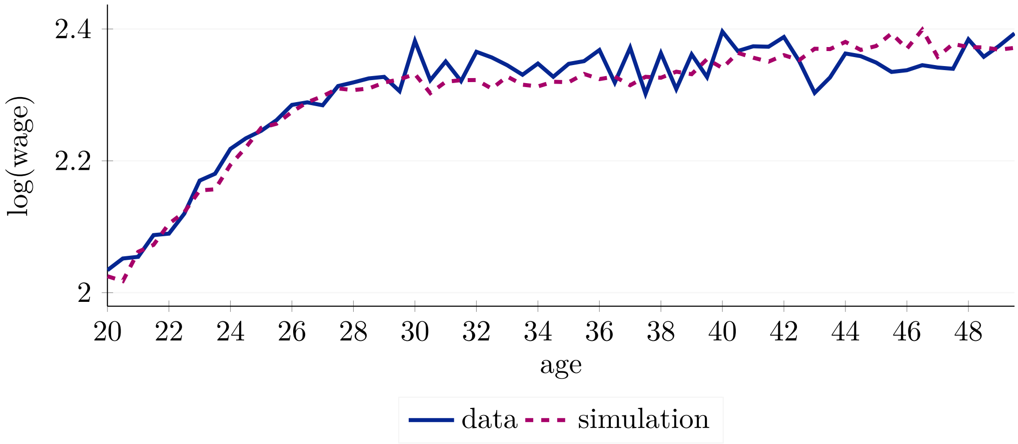

Figure 3.8: Mean log wage rates over the life-cycle

Notes: Comparison of the observed and simulated log wages over the life-cycle. Observed log wages are in solid lines. Log wages simulated based on the estimated model are in dashed magenta lines.

The estimated parameters for leisure preferences, depending on the age of the youngest child, imply that as the child gets older, especially after the age of ten, utility from part-time leisure slightly decreases. In contrast, for utility derived from working full-time increases for almost all ages of the youngest child. It is also notable that the parameters for the presence of a partner and persistence are both small and insignificant for part-time, but large and significant for full-time work.

Table 3.7 reports the estimated parameters of the wage function and the human capital process. Individuals without any on-the job human capital receive average wages of €6.69.[77] An additional year of full-time employment increases wages by 21.2% at the beginning of the working life, while after working full-time for ten years, an additional year only results in a wage increase of about 1.8%. Human capital depreciates with a yearly rate of 4.7%, which is important when regarding part-time employment. Human capital accumulation for part-time workers corresponds to only 5.7% of the full-time accumulation. Working part-time can only increase wages for individuals with a very low human capital stock, while for other groups, human capital accumulation in part-time work can barely compensate for depreciation and, thus, might even lead to losses in human capital. Blundell et al. (2016) term this effect in combination with the persistence in working choices, the part-time penalty. In general, the estimated parameters are close to their findings, although for Germany and not the United Kingdom.[78]

[77] The value is for 2000 and corresponds to $10.50 in 2018.

[78] Blundell et al. (2016) find slightly lower values for $\gamma_{wage,\,e}$ of 0.152 for women with secondary education and 0.229 for women with high school education. For η, their yearly values translate to half-yearly values of 0.0414 (secondary education) and 0.0289 (high school education) and, respectively, for

Table 3.6: Estimates of preference parameters

	Part-Time Employment		Full-Time Employment	
	Coeff.	St. Error	Coeff.	St. Error
	(1)	(2)	(3)	(4)
General ($\gamma_{\cdot}$)	0.306296	(0.012680)	0.457867	(0.025300)
Partner ($\gamma_{\cdot,\,partner}$)	0.002645	(0.058486)	−0.947004	(0.025304)
Y. child's age, cons. ($\gamma_{\cdot,\,ac0}$)	0.041199	(0.058486)	−0.581309	(0.012682)
Y. child's age, lin. ($\gamma_{\cdot,\,ac1}$)	−0.022504	(0.000376)	0.021216	(0.000414)
Y. child's age, qu. ($\gamma_{\cdot,\,ac2}$)	0.003229	(0.000075)	−0.003966	(0.000117)
State dependency	0.029302	(0.058473)	0.981887	(0.012682)
Risk preferences (γ_c)	2.997577 (0.051999)			

Notes: The table reports estimated preference parameters. The first row reports the overall taste for part-time and full-time employment, the second row the taste for part-time and full-time when a partner is present. Rows three, four and five report the quadratic modeled preferences for both employment states depending on the age of the youngest child present. Row six reports the added utility when choosing the same employment state as last period and row seven the CRRA parameter.

In addition to parameters for the real offer rate for part- and full-time employment, the parameter determining expectations (α) is reported in table 3.8. For an easier interpretation of these results, the yearly real and expected offer rates are depicted in figure 3.9. Without any potential non-employment time (ρ), individuals receive offers for part-time employment with a probability of 17.1% and for full-time employment with a probability of 50.8% in a given year.[79] While the probability of receiving a part-time offer only decreases slightly over time, the probability of full-time employment strongly decreases with more potential years spent non-employed. With ten years of potential non-employment, the yearly offer rates drop to 9.0% for part-time employment and 19.9% for full-time employment.

Agents have strongly biased beliefs regarding their opportunity to be able to return to employment, with an overestimate of roughly 66%. Given the small standard error, a null-hypothesis that individuals have rational expectations can be rejected at all common significance levels. The estimated parameter also aligns perfectly with the evidence presented in section 3.4. If the two-year values of table 3.4 are transformed

λ to half-yearly values of 0.0766 and 0.0487.

[79]Since the decision period in the model is semi-annual, the yearly job offer rates are computed as follows: $\pi^{yearly} = 1 - (1 - \pi^{half\text{-}yearly})^2$.

Table 3.7: Wage and human capital parameters

	Coeff. (1)	St. Error (2)
Intercept ($\gamma_{wage,\,const.}$)	6.6924	(0.007703)
Returns to experience ($\gamma_{wage,\,e}$)	0.2786	(0.003056)
Depreciation rate (η)	0.0240	(0.000961)
Human capital accum. while in part-time (λ)	0.0285	(0.001635)
Variance wage shock (σ_{xi})	0.2498	(0.002937)

Table 3.8: Employment offers

	Part-Time Employment		Full-Time Employment	
	Coeff. (1)	St. Error (2)	Coeff. (3)	St. Error (4)
$\gamma_{JO\cdot,c}$	−2.318044	(0.013012)	−0.855202	(0.001692)
$\gamma_{JO\cdot,\rho_1}$	−0.084939	(0.000336)	−0.087805	(0.001036)
$\gamma_{JO\cdot,\rho_2}$	0.001395	(0.000059)	−0.004080	(0.000058)
α		1.657566	(0.013012)	

to half-yearly values, the average α is about 1.31. However, as mentioned earlier, this value might be driven by individuals who are long-term unemployed and have lost faith in finding new employment. For groups with higher expectations of future employment opportunities, the half-yearly α lies between 1.67 and 2.43.[80] The estimate from the structural model aligns greatly with the sample after excluding the most pessimistic individuals.

3.7.3 Overconfidence vs. rational Expectations

Given the estimated parameters, it is possible to quantify the costs of biased expectations. To do so, I randomly draw 10,000 individuals from my sample and simulate their life-cycle decisions, once with biased expectations and once with rational expectations,

[80] For the group stating a probability greater than or equal to 30% for finding employment within two years, the half-yearly α equals 1.67, for the group stating the probability greater than or equal to 50%, it is 1.79, and for the group stating the probability greater than or equal to 80% it is 2.43.

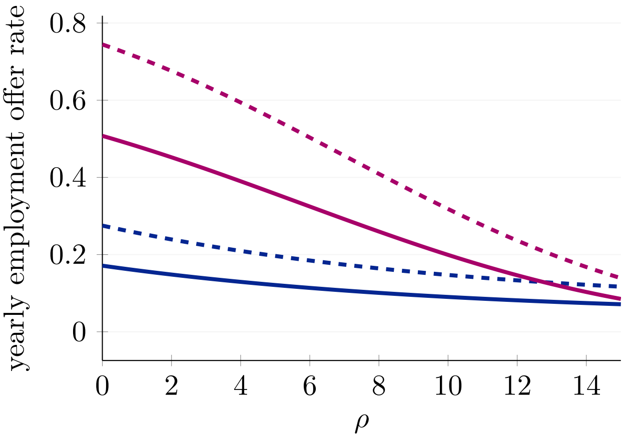

Figure 3.9: Employment offers and expectations

Notes: Employment offers varying with potential non-employment time ρ. Offer for part-time employment are in blue, while offers for full-time employment are in magenta. Real offer rates are in solid lines, expected rates are in dashed lines.

holding everything else constant. Table 3.9 reports the differences between biased and rational expectations regarding child-related career breaks. In regime I, overconfidence causes these breaks to last 5.8 months longer, an increase of 11.2%. In regime II, breaks are 7.8 months longer, corresponding to an increase of 17.6%, and in regime III, biased beliefs result in breaks lasting 4.7 months longer, a rise of 13.9%. All these numbers include mothers who entered employment at some point after having a child. The results of rows four and five of table 3.9 show that biased beliefs also affect behavior with respect to entering employment at all after having a child. For regime I, over 20.4% fewer women enter employment after having a child when having biased beliefs. Due to the higher percentage of working mothers in the other two regimes, their numbers are even higher. In regime II, 40.6% more mothers never return to employment after having a child when overestimating future employment opportunities, while in regime III, this increase is about 29%.

To quantify the costs of biased beliefs, I compute the net present values of earnings, income and consumption, from the birth of the first child through the end of the model. Table 3.10 provides an overview of these costs depending on the respective regime. Column (1) refers strictly to labor market earnings and does not include any benefits. Depending on the regime, life-cycle earnings decrease between 18% and 12%, when overestimating job offer probabilities. From a public economics point of view, these results are striking, since lower labor market earnings correspond to lower taxable

Table 3.9: Overconfidence and child related career breaks

	Regime I	Regime II	Regime III
	(1)	(2)	(3)
Avg. career break ($\alpha = 1.66$)	4.8004	4.3241	3.1776
Avg. career break ($\alpha = 1$)	4.3137	3.6776	2.7900
Share of mothers, not enter. emp. ($\alpha = 1.66$)	15.66%	13.19%	7.64%
Share of mothers, not enter. emp. ($\alpha = 1$)	13.00%	9.38%	5.92%

Notes: Results are based on 10,000 simulated life-cycles, once with the estimated beliefs of job offer expectations ($\alpha = 1.66$) and once with rational expectations ($\alpha = 1$). Average career breaks are in years.

income. However, costs are lower for individuals since maternity leave, unemployment, and child benefits mitigate some of the lost earnings. As column (2) of table 3.10 shows, individual income only drops by 4.13% for regime I, 7.39% for regime II, and 5.06% for regime III. These numbers are still considerably large, compared to overall career costs of children of 35% found by Adda et al. (2017). Accommodating for the potential partners' income only slightly reduces the career costs. As stated before, husbands are assumed to work full-time and do not interrupt their working careers due to the birth of a child. Since most women do work part-time, their contribution to the overall household earnings is also not particularly high under rational expectations, thus limiting the reduction in consumption when women have biased beliefs.

3.8 Conclusion

The birth of a child strongly impacts the working careers of women, especially since a majority of mothers remain at home for an extended period of time before re-entering employment. The length of such career breaks is influenced by the expectations of future employment possibilities. Overestimating these possibilities might cause mothers to not return within the maternity leave period during which their employment is protected, although non-optimal with rational expectations. Thus, upwards biased beliefs cause longer career interruptions, a higher fraction of mothers never returning to employment, and higher career costs of children.

I develop a structural life-cycle model of female labor supply and human capital accumulation, allowing for non-rational expectations of future job arrival rates. To identify expectations within the model, I derive a novel identification strategy that allows re-

Table 3.10: Costs of biased expectations

	Earnings (1)	Income (2)	Consumption (3)
Regime I	−16.36%	−4.13%	−3.39%
Regime II	−18.28%	−7.39%	−4.14%
Regime III	−12.41%	−5.06%	−3.07%

Notes: Results are based on 10,000 simulated life-cycles, once with the estimated beliefs of job offer expectations ($\alpha = 1.66$) and once with rational expectations ($\alpha = 1$). Earnings refers to direct labor market earnings. Income additionally includes maternity leave, unemployment and child benefits. Consumption includes income and additionally income of the partner, but are weighted by household members like in the utility function. Values correspond to net present values from the birth of the first child until the end of the model.

covering the key parameters from observed labor supply choices. The strategy exploits a discontinuity in the future expected value of non-employment caused by the end of employment protection. In combination with maternity leave reforms that change the duration of the employment protection after the birth of a child, it is possible to separately identify expectations, job-arrival rates, and preferences. I estimate the model using survey data from the German Socio-Economic Panel Study, since the German setting provides the necessary variation for identification.

Indeed, estimations show that mothers highly overestimate their chances on the labor market, since they expect the half-yearly job arrival rate to be 66% higher than the actual rate, on average. Comparing simulations with the estimated preference parameters – one restricted to rational expectations and one with the estimated expectations – shows that overconfidence prolongs career breaks between 4.7 and 7.8 months on average, depending on the length of employment protection. This results in a reduction of lifetime earnings from employment between 12% and 18%. Some of these losses are mitigated by various benefits and earnings of a potential husband, thus actual consumption losses only range from 3.1% to 4.1%.

The results have important implications from a public economics perspective. In addition to increasing social security spending, prolonged career breaks cause lower earnings from employment that translate directly into forgone tax revenue. The consequences for the individual are also substantial. The income loss causes reduced pension benefits, thus, contributing to an increased risk of poverty in retirement. Since biased expectations can be interpreted as market failures and because of their far-reaching consequences, interventions by policy makers might be justified. Possible policies could aim to provide better information about labor market conditions, for example by sending official information letters to new families, or financial incentives to return within the period the individual's job is protected.

Overall, some caution is appropriate when interpreting these results, since the model only estimates an average bias for all individuals and does not model an explicit retirement decision. Adding heterogeneity in expectations might be valuable, since individuals most likely exhibit different degrees of overconfidence causing different magnitudes of career costs. While heterogeneity based on observables can be estimated using the presented strategy, including unobservable heterogeneity demands a strong refinement of the identification approach. Extending the model by including a retirement decision presumably results in an increase in the costs of overconfidence, since the lost earnings from employment result in a lower average pension income. However, it is not immediately evident if individuals will postpone their retirement to overcome these losses. Future work might incorporate these elements into the model and, to simulate the effects of possible policies, aim to reduce the cost of overconfidence. These can be, for example, an increase in in-work benefits conditioned on returning within the employment protection or further prolonging employment protection without prolonging maternity benefits.

General Conclusion

The three chapters of this dissertation provide new insights in modeling and estimating dynamic discrete choice models. Building on the previous literature on identification issues, several strategies are presented to estimate important aspects of dynamic decision processes. A focus lies on hyperbolic discounting and biased expectations, two elements that are typically ignored in the vast majority of literature on female life-cycle employment. Furthermore, the empirical analyses demonstrate that individual behavior is better described by the proposed models than by the dominant models in the literature. It is also shown that these elements have an economic meaning, as they increase the costs of child-related career breaks.

The first chapter provides new exclusion restrictions to identify the exponential discount factor in dynamic discrete choice models. It demonstrates how exogenous changes in restriction probabilities can be used to recover time preferences from choice data. These restriction probabilities are common for this model class, as they, for example, represent job offers in a model of labor supply. Unemployed individuals have to rely on these offers to be able to choose employment from their choice set. If they do not receive an employment offer, they are restricted to stay non-employed. The chapter first discusses under which conditions potentially unobserved restriction probabilities, like job offers, can be recovered from the data. It is proven that if at least one choice exists that does not cause a potentially restricted choice set in the subsequent period, then all restriction probabilities can be recovered from observed choices. Following the labor supply example, this condition is fulfilled if employed individuals do not have to rely on a job offer to continue to work in the next period. For other cases, it is necessary to know at least one restriction probability to identify the remaining probabilities. Because this condition does not guarantee identification, chapter 1 also discusses the most common cases for which it can be achieved.

Once the restriction probabilities are known, exogenous variations of these can be exploited to identify the exponential discount factor. In the context of labor supply, this can be provided by a labor market policy that helps to bring firms and unemployed individuals together, thus, temporarily increasing the job offer rate. For the exclusion restriction to hold, changes in the restriction probabilities are not allowed to directly impact the utility function. The underlying idea is that when comparing individuals with different restriction probabilities in the future, but equal utility functions in the present, changes in choice probabilities can be linked to how economic agents account for their future in their decision making. Interestingly, the derived formula that identifies the exponential discount factor does not depend on functional form assumptions of the utility function. It only depends on the changes in the restriction probabilities and the observed transition choice probabilities.

Chapter 2 further extends this identification strategy and analytically shows that parameters of a hyperbolic discounting model can be recovered within a three-period model of dynamic discrete choice. Although, for models with more periods, no analytical expression can be derived for any parameter, it is argued that identification carries over. The underlying concept is to compare observed choice probabilities of two groups with different restriction probabilities at different points in time. When it is possible to observe at least two of these points, the present bias, in addition to the exponential discounting parameter of future periods, can be backed out. This is a direct result of the different weights used for utilities of different future periods.

In contrast to the first chapter, the derivation of an identification strategy is complemented with an empirical analysis. Focusing on female labor supply, and especially on child-related career breaks, a dynamic discrete choice model of female labor supply, human capital accumulation, and labor market frictions is developed. Women choose semi-annually from a choice set including: non-working, part-time employment, and full-time employment. When previously non-employed, individuals have to receive a job offer to be able to choose part-time or full-time employment. The chapter exploits multiple maternity leave reforms that provided different time horizons for employment protection. Comparing the employment behavior of mothers in these different regimes identifies both parameters of a model with quasi-hyperbolic time preferences.

To estimate the model, we use data from the German Socio-Economic Panel, finding a present-bias of 0.77 that significantly deviates from an exponential model, which assumes a parameter value of 1. Interestingly, the yearly discount factor for two subsequent future periods is close to values usually assumed for exponential discounters in the literature, with a value of 0.92. This result implies that women do not tend to be myopic per se, since even utilities 20 years away receive an overall weight of 0.137 in the lifetime utility. Rather, it is that instantaneous utility is valued much more than any future utility, thus providing evidence that women postpone their re-entry into the labor market, by overvaluing current leisure over future career advancements.

The first two chapters focus on identifying average time preferences within a population. For future research, it might be interesting to focus on identification of distributions of time preference parameters, since heterogeneity should be expected. In addition, chapter two assumes that individuals are naïve hyperbolic discounting, that is they are not aware of their time-inconsistent behavior. Suggestive evidence provided at the beginning of the chapter seems to support this hypothesis, since women strongly underestimate the length of their career breaks, even when enjoying employment protection for three years. However, it might be interesting to estimate the degree of naïvete, since commitment devices have much more power when individuals are aware of their time-inconsistencies. Future research might also look more into the career costs of hyperbolic discounters and how policies can help to mitigate these.

Chapter 3 concentrates on another important margin of dynamic choice, expectations about future states of the world. It represents one of the first analyses in the literature on female labor supply that deviates from the rational expectations hypothesis. This hypothesis assumes that individuals, on average, correctly predict future events. After presenting some suggestive evidence that economic agents systematically mispredict their future employment opportunities, a structural life-cycle model of labor supply, human capital accumulation, and labor market frictions is developed. Nesting the rational expectations framework, it allows job offer expectations to deviate from their real probability. Overestimating these probabilities might cause mothers to not return within the maternity leave period, when their employment is protected, although non-optimal with rational expectations. Thus, upwards biased beliefs cause longer career interruptions, a higher fraction of mothers to never return to employment, and higher career costs of having children.

To recover expectations within the model, an identification strategy is developed that allows estimating the key parameters from observed labor supply choices. The strategy exploits a discontinuity in the future expected value of non-employment caused by the end of an employment protection. Similar to chapter 2, multiple maternity leave reforms are exploited to separately identify preferences, real job offer rates, and expectations. Estimations provide evidence that mothers strongly overestimate their chances on the labor market, since they expect the half-yearly job arrival rate to be 66 % higher than the actual rate. The structural framework allows for quantifying the costs stemming from these biased beliefs by comparing simulations, one time with the estimated expectations and one time with rational expectations. The results are striking. Overestimating future employment opportunities prolongs career breaks on average between 4.7 and 7.8 months, depending on the length of employment protection. This causes a reduction of lifetime earnings from employment, as measured at the first birth of a child, between 12% and 18%. Some of these losses are mitigated by various benefits and the earnings of a potential husband, thus consumption losses only range from 3.1 % to 4.1 %.

From a public economics perspective, these results have important implications, since longer career breaks lower tax revenues and increase social security spending. From an individual perspective, the losses in earnings from employment also translate to lower pension benefits, thus contributing to an increased risk of poverty in retirement. Policies seeking to mitigate these negative results stemming from biased expectations, should consider providing better information about the labor market conditions to new families, for example by sending official information letters. Another possible policy approach is to provide target financial incentives to return within the period when the individual's job is protected. In addition to investigating these policy interventions, future research could focus on recovering a distribution of expectations, since chapter 3 only identifies the average bias. Adding heterogeneity in expectations might be valuable, since individuals most likely exhibit different degrees of overconfidence causing different magnitudes of career costs. Building on the presented identification approach, it might also be beneficial to look at pension income expectations in retirement models, thus improving recommendations for reforming pension systems.

Appendix A

Appendix to Chapter 2

Appendix A.1: Arrival of children, childcare costs, and job destruction

The arrival probability of a child depends on the woman's education, age, and the presence of other children in the household. To estimate these probabilities, linear probability models that are separately estimated for both education groups and for the presence of other children are used. For the arrival of a first child, a polynomial of order three in women's age is employed, while for additional children, an additional polynomial of order two for the age of the present sibling, and an interaction term of the woman's age and the sibling's age are estimated. Figure A.1 compares the estimated with the observed process. The figure reports the fraction of mothers over the lifecycle

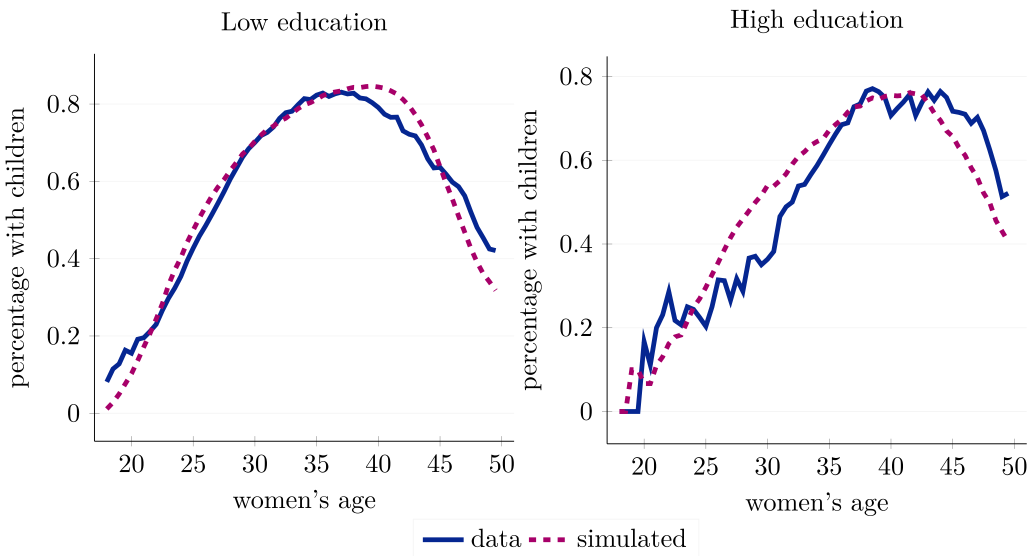

Figure A.1: Share of mothers

Notes: Data in solid blue lines, simulations in dashed pink lines.

for the observed data and the simulated processes. The estimated linear probability models approximate having children well and the overall fit is reasonably tight for both education groups.

Childcare costs are directly estimated from the data and are for a day for children under three €9.72 when working part-time, and €21.04 when working full-time. For children between 3 and 6, the daily costs are lower at €4.76 for part-time work and €8.88 for full-time.

The half-yearly probability of an involuntary job separation is estimated using answers of individuals who transitioned out of employment. The SOEP asks these groups to list specific reasons for their transitions, like plant closures, lay-offs, and personal reasons, among others. We use only exogenous causes to estimate a semi-annual separation rate of 3.23%.

Appendix A.2: Overview of Moments

Table A.1: Employment rates by education

	Low Education				High Education			
Moment	Data	Sim	SE Data	SE Diff	Data	Sim	SE Data	SE Diff
	Full-time Employment Rates							
Overall	0.3440	0.3097	0.0080	4.3030	0.3994	0.3433	0.0215	2.6009
Mothers	0.1527	0.1399	0.0083	1.5485	0.1495	0.1266	0.0203	1.1285
Youngest child [0, 3[	0.0525	0.0404	0.0056	2.1565	0.0567	0.0812	0.0159	1.5397
Youngest child [3, 6[	0.1217	0.0831	0.0100	3.8567	0.1143	0.0979	0.0242	0.6778
Youngest child [6, 11[	0.1539	0.1471	0.0111	0.6095	0.1833	0.1102	0.0327	2.2333
	Part-time Employment Rates							
Overall	0.2416	0.1999	0.0068	6.1674	0.2324	0.1968	0.0154	2.3182
Mothers	0.3125	0.2842	0.0096	2.9423	0.3244	0.3223	0.0235	0.0890
Youngest child [0, 3[	0.1049	0.1026	0.0068	0.3421	0.1311	0.1479	0.0206	0.8157
Youngest child [3, 6[	0.3315	0.2229	0.0138	7.8867	0.3514	0.2203	0.0361	3.6302
Youngest child [6, 11[	0.4118	0.4576	0.0153	3.0019	0.4317	0.4458	0.0382	0.3694

Notes: This table reports average employment rates. Data moments and moments from the estimated model are presented. SE Data lists the standard errors of the data moment, based on 1001 bootstraps. SE Diff lists the deviation from the data and the simulations in terms of standard errors of the data moment.

Table A.2: 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	0.0453	0.0171	0.6867	0.0075	0.0266	0.0074	2.5797	0.0228	0.0187	0.0048	0.8470
7-12 months	0.0670	0.0446	0.0162	1.3819	0.0286	0.0394	0.0145	0.7507	0.0241	0.0260	0.0053	0.3586
13-18 months	0.0919	0.0343	0.0225	2.5531	0.0976	0.0611	0.0215	1.7018	0.0280	0.0343	0.0055	1.1478
19-24 months	0.0865	0.0282	0.0205	2.8426	0.0683	0.0207	0.0175	2.7276	0.0644	0.0613	0.0079	0.3837
25-30 months	0.1027	0.0280	0.0229	3.2621	0.0585	0.0231	0.0175	2.0258	0.0672	0.0849	0.0082	2.1648
Part-time Employment Rates												
0-6 months	0.0104	0.0411	0.0071	4.3370	0.0522	0.0453	0.0245	0.2832	0.0342	0.0547	0.0059	3.4639
7-12 months	0.0491	0.0629	0.0155	0.8942	0.0286	0.0632	0.0179	1.9320	0.0649	0.0800	0.0080	1.8905
13-18 months	0.0919	0.0363	0.0219	2.5371	0.0732	0.0751	0.0188	0.1023	0.0951	0.1103	0.0100	1.5164
19-24 months	0.1622	0.0548	0.0285	3.7670	0.1024	0.0332	0.0218	3.1783	0.1539	0.1592	0.0120	0.4355
25-30 months	0.1351	0.0683	0.0286	2.3401	0.0976	0.0382	0.0218	2.7214	0.1791	0.1643	0.0129	1.1465

Notes: This table reports average employment rates depending on the employment protection regimes. Data moments and moments from the estimated model are presented. SE Data lists the standard errors of the data moment, based on 1001 bootstraps. SE Diff lists the deviation from the data and the simulations in terms of standard errors of the data moment.

Table A.3: Log wage regressions on accumulated experience and lagged wages

Moment	Low Education				High Education			
	Data	Sim	SE Data	SE Diff	Data	Sim	SE Data	SE Diff
Constant	0.4464	0.4080	0.0415	0.9264	0.4343	0.4652	0.1152	0.2679
$\ln(w_{t-1})$	0.7846	0.7823	0.0121	0.1918	0.8534	0.8003	0.0290	1.8321
Log accumulated working years	0.1261	0.2144	0.0935	0.9455	-0.0788	0.1992	0.2492	1.1157
Lagged log accumulated working years	-0.1058	-0.1748	0.0832	0.8296	0.0646	-0.1752	0.2226	1.0772
Variance of residuals	0.0484	0.0008	0.0719	0.6610	0.0515	0.0008	0.3056	0.1657

Notes: This table reports coefficients from a log wage regression on the four listed covariates. Data moments and moments from the estimated model are presented. SE Data lists the standard errors of the data moment, based on 1001 bootstraps. SE Diff lists the deviation from the data and the simulations in terms of standard errors of the data moment.

Table A.4: Transition rate into employment

	Low Education				High Education			
Moment	Data	Sim	SE Data	SE Diff	Data	Sim	SE Data	SE Diff
Rate	0.0950	0.1155	0.0030	6.7951	0.0987	0.1227	0.0083	2.8905

Notes: This table reports transition rates from non-employment to employment. Data moments and moments from the estimated model are presented. SE Data lists the standard errors of the data moment, based on 1001 bootstraps. SE Diff lists the deviation from the data and the simulations in terms of standard errors of the data moment.

Table A.5: 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	1.9700	0.0164	3.1839	2.4855	2.3923	0.0399	2.3352
Variance	0.1860	0.0096	0.0100	17.6698	0.1231	0.0078	0.0277	4.1673

Notes: This table reports log wage rates and their variance at the entrance in the working life. Data moments and moments from the estimated model are presented. SE Data lists the standard errors of the data moment, based on 1001 bootstraps. SE Diff lists the deviation from the data and the simulations in terms of standard errors of the data moment.

Table A.6: Further wage moments

	Low Education				High Education			
Moment	Data	Sim	SE Data	SE Diff	Data	Sim	SE Data	SE Diff
				Full-time Workers				
Mean	2.3140	2.2498	0.0094	6.8629	2.6189	2.6110	0.0262	0.3018
				Part-time Workers				
Mean	2.1074	2.2136	0.0140	7.6004	2.4038	2.5337	0.0432	3.0065
				First Differences log wage regression				
experience	0.2151	0.2151	0.0265	0.0017	0.1819	0.1901	0.0699	0.1173
				Yearly differences in log wages				
full-time	0.0243	0.0384	0.0024	5.8264	0.0261	0.0303	0.0063	0.6753

Notes: This table reports further log wage moments. The first two rows present average log wages of full-time and part-time workers, respectively. The third row presents the coefficient of a log wage regression on experience in first differences. The last row presents average yearly changes in the log wages. Data moments and moments from the estimated model are presented. SE Data lists the standard errors of the data moment, based on 1001 bootstraps. SE Diff lists the deviation from the data and the simulations in terms of standard errors of the data moment.

Appendix B

Appendix to Chapter 3

Appendix B.1: Suggestive Evidence: Additional Information

Appendix B.1.1 SOEP Questions

The suggestive evidence of section 3.4 is based on the following two SOEP questions. Only individuals who respond (b), (c), or (d) in the first question are asked the second question.

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

(a) No, definitely not

(b) Probably not

(c) Probably

(d) Yes, definitely

The original questionnaire also underlines the words “next two years” at the following question.

How likely is it that one or more of the following occupational changes will take place in your life within the next two years?

	0	10	20	30	40	50	60	70	80	90	100
Start paid work	□	□	□	□	□	□	□	□	□	□	□
Become self-employed	□	□	□	□	□	□	□	□	□	□	□
Receive further education	□	□	□	□	□	□	□	□	□	□	□

Appendix B.1.2 Registered Unemployment Rate for Germany between 1999 and 2007

A possible explanation for the gap between the stated likelihood to find employment and the realizations is that individuals were affected by a shared macro shock. If the economy falls into an unexpected recession, it is harder for everyone to find employment and, thus, it is natural to expect a gap between stated preferences and realizations. Figure B.1 shows that this is unlikely to drive the results of section 3.4. It plots the unemployment rate for the relevant years of table 3.4. The questions were first asked in 1999 and then again in 2001, 2003, and 2005. Although the SOEP interview can be at any time throughout the year, the majority of interviews are in spring. The possible interview times are indicated by the shaded pink areas in the figure. The gray area marks a recession according to the definition of the OECD.

Overall the unemployment rate did not fluctuate much, staying mostly around 10 %. In two of the four years the questions were asked, a recession followed, while in the other two years a decrease in the unemployment rate followed. Overall, this should at least partly balance the macro influence on the realizations.

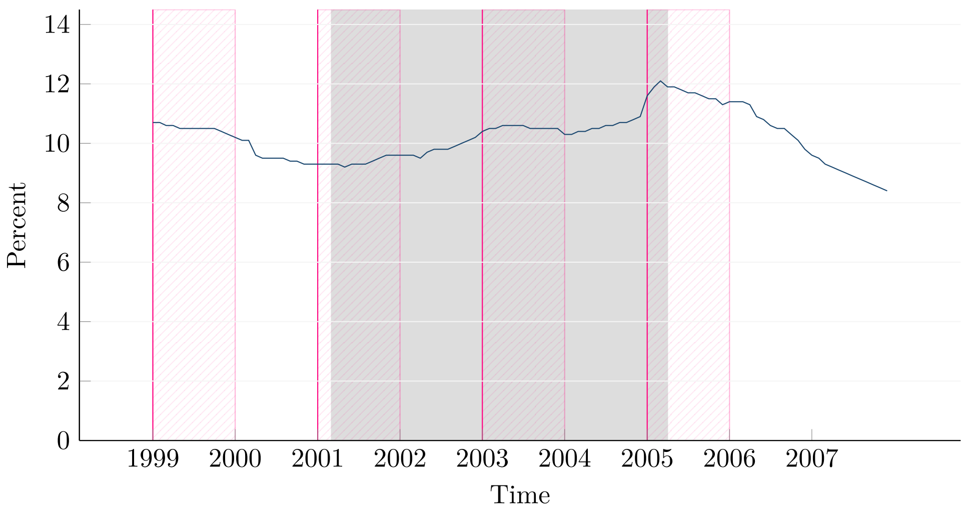

Figure B.1: Registered unemployment rate for Germany

Notes: The blue line depicts the monthly registered unemployment rate in Germany. The gray area marks recessions as defined by the OECD. The shaded pink areas indicate the year over which subjects were asked about their employment expectations. Sources: Federal Reserve Bank of St. Louis (2017a), Federal Reserve Bank of St. Louis (2017b).

Appendix B.2: External Processes

Appendix B.2.1 Family Dynamics

The underlying binary processes of marriage, divorce, and the arrival of children are modeled as linear probability models, the parameters of which are estimated via the method of simulated moments using the same optimization algorithm as discussed in subsection 3.6.2. When simulating these transitions, predicted probabilities below 0 are reset to 0, and predicted probabilities above 1 are reset to 1.

The arrival probability of husbands depends on a three-order polynomial in the woman's age. The divorce probability additionally depends on the presence of children. If no children are present, the probability only depends on a fourth-order polynomial of the woman's age, while if children are present, it additionally depends on a second-order polynomial of the age of the youngest child and an interaction term between the mother's and the youngest child's age.

The probabilities of the arrival of children are separately estimated for the first child, additional children and for the presence of a partner. The models for the first child include a second-order polynomial in woman's age if a partner is present and a fourth-order polynomial in the woman's age if no partner is present. For additional children, further terms are added. In a marriage, an additional birth depends on a fourth-order polynomial of the mother's age and a second-order polynomial of the age of the youngest child. Furthermore, interaction terms of the two ages up to the third-order are added. When there is no partner present, the probability depends on a fifth-order polynomial of the woman's age, a fourth-order polynomials of the youngest child's age and interaction terms of the two ages up to a fourth-order polynomial.[81] Women are not allowed to have a first child after the age of 37.5 and an additional child after 38 corresponding to 99 % of the observed births in the sample.

In addition to figures B.2 and B.3, table B.1 illustrates how the estimated processes (dashed lines) compare to the real data (solid lines). Figure B.2 shows how family types vary over the age of the mother. The estimated processes nicely fit marriages and the arrival of children over the whole sample. Similarly well fitted is the age of the youngest child, as well as the age of the youngest child when an additional child arrives. These processes are of special importance for two reasons. First, leisure preferences and, thus, employment rates depend strongly on the age of the youngest child. Second, as long as a mother has an additional child during her maternity leave period, the leave period resets.[82] As table B.1 shows, the distribution of the sibling's age when a new child is

[81] Since the model's decision period is a half-year, women are not able to have an additional child if the youngest child has not yet reached the age of one.

[82] This is fully integrated into the model.

born is tightly fitted by the simulated data.

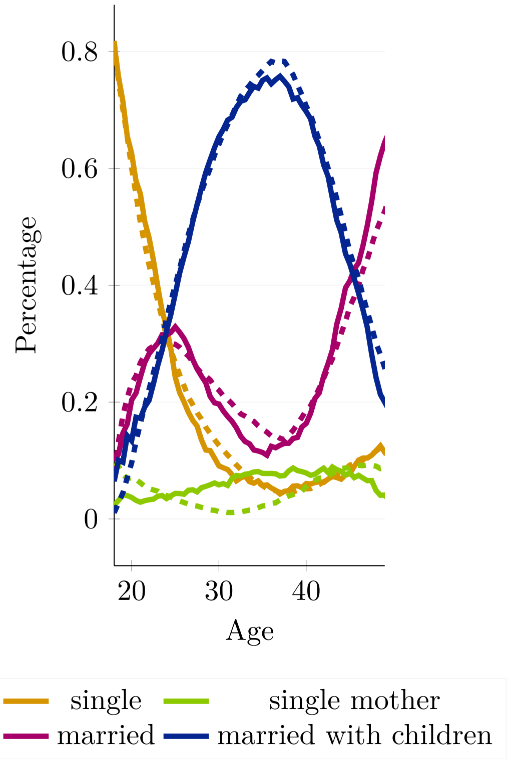

Figure B.2: Family dynamics

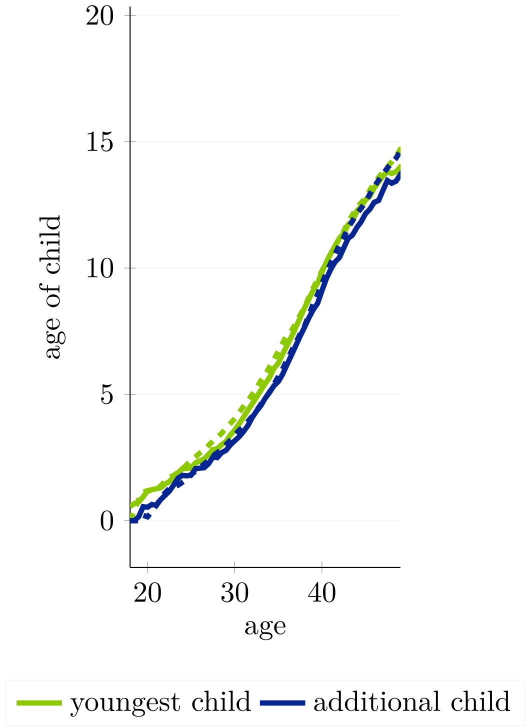

Figure B.3: Age of youngest child

Notes: Distribution of family types by age of woman. Data in solid lines, simulations in dashed lines.

Besides affecting women's leisure preferences, husbands add to the household income. Almost all husbands work over the whole lifecycle, which is indicated by an overall sample employment rate of 92 % with below 2 % working part-time. To estimate the husband's half-yearly gross income, the observed average income depending on the wife's age are weighted with the respective sample employment rate. These incomes are then used as moments to estimate a linear regression model via the method of simulated moments. Figure B.4 illustrates that the simulated values closely fit the observed incomes.

Table B.1: Additional child moments

	Low education			
Moment	Data	Sim	SE Data	SE Diff
(1)	(2)	(3)	(4)	(5)
share of add. children	0.5389	0.5442	0.0111	0.4790
age young. sibl. $<$ P10	0.1361	0.1704	0.0131	2.6161
age young. sibl. $<$ P25	0.2571	0.2707	0.0160	0.8447
age young. sibl. $<$ P50	0.5156	0.5119	0.0182	0.2073
age young. sibl. $<$ P75	0.7701	0.7527	0.0157	1.1077
age young. sibl. $<$ P90	0.9007	0.9128	0.0113	1.0808

Notes: Additional moments to estimate family dynamics. Row 1 reports the share of children born when another child is present. Rows 2 - 6 reports the share of additional children born when the youngest sibling's age under the, in the data observed, percentile.

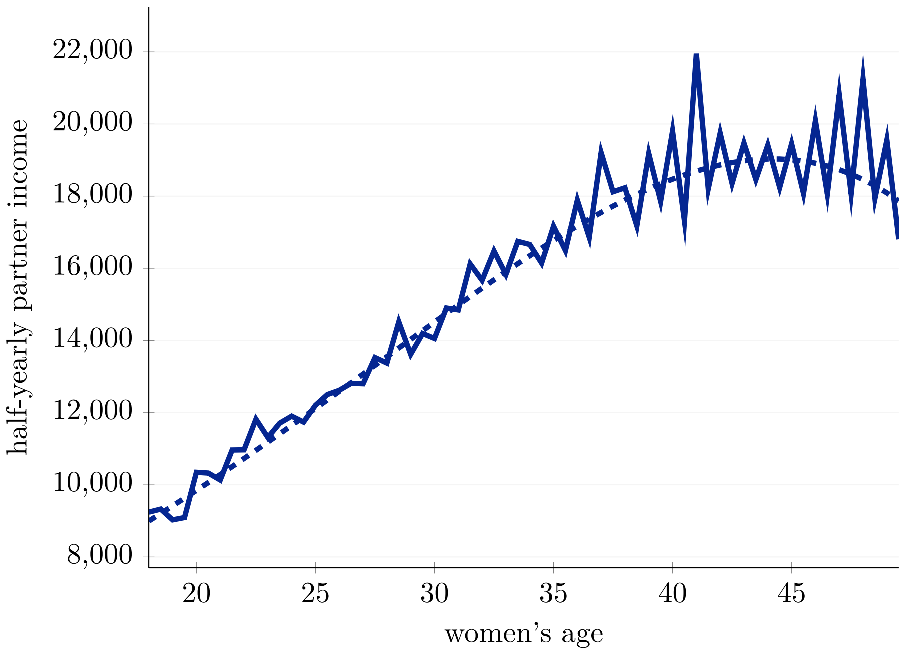

Figure B.4: Husband's income

Notes: Potential husband's half-yearly gross income. Data in solid lines, simulations in dashed lines.

List of Tables

List of Figures

German Summary

Diese Dissertation präsentiert neue Ergebnisse für die Modellierung und Schätzung von dynamischen diskreten Entscheidungsmodellen. Aufbauend auf der Literatur zu Identifikationsstrategien dieser Modellgruppe, werden neue Methoden erarbeitet, die eine Schätzung von wichtigen Aspekten der dynamischen Entscheidungsfindung ermöglichen. Ein Fokus liegt dabei auf der Identifizierung von Zeitpräferenzen in hyperbolischer Form und Erwartungen über zukünftige Arbeitsmöglichkeiten. Dies sind zwei Aspekte, die bisher weitgehend in den Analysen des Arbeitsangebotes von Frauen unberücksichtigt blieben. Mit empirischen Analysen zeigt die Dissertation, dass die neu vorgeschlagenen Modelle besser als die bisherigen Ansätze das Verhalten von Individuen beschreiben. Diese Ergebnisse sind ökonomisch von großer Bedeutung, da sie zeigen dass die Karrierekosten durch Kinder bisher unterschätzt wurden.
Das erste Kapitel zeigt neue Möglichkeiten auf, wie der exponentielle Diskontfaktor in dynamischen Entscheidungsmodellen identifiziert werden kann. Mit Hilfe exogener Variation in den sogenannten Einschränkungswahrscheinlichkeiten ist es möglich die Zeitpräferenzen ausschließlich aufgrund von beobachteten Entscheidungen zu schätzen. Als Einschränkungswahrscheinlichkeiten werden dabei die Wahrscheinlichkeiten bezeichnet, mit denen ein Individuum in zukünftigen Perioden in seiner Wahl eingeschränkt ist. Im Kontext von Arbeitsangebotsmodellen können Veränderungen in der Wahrscheinlichkeit ein konkretes Jobangebot zu erhalten zur Identifizierung genutzt werden. Diese stellen in einem dynamischen diskreten Entscheidungsmodell eine Einschränkungswahrscheinlichkeit dar, da ein in der jeweiligen Periode arbeitsloses Individuum ohne Jobangebot nicht die Möglichkeit hat, sich für eine Arbeitsaufnahme zu entscheiden. Die zugrunde liegende Idee der Identifizierung ist, dass Variation in der Einschränkungswahrscheinlichkeit zwar zukünftigen erwarteten Nutzen ändert, jedoch keinen Einfluss auf den aktuellen Nutzen hat. Es ist somit möglich, die Unterschiede im Entscheidungsverhalten zweier Gruppen mit unterschiedlichen Einschränkungswahrscheinlichkeiten auf die durchschnittlichen Zeitpräferenzen zurückzuführen. Das Kapitel leitete eine Formel zur Identifizierung des exponentielle Diskontfaktors her, die ausschließlich auf den beobachtbaren Entscheidungen und Übergangswahrscheinlichkei-

ten und somit nicht auf der funktionellen Form des Modells basiert.
Kapitel 2 erweitert den Ansatz des ersten Kapitels und zeigt Identifikationsstrategien auf, um hyperbolische Zeitpräferenzen zu identifizieren. Nachdem eine analytische Beweisführung für die Identifikation in einem Drei-Perioden-Modell diskutiert wird, wird diese auf ein Mehrperiodenmodellen erweitert. Die theoretische Analyse wird durch eine empirische Untersuchung ergänzt. Diese Untersuchung konzentriert sich auf das dynamische Arbeitsangebot von Frauen innerhalb eines diskreten Entscheidungsmodells, welches unter anderem die Entwicklung von Humankapital und Arbeitsmarktfriktionen berücksichtigt. Frauen wählen ihr Arbeitsangebot innerhalb dieses Modells jedes halbe Jahr. Die drei Wahlmöglichkeiten sind 1) keine Arbeit anbieten, 2) Teilzeit oder 3) Vollzeit arbeiten. Aktuell nicht beschäftigte Frauen, können dieses Arbeitsangebot jedoch nur realisieren, wenn sie ein Jobangebot erhalten. Für die Schätzung der Zeitpräferenzen werden mehrere Elternzeitreformen genutzt. Diese Reformen verlängerten jeweils den Zeitraum des Rückkehrrechts in die vorherige Arbeitsstelle und stellen somit zeitliche Variation dar, kein Jobangebot zu benötigen um in Arbeit zurückzukehren.
Das Modell wird mit Hilfe des Sozio-oekonomisches Panel (SOEP) geschätzt. Der Parameter, der die zusätzliche Diskontierung zwischen aktuellem Nutzen und zukünftigen Nutzen widerspiegelt, wird auf 0.77 geschätzt und ist signifikant unterschiedlich von 1, dem Wert in einem exponentiell diskontierendem Modell. Der exponentielle Diskontfaktor des geschätzten Modells, das heißt, die Diskontierung zwischen zwei zukünftigen Perioden, wird auf 0.92 geschätzt. Diese Ergebnisse zeigen deutlich, dass Frauen in diesem Modell zukünftigen Nutzen nicht einfach stark diskontieren, sondern sich vielmehr zeitinkonsistent verhalten. Dieses Verhalten kann zu Verzögerungen beim Wiedereintritt in das Erwerbsleben nach kinderbedingten Erwerbsunterbrechung führen, die negative Auswirkungen auf das Lebenszeiteinkommen haben.
Das letzte Kapitel beschäftigt sich mit den individuellen Erwartungen bezüglich zukünftigen Jobangeboten. Es stellt eine der ersten Analysen in der Literatur über das Arbeitsangebot von Frauen dar, welche von der klassischen Annahme der rationalen Erwartungen abweicht. Zunächst wird deskriptiv gezeigt, dass Frauen im Allgemeinen überschätzen, wie leicht es ist nach einer Erwerbsunterbrechung wieder in eine Beschäftigung zurückzukehren. Danach wird ein dynamisches Entscheidungsmodell entwickelt, welches dem in Kapitel 2 ähnlich ist. Einen wesentlichen Unterschied stellt die explizite Modellierung von Erwartungen dar, welche Abweichungen von der rationalen Erwartungshypothese zulässt. Mit Hilfe von Elternzeitreformen können Erwartungen, die realen Wahrscheinlichkeiten ein Jobangebot zu erhalten und Präferenzen getrennt identifiziert werden.
Erneut wird das vorgeschlagene Modell mit den beobachteten Entscheidungen im SOEP geschätzt. Ergebnisse zeigen, dass Frauen die Jobangebotswahrscheinlichkeit stark über-

schätzen, da sie im Durchschnitt ein Jobangebot 1,6 mal häufiger erwarten, als sie es in der Realität erhalten. Das strukturelle Modell erlaubt die Kosten dieser verzerrten Erwartungen zu quantifizieren. Die Ergebnisse zeigen, dass aufgrund der Verzerrung kinderbedingte Karriereunterbrechungen bis zu acht Monaten länger als unter nicht verzerrten Erwartungen sind und das Lebenszeit-Arbeitseinkommen um bis zu 18% sinkt, gemessen zum Zeitpunkt der Geburt des ersten Kindes. Teilweise werden diese Verluste durch Sozialleistungen und das Einkommen des Ehemannes kompensiert. Da die meisten Ehemänner weiterhin Vollzeit arbeiten, wenn Kinder im Haushalt präsent sind und das Steuersystem zweitverdienende Ehepartner stark besteuert, belaufen sich die Konsumeinbußen daher nur auf durchschnittlich 3,6%. Das Kapitel macht deutlich, dass die Annahme der rationalen Erwartungen in dynamischen Entscheidungsmodellen nicht aufrecht erhalten werden kann und das ökonomisch bedeutende Effekte entstehen, sobald diese gelockert wird.

Bibliography

Abbring, J. H. & Daljord, O. (2016). Identifying the Discount Factor in Dynamic Discrete Choice Models. *Centre for Economic Policy Research Discussion Paper 11133.*

Adams, B., Bauman, L., Bohnhoff, W., Dalbey, K., Ebeida, M., Eddy, J., Eldred, M., Hough, P., Hu, K., Jakeman, J., Swiler, L., & Vigil, D. (2013). *DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 5.4 User's Manual.* Technical report, SAND2010-2183.

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

Altonji, J. G. & Segal, L. M. (1996). Small-Sample Bias in GMM Estimation of Covariance Structures. *Journal of Business and Economic Statistics*, 14(3), 353–366.

Altug, S. & Miller, R. A. (1998). The Effect of Work Experience on Female Wages and Labour Supply. *Review of Economic Studies*, 65(1), 45–85.

Anderson, D. J., Binder, M., & Krause, K. (2002). The Motherhood Wage Penalty: Which Mothers Pay It and Why? *American Economic Review*, 92(2), 354–358.

Attanasio, O., Kovacs, A., & Molnar, K. (2017). Euler Equations, Subjective Expectations and Income Shocks. *Norwegian School of Economics Discussion Paper 05/107.*

Attanasio, O., Low, H., & Sanchez-Marcos, V. (2008). Explaining Changes in Female Labor Supply in a Life-Cycle Model. *American Economic Review*, 98(4), 1517–1552.

Baily, M. N. (1978). Some Aspects of Optimal Unemployment Insurance. *Journal of Public Economics*, 10(3), 379–402.

Bellman, R. (1957). *Dynamic Programming.* Princeton University Press.

Ben-David, I., Graham, J. R., & Harvey, C. R. (2013). Managerial Miscalibration. *Quarterly Journal of Economics*, 128(4), 1547–1584.

Bernheim, B. D. (1988). Social Security Benefits: An Empirical Study of Expectations and Realizations. *Issues in Contemporary Retirement*, (pp. 312–345).

Bhargava, B. S. & Manoli, D. (2015). Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment. *American Economic Review*, 105(11), 3489–3529.

Blau, F. D. & Ferber, M. A. (1991). Career Plans and Expectations of Young Women and Men: The Earnings Gap and Labor Force Participation. *Journal of Human Resources*, 26(4), 581–607.

Blundell, R., Costas-Dias, M., Meghir, C., & Shaw, J. M. (2016). Female Labour Supply, Human Capital and Welfare Reform. *Econometrica*, 84(5), 1705–1753.

Chan, M. K. (2017). Welfare Dependence and Self-Control: An Empirical Analysis. *Review of Economic Studies*, 84(4), 1379–1423.

Chetty, R. (2006). A General Formula for the Optimal Level of Social Insurance. *Journal of Public Economics*, 90(10-11), 1879–1901.

Cogan, J. F. (1981). Fixed Costs and Labor Supply. *Econometrica*, 49(4), 945–963.

Davidson, R. & MacKinnon, J. G. (2000). Bootstrap Tests: How Many Bootstraps? *Econometric Reviews*, 19(1), 55–68.

De Bondt, W. F. M. & Thaler, R. H. (1995). Financial Decision-Making in Markets and Firms: A Behavioral Perspective. In R. Jarrow, V. Maksimovic, & W. T. Ziemba (Eds.), *Handbooks in Operations Research and Management Science*, Volume 9, Chapter 13, (pp. 385–410). Amsterdam: North Holland.

DellaVigna, S., Lindner, A., Reizer, B., & Schmieder, J. F. (2017). Reference-Dependent Job Search: Evidence from Hungary. *The Quarterly Journal of Economics*, 132(4), 1969–2018.

DellaVigna, S. & Malmendier, U. (2006). Paying Not to Go to the Gym. *American Economic Review*, 96(3), 694–719.

DellaVigna, S. & Paserman, M. D. (2005). Job Search and Impatience. *Journal of Labor Economics*, 23(3), 527–588.

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

Dolls, M., Doerrenberg, P., Peichl, A., & Stichnoth, H. (2016). Do Savings increase in Response to Salient Information about Retirement and Expected Pensions? *NBER Working Paper 22684*.

Duflo, E. & Saez, E. (2003). The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment. *Quarterly Journal of Economics*, 118(3), 815–842.

Eckstein, Z. & Wolpin, K. I. (1989). Dynamic Labour Force Participation of Married Women and Endogenous Work Experience. *Review of Economic Studies*, 56(3), 375–390.

Ejrnaes, M. & Kunze, A. (2013). Work and Wage Dynamics around Childbirth. *Scandinavian Journal of Economics*, 115(3), 856–877.

Fang, H. & Silverman, D. (2009). Time-Inconsistency and Welfare Program Participation: Evidence from the NLSY. *International Economic Review*, 50(4), 1044–1077.

Fang, H. & Wang, Y. (2015). Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions. *International Economic Review*, 56(2), 565–596.

Federal Reserve Bank of St. Louis (2016). Organization for Economic Co-operation and Development, Consumer Price Index of All Items in Germany [DEUCPIALLMINMEI].

Federal Reserve Bank of St. Louis (2017a). OECD based Recession Indicators for Germany from the Period following the Peak through the Trough [DEUREC].

Federal Reserve Bank of St. Louis (2017b). Organization for Economic Co-operation and Development, Registered Unemployment Rate for Germany [LMUNRRTTDEM156S].

Francesconi, M. (2002). A Joint Dynamic Model of Fertility and Work of Married Women. *Journal of Labor Economics*, 20(2), 336–380.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Preference: A Critical Time Review. *Journal of Economic Literature*, 40(2), 351–401.

Gallant, R. & Tauchen, G. (1996). Which Moments to Match? *Econometric Theory*, 12(4), 657–681.

Garfield, J. & Ahlgren, A. (1988). Difficulties in Learning Basic Concepts in Probability and Statistics: Implications for Research. *Journal for Research in Mathematics Education*, 19(1), 44–63.

Gourieroux, C., Monfort, A., & Renault, E. M. (1993). Indirect Inference. *Journal of Applied Econometrics*, 8, 85–118.

Gronau, R. (1988). Sex-Related Wage Differentials and Women's Interrupted Labor Careers - the Chicken or the Egg. *Journal of Labor Economics*, 6(3), 277–301.

Gustman, A. L. & Steinmeier, T. L. (2012). Policy Effects in Hyperbolic vs. Exponential Models of Consumption and Retirement. *Journal of Public Economics*, 96(5-6), 465–473.

Haan, P. & Prowse, V. (2014). Longevity, Life-Cycle Behavior and Pension Reform. *Journal of Econometrics*, 178(3), 582–601.

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

Hamermesh, D. S. (1985). Expectations, Life Expectancy, and Economic Behavior. *Quarterly Journal of Economics*, 100(2), 389–408.

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

Heckman, J. J. & Macurdy, T. E. (1980). A Life Cycle Model of Female Labour Supply. *Review of Economic Studies*, 47(1), 47–74.

Heckman, J. J. & Macurdy, T. E. (1982). Corrigendum on A Life Cycle Model of Female Labour Supply. *Review of Economic Studies*, 49(4), 659–660.

Hoffman, M. & Burks, S. V. (2017). Worker Overconfidence: Field Evidence and Implications for Employee Turnover and Returns from Training. *NBER Working Paper 23240.*

Hurd, M. D., Smith, J. P., & Zissimopoulos, J. M. (2004). The Effects of Subjective Survival on Retirement and Social Security Claiming. *Journal of Applied Econometrics*, 19(6), 761–775.

International Labour Organization (2012). *Maternity and Paternity Leave - Law and Practive Across the World*. Technical report.

Kassenboehmer, S. C. & Schatz, S. G. (2017). Re-employment Expectations and Realisations: Prediction Errors and Behavioural Responses. *Labour Economics*, 44, 161–176.

Keane, M. P., Todd, P. E., & Wolpin, K. I. (2011). The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications. In O. A. Card & David (Eds.), *Handbook of Labor Economics*, Volume 4, Part A, Chapter 4, (pp. 331–461). Elsevier Inc.

Keane, M. P. & Wolpin, K. I. (1997). The Career Decisions of Young Men. *Journal of Political Economy*, 105(3), 473.

Keane, M. P. & Wolpin, K. I. (2010). The Role of Labor and Marriage Markets, Preference Heterogeneity, and the Welfare System in the Life Cycle Decisions of Black, Hispanic, and White Women. *International Economic Review*, 51(3), 851–892.

Kimmel, J. & Kniesner, T. J. (1998). New Evidence on Labor Supply: Employment versus Hours Elasticities by Sex and Marital Status. *Journal of Monetary Economics*, 42(2), 289–301.

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

Kleven, H. J. & Waseem, M. (2013). Using Notches To Uncover Optimization Frictions and Structural Elasticities: Theorie and Evidence from Pakistan. *Quarterly Journal of Economics*, 128(2), 669–723.

Kruger, J. & Dunning, D. (1999). Unskilled and Unaware of it: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134.

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*, 112(2), 443–478.

Low, H., Meghir, C., & Pistaferri, L. (2010). Wage Risk and Employment Over the Life Cycle. *American Economi Review*, 100, 1432–1467.

Magnac, T. & Thesmar, D. (2002). Identifying Dynamic Discrete Decision Processes. *Econometrica*, 70(2), 801–816.

Malmendier, U. & Taylor, T. (2015). On the Verges of Overconfidence. *Journal of Economic Perspectives*, 29(4), 3–8.

Manski, C. F. (1990). The Use of Intentions Data to Predict Behavior: A Best-Case Analysis. *Journal of the American Statistical Association*, 85(412), 934–940.

Manski, C. F. (2004). Measuring Expectations. *Econometrica*, 72(5), 1329–1376.

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

Mincer, J. (1962). Labor Force Participation of Married Women: A Study of Labor Supply. In Universities-National Bureau Committee for Economic Research (Ed.), *Aspects of Labor Economics*, Volume 1 (pp. 63–106). Princeton University Press.

Mirrlees, J. (1971). An Exploration in the Theory of Optimal Income Taxation. *Review of Economic Studies*, 38(2), 175–208.

Moore, D. A. & Healy, P. J. (2008). The Trouble with Overconfidence. *Psychological Review*, 115(2), 502–517.

O'Donoghue, T. & Rabin, M. (1999). Doing It Now or Later. *American Economic Review*, 89(1), 103–124.

OECD (2017). *Pensions at a Glance 2017: OECD and G20 Indicators*. Paris: OECD Publishing.

OECD (2018). *OECD Labour Force Statistics 2017*. Paris: OECD Publishing.

Phelps, E. S. & Pollak, R. A. (1968). On Second-Best National Saving and Game-Equilibrium Growth. *Review of Economic Studies*, 35(2), 185–199.

Pollak, R. A. (1968). Consistent Planning. *Review of Economic Studies*, 35(2), 201–208.

Rees-Jones, A. (2018). Quantifying Loss-Averse Tax Manipulation. *Review of Economic Studies*, 85(2), 1251–1278.

Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 55(5), 999–1033.

Rust, J. (1994). Structural estimation of markov decision processes. In *Handbook of Econometrics*, Volume 4 Chapter 51, (pp. 3081–3143).

Saez, E. (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, 2(3), 180–212.

Sandell, S. H. & Shapiro, D. (1980). Work Expectations, Human Capital Accumulation, and the Wages of Young Women. *Journal of Human Resources*, 15(3), 335–353.

Schoenberg, U. & Ludsteck, J. (2014). Expansions in Maternity Leave Coverage and Mothers' Labor Market Outcomes after Childbirth. *Journal of Labor Economics*, 32(3), 469–505.

Shaw, L. B. & Shapiro, D. (1987). Women's work plans: contrasting expectations and actual work experience. *Monthly Labor Review*, 110, 7–13.

Sheran, M. (2007). The career and family choices of women: A dynamic analysis of labor force participation, schooling, marriage, and fertility decisions. *Review of Economic Dynamics*, 10(3), 367–399.

Slovic, P. (2000). *The Perception of Risk*. Earthscan.

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

Spinnewijn, J. (2015). Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs. *Journal of the European Economic Association*, 13(1), 130–167.

Strotz, R. (1956). Myopia and Inconsistency in Dynamic Utility Maximization. *Review of Economic Studies*, (23), 165–180.

Svenson, O. (1981). Are We All Less Risky and More Skillful than Our Fellow Drivers? *Acta Psychologica*, 47, 143–148.

Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207–232.

Tversky, A. & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–31.

Van der Klaauw, W. (1996). Female Marital Labour Status Supply Model and Decisions: A Life-Cycle Model. *Review of Economic Studies*, 63(2), 199–235.

Vaskovics, L. A. & Rost, H. (1999). Väter und Erziehungsurlaub. In *Schriftenreihe des Bundesministeriums für Familie, Senioren, Frauen und Jugend*, Number 179. Stuttgart: Kohlhammer.

Wagner, G. G., Frick, J. R., & Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements. *Schmollers Jahrbuch*, 127(1), 139–169.

Weinstein, N. D. (1980). Unrealistic Optimism About Future Life Events. *Journal of Personality and Social Psychology*, 39(5), 806–820.

Wrohlich, K. (2011). Labor Supply and Child Care Choices in a Rationed Child Care Market. *DIW Berlin Discussion Paper No. 1169.*

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

Declaration

Erklärung gem. §4 Abs. 2 der Promotionsordnung

Hiermit erkläre ich, dass ich mich noch keinem Promotionsverfahren unterzogen oder um Zulassung zu einem solchen beworben habe, und die Dissertation in der gleichen oder einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule nicht bereits zur Überprüfung vorgelegen hat.

Berlin, November 2018

Ulrich Schneider

Erklärung gem. §10 Abs. 3 der Promotionsordnung

Hiermit erkläre ich, dass ich für die Dissertation folgende Hilfsmittel und Hilfen verwendet habe: Stata, MATLAB, R und C++.

Auf dieser Grundlage habe ich die Arbeit selbstständig verfasst.

Berlin, November 2018

Ulrich Schneider