Essays in Public Finance and Family Economics

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt von

Tim Obermeier

im Frühjahrssemester 2019

Abteilungssprecher:	Prof. Dr. Jochen Streb
Referent:	Prof. Dr. Eckhard Janeba
Korreferent:	Prof. Dr. Michèle Tertilt

Tag der Verteidigung: 1. August 2019

Meiner Familie

Acknowledgements

Most importantly, I want to thank my advisors Eckhard Janeba and Michèle Tertilt for their exceptional guidance and support throughout my PhD studies and many insightful discussions about each of my research projects.

I am further very grateful to the co-authors of the thesis chapters. It was a great experience to start into the research phase of the PhD together with Mario Meier and jointly start working on a project that would – after countless discussions - eventually become our first paper. For the third chapter, I enjoyed working with Hans-Martin von Gaudecker and Holger Stichnoth. Both of them have also in general given me important support and advice throughout the years.

Special thanks also go to Andreas Peichl for his support. In addition, I am grateful to Richard Blundell for enabling a very enjoyable and productive research stay at IFS towards the end of the PhD.

My projects have benefited a lot from the research environment at the University of Mannheim. The weekly group meetings with Michèle have been a great opportunity to discuss research and have constantly provided feedback. I want to thank everyone who was involved – Michèle, Hanno, Fabian, Frederick, Xiaodi and Xue. The Macro and Public Finance groups have also provided important feedback in many presentations over the years.

Furthermore, I am grateful to Sonja Collet, Gabriele Zorell and Angelika Abuja for administrative support. I also appreciated the support of the staff at the Institute for Employment Research (IAB) and benefited from financial support through CRC TR 224 and from computational resources provided by the state of Baden-Wuerttemberg through bwHPC.

Finally, I want to thank Mario, Tobias, Hanno, Carla, Xin, Laura, Majed, Lukas and Vahe, who made the time of writing the thesis much more enjoyable, Wouter for his companionship and encouragement and for always being there for me and, last but not least, my family - in particular, Manfred, Renate and Britta.

Contents

List of Figures

List of Tables

General Introduction

This dissertation consists of three chapter, each of which can be read independently. Each chapter analyzes a topic broadly related to labor markets or the role of the government in the labor market: the first chapter is concerned with the optimal design of unemployment insurance programs, the second chapter with the effects of reforms to the progressivity of the income tax and the third with women's occupational choice over the life-cycle.

How governments should design the tax and transfer system and social insurance programs is a complicated question. Many issues are controversial, both in the academic and popular debate. How high should unemployment benefits be and for how long should they be paid? Should government increase income tax progressivity, given high levels of inequality? The first two chapters of the thesis add to these debates by studying specific aspects of these policies. Chapter 1 studies how unemployment insurance affects the hiring prospects of the long-term unemployed, by influencing firm decisions about which job candidates to invite for interviews and hire. The second chapter considers how income tax progressivity affects the *within-family* allocation of consumption and leisure, in addition to the well-studied effects on the *across-family* distribution, as well as marriage and divorce. In both chapters, the analysis is based on quantitative models, which allows to study large-scale hypothetical reforms and to assess their welfare implications.

The third chapter deals with fertility and women's careers. Having children has a substantial impact on women's labor supply and earnings, which has recently received renewed attention in the literature. We add to this literature by studying the *tasks* that women perform in their occupation. Tasks have been shown to be intricately linked to earnings and human capital. How do the tasks women perform evolve over the life-cycle and across cohorts? Do long career breaks lead to occupational 'downgrading'? Do women with long career breaks eventually perform different tasks than before childbirth? We empirically analyze these questions using large German datasets on labor market histories and fertility as well as on tasks that different occupations require.

In the following, I will present each of the three chapters in more detail.

Chapter 1: Employer Screening and Optimal Unemployment Insurance

This chapter (which is co-authored with Mario Meier) studies how firms' screening behavior affects the optimal design of unemployment policies. In our model, firms receive multiple applications and make selective interview decisions among applicants, combining the information from a noisy signal about productivity and unemployment duration. UI benefits change the search effort of workers, the pool of applications that firms receive and how informative duration is relative to the signal. We estimate the model using German administrative employment records and information on job search behavior, vacancies and applications. The model matches the observed decline in search effort, job finding rates and interview rates with increased unemployment duration. We find that allowing for employer screening is quantitatively important for the optimal design of unemployment insurance. Benefits should be paid for a longer period of time and be more generous in the beginning, but more restrictive afterwards, compared to the case where we treat the hiring and interview decisions of firms as exogenous.

Chapter 2: The Marriage Market, Inequality and the Progressivity of the Income Tax

This chapter studies how the progressivity of the income tax affects intra-household inequality and the marriage market. Tax progressivity increases the after-tax earnings of the lower-earning spouse and improves their bargaining position in marriage. This mechanism reduces inequality in consumption and leisure *within* households. In addition, tax progressivity can change who is single and who marries whom. I study these effects in an equilibrium search and matching model with intra-household bargaining, labor supply and savings. The model is calibrated to data from the Netherlands and used to study a hypothetical reform which increases progressivity by 40% relative to the current system. The reduction of intra-household inequality accounts for 24.77% of the reduction in inequality in private consumption due to the reform, and 11.43% of the reduction in inequality in utility from private and public consumption, leisure and home production. Changes in the composition of couples and singles, due to endogenous marriage and divorce, have small implications for inequality.

Chapter 3: Fertility and Women's Occupational Choice: Germany, 1975 - 2007

This chapter (which is co-authored with Hans-Martin von Gaudecker and Holger Stichnoth) studies the relationship between fertility and women's occupational choice. We combine German administrative data on employment, earnings and fertility with detailed survey data on the tasks that workers typically perform in an occupation. We use the data to study life-cycle profiles of employment and the intensity of analytical, interactive, routine and manual tasks. We analyze the dynamics of task intensities around childbirth in an event-study framework. Our findings suggest that mothers, on average, increase their analytical task intensities less than predicted in the absence of children. A simple back-of-the-envelope calculation suggests that this could generate a 7% reduction in the hourly wage. Compared to childless women, mothers switch their occupation more often, particularly those with a long career break, switch more towards family-friendly occupations, and eventually have a greater distance between their current and their pre-birth tasks.

Chapter 1

Employer Screening and Optimal Unemployment Insurance

1.1 Introduction

Most governments provide substantial levels of insurance against unemployment. Commonly, unemployment insurance systems pay benefits for a finite period of time and individuals move to more restrictive assistance schemes after benefits have expired. How high benefits should be, and for how long they should be paid, is controversial. While benefits typically expire after six months in the US, they are often paid for years in European countries. At the same time, several European countries have reformed their unemployment insurance system and substantially lowered the benefits for the long-term unemployed.[1]

Recent evidence from field experiments shows that employers take unemployment duration into account when deciding which applicants to interview for vacancies. Kroft et al. (2013) document that the probability of being invited for an interview falls by almost 50% within the first six months of unemployment in the US and find that this result can best be explained as screening behavior, which refers to the notion that firms infer lower productivity of a worker from a long unemployment spell. An aspect that has received little attention in the literature on UI design is that such statistical discrimination against the long-term unemployed could be endogenous to the UI system. Since UI policy affects the search effort of workers, the beliefs of firms and the pool of applicants they can choose from, and therefore their interviewing decisions, could adjust as well in equilibrium. The goal of this chapter is to study the impact of screening on the optimal schedule and in particular the role of the

[1]During the labor market reforms between 2000 and 2005, Germany reduced the benefit level for the long-term unemployed from 50-60% of the pre-unemployment wage to a fixed payment, which is 404 euros for singles in 2016, not including additional rent support. In Sweden, the unemployed get 80% of their pre-unemployment wage forever, but the payment is capped. In 2001, the government introduced duration-dependent caps, with a lower cap for the long-term unemployed (see Kolsrud et al. (2018) for details). In 2010, Denmark reduced the potential benefit duration from 4 to 2 years (afterwards, individuals may still receive welfare benefits).

equilibrium effects. We build a quantitative model of the job search and recruitment process and use the model to analyze optimal UI schedules.

The key feature of our model is that firms receive multiple applications from workers and only observe unemployment duration and a noisy signal about productivity. Firms rank workers by their expected productivity and workers with a long unemployment spell are less likely to be considered for interviews. Workers decide on their search effort and savings. Hiring and interview decisions are endogenous. They depend on how many applications a vacancy receives, the unemployment duration and the noisy signal of each application, and firms' beliefs about the productivity of the long-term unemployed. UI benefits influence how many applications firms receive on average as well as the composition of these applications, since e.g. low-productivity types react more strongly to benefit increases for the long-term unemployed. In addition, generous UI benefits reduce the information contained in unemployment duration, since they make it more likely that high-productivity types stay unemployed longer. Thus, in equilibrium the interview decisions of firms adjust as well in equilibrium when the search behavior of workers changes.

We estimate the model using German administrative data on job finding rates and survey data on search effort, vacancies, applications and savings. In particular, we use a comprehensive survey of establishments (the German Job Vacancy Survey) which contains information about the recruitment process. Vacancies on average receive 15 applications. When there is just one applicant for a vacancy, the probability that the applicant is interviewed is close to one. However, this probability drops to about 55% when there are 5 applicants, which is the median number of applications, and to 35% at the mean number of applications of 15. The Job Vacancy Survey also provides direct survey evidence that firms take workers' unemployment duration into account. About 45% of the establishments that consider unemployment applicants state that they are not willing to consider individuals with durations higher than 12 months. Our estimated model can match the empirical features of the job search and hiring process, namely the decline in job finding rates, the applications-per-vacancy ratio, the decline in interview rates and the decline in the job search effort of agents. We then use the estimated model to analyze the optimal unemployment insurance system and investigate the role of the equilibrium effects.

Our policy analysis is concerned with three features of an unemployment insurance system: the initial benefit level (first level), the length for which individuals are allowed to receive this level (potential benefit duration), and a second level for the long-term unemployed (second level). Benefit levels are always replacement rates in terms of the past wage. We find that the optimal schedule pays 73% for 42 months and drops close to zero afterwards. If we restrict the model to allow only for one application per vacancy, which shuts down the information friction, the optimal schedule pays 63% for 20 months and 27% afterwards. Thus, our first main result is that introducing employer screening matters substantially for optimal policy, relative to the case without screening.

We then use the model to assess how important the equilibrium channels of changing unemployment insurance benefits are relative to partial equilibrium effects. The equilibrium effects refer to changes in the probability of being hired conditional on applying to a firm. Our model features three channels through which unemployment policy can affect hiring probabilities. First, the information contained in unemployment duration depends on how different the shares of low and high types at that duration are. When changes to the unemployment insurance system increase the relative share of applications at high durations that come from high types, firms will take this into account and interview individuals with high durations more often. Second, unemployment insurance policy affects the overall applications-to-vacancy ratio. When there are more applications per vacancy, the long-term unemployment have worse job prospects because it becomes more likely that the firm has at least one applicant with a higher expected productivity. Third, unemployment insurance policy affects the composition of the pool of applicants, holding the overall ratio of applications per vacancy and firms' beliefs about productivity constant. For example, if policy reduces the search effort of individuals with low durations, this will increase the job prospects of individuals with high durations. In addition to these equilibrium adjustments, the partial equilibrium trade-off is between providing insurance and distorting the search effort of workers. Introducing employer screening, relative to a case with full information, interacts with this trade-off even in the absence of equilibrium effects. Moral hazard is represented by the responsiveness of workers to benefits and as workers anticipate their lower job chances in the future due to screening, or actually experience them after becoming long-term unemployed, their responsiveness to benefits changes.

To isolate the role of equilibrium effects, we analyze the case where hiring probabilities decline with duration as under the current German benefit schedule, but are assumed to be invariant to policy. This corresponds to the partial equilibrium effects of employer screening, where falling hiring probabilities change workers' search incentives, but these hiring probabilities itself are treated as exogenous. Calculating the optimal schedule yields 64% for 26 months and 21% afterwards. Also allowing hiring rates to adjust, which was our previous experiment, leads to 73% for 42 months and almost 0 afterwards. Under the current schedule, the hiring probability declines from 0.3 to 0.15 after 12 months. Under the optimal schedule, this decline is more gradual and hiring rates decline to about 0.22 after 12 months. Our second main result is therefore that the equilibrium effects - the adjustment of hiring rates - turn out to be fairly important, especially for the length of the first step and the level of the second step.

In addition, our results show that even when allowing for employer screening, the second benefit level for the long-term unemployed is relatively low. In general, with duration dependence - which refers to declining job-prospects over the spell -, it is theoretically open if benefits for the long-term unemployed are higher or lower than for the short-term unemployed, primarily because duration dependence decreases the moral hazard cost of providing benefits for the long-term unemployed. This is due to the fact that as the overall job finding rates of the long-term unemployed decrease, they become less responsive to benefits. Therefore, it could be the case that introducing employer screening, relative to the case without screening, makes it optimal to provide high levels of insurance for the long-term unemployed. Quantitatively, in the case of fixed hiring rates, we find that this effect mainly increases the length for which workers can receive the first level, but has a smaller effect on the levels. Taking the adjustments of hiring rates into account, the optimal level for the long-term unemployed is even lower than in the case without screening. These results suggest that while employer screening increases the length for which benefits should be paid, it does not necessarily provide a reason for giving high benefits to job-seekers with very long durations.

Related Literature. This chapter contributes to the literature on optimal unemployment insurance by providing a model of the hiring process that can be used to quantify the impact of employers' screening behavior on optimal benefit schedules. Many papers in the literature focus on partial equilibrium models and distortions in search effort, where unemployment insurance is a trade-off between moral hazard and consumption smoothing, e.g. Baily (1978), Gruber (1997), Chetty (2006), Chetty (2008). The optimal schedule is often argued to be declining with duration or flat, as in Hopenhayn and Nicolini (1997) and Shimer and Werning (2008), respectively. Related to our approach, Lentz (2009) estimates a search model with savings to analyze optimal unemployment insurance levels. In Schmieder and Von Wachter (2016) the authors extend the standard optimal unemployment insurance setting to a case where not only benefit levels but also the benefit duration is optimally chosen by the planner. The policies we look at are comparable to their setting. Most related to this chapter, Lehr (2017) and Kolsrud et al. (2018) theoretically show that allowing for firms' screening behavior changes the optimality conditions for benefit levels by introducing an externality term, so that the standard Baily-Chetty formula does not hold. The contribution of this chapter is that we build a quantitative model that can match the relevant empirical features of the recruitment process and use the model to assess the role of the equilibrium effects relating to employer screening. Our results suggest that these equilibrium effects are quantitatively important and should be taken into account when designing unemployment policies.

There is relatively little other work on the implications of duration dependence for optimal policy. Shimer and Werning (2006) investigate optimal unemployment insurance in a setting with exogenously falling wages or job arrival rates. Pavoni (2009) focuses on human capital depreciation. These papers analyze duration dependence in models where duration dependence is exogenous and invariant to unemployment insurance policy while screening, on the other hand, is endogenous to the benefit system. As a result, screening has different policy implications than other forms of duration dependence since we find that the equilibrium adjustments of the hiring rates are quite important.

In addition, this chapter is related to the literature on duration dependence and recruitment behavior. Lockwood (1991) was an early paper in this literature. In his setting, firms test the unemployed before hiring and a high unemployment duration can be a bad signal. The idea of ranking applicants by unemployment duration was first explored by Blanchard and Diamond (1994), who assume that firms with multiple applications always hire the applicant with the shortest unemployment duration. Recently, the results from the audit studies have led to a growing amount of work on the role of employer screening. Jarosch and Pilossoph (2018) investigate the quantitative link between the decline in callback rates and duration dependence and emphasize that statistical discrimination may not always lead to lower job-finding rates. Doppelt (2016) models the role of information contained in the history of unemployment spells, thereby stressing the life-cycle dimension. Fernandez-Blanco and Preugschat (2018) consider a directed search model with endogenous wages, in which firms rank applicants by unemployment duration. There are two important features of our model relative to these papers. First, in our setting, firms rank multiple applicants according to unemployment duration and a signal, whereas previous models of ranking assume that firms only use duration. As a result, unemployment insurance changes how informative duration is relative to the signal. When policy makes the selection of types by duration weaker, firms rank applicants less by unemployment duration and more by the signal. Second, we integrate search effort and savings, which are crucial for the analysis of optimal unemployment insurance.

There have been recent studies that emphasize the role of equilibrium effects and market externalities, e.g. Michaillat (2012), Landais et al. (2018) and Landais et al. (2016), Marinescu (2017) and Lalive et al. (2015). These papers argue that search externalities among job seekers might be important for job outcomes which in turn has implications for the design of unemployment insurance benefits. Our concept of multiple applications generates search externalities among job seekers and the higher the applications-per-vacancy ratio the more important are search externalities. Hagedorn et al. (2019) argue that unemployment benefit extensions can have externalities on labor demand and decrease the incentive to create vacancies. Our model also allows for vacancy creation to close the model and to account for this effect.

The rest of the paper is organized as follows. In Section 1.2, we focus on the data and some descriptive facts. Sections 1.3 presents the model and policy problem. Section 1.4 describes the estimation and discusses estimation results and model fit. In Section 1.5, we discuss welfare and the corresponding policy results. In Section 1.6, we discuss some extensions of our model and conclude in Section 1.7.

1.2 Data & Descriptive Facts

This section presents the data we use and empirical facts about job search behavior and the hiring process.

1.2.1 Data

In this chapter we consider the case of Germany. In Germany most unemployed receive unemployment benefits for up to 12 months of unemployment and are eligible for unemployment assistance if they stay unemployed for longer than 12 months. Older individuals are eligible for longer unemployment insurance payments, but we restrict to individuals that receive 12 months of benefits. Unemployed individuals receive benefits that amount to 60% or 67% of their past wage, depending on their marital status. After individuals run out of unemployment insurance (UI) they receive means-tested unemployment assistance benefits (UA) which are on average around 40% of the past wage for the average unemployed. Unemployment benefits are financed by social security contributions of workers and firms.[2]
The German setting allows us to base the design and estimation of our model on several datasets that contain information on job-finding rates, search effort and vacancies. First, we use the German social insurance data (IEB) which provides us with information on the characteristics of the unemployed; in particular the length of their unemployment spell and their wage history. The data contains all individuals that were ever unemployed or regularly employed through an employment relationship that is subject to social insurance. We have access to a 2% random sample of the population and restrict ourselves to unemployment spells starting in the years from 2000 until 2011. Second, we use the IZA Evaluation dataset (IZA ED) which is a representative survey performed among UI entrants between June 2007 and May 2008. The data is a panel where participants were interviewed up to four times after their unemployment spell has realized. The first interview took place close to the beginning of unemployment. Additional interviews took place six, twelve and thirty-six months after the start of the UI spell, respectively. Participants are asked about their individual search effort, e.g. the number of applications or number of search channels, and they are asked to report their reservation wage. Third, we use the IAB Job Vacancy Survey (JVS) which is a representative survey conducted among firms on open vacancies and hiring decisions made by firms. The survey contains information on whether unemployed applicants were hired and how many applicants firms invite to an interview. Fourth, we use the Bundesbank Panel on Household Finances (PHF), which contains information on savings, liquid assets and

[2]The German unemployment insurance system compares relatively well to unemployment insurance schemes in other developed countries, like the US or many other European countries. However, the US system has somewhat less generous potential benefit durations and replacement rates than Germany and no unemployment assistance system. For further details on the institutions in Germany we refer the reader to the appendix.

debt levels. In the data individuals are also asked to report whether they are unemployed or employed.

Table 1.1 summarizes some of the main characteristics of the data sources. The average monthly re-employment wage after unemployment for job seekers is 1,606 euros. The re-employment wage is defined as the average monthly earnings an individual receives in the year after the UI spell has ended. Table 1.1 also reports some observable characteristics of unemployed job seekers. In the IZA ED data, individuals use roughly four to five search channels, where most people in the sample look for job advertisements, ask friends or relatives for jobs or use online search. Many individuals are also offered help from the local employment agencies. Table 1.1 shows that agents send out 13 applications on average at the beginning of the UI spell. From the PHF dataset we extract some information regarding assets, in particular liquid assets, of the unemployed. In Table 1.1 we show different quantiles from the net liquid asset distribution of the unemployed in the sample. We see that asset holdings are indeed very heterogeneous where nearly half of the individuals barely have any assets.[3] In contrary, 10% of individuals have more than 40,000 euros in liquid assets. Net assets, which also include real estates, are on average larger. Finally, the JVS shows that firms receive on average 15 applications and that it takes around two months to fill an open vacancy.[4]

1.2.2 Descriptive Facts

Standard job search models assume that job finding rates are only determined by agents' search effort, potentially with declining job prospects in the form of duration dependence or heterogeneity in job finding rates.[5] However, whether agents find jobs to exit unemployment also requires a firm to actually hire the job seeker. This drives a wedge between the search effort of an agent and the job finding rate of an agent. In addition, firms' hiring probabilities are potentially dependent on the policy context. Hence, in the following we provide some evidence on job seekers search effort as well as on firms screening and interview decisions. Based on this evidence we build a job search model that incorporates all of the discussed features and makes distinct predictions along the evidence that we provide.

Job finding rates. The job finding rate of unemployed job seekers in Germany is shown in Figure 1.1 panel (a). In the first months of unemployment, exit rates out of unemployment are above 10%. However, job finding rates decrease throughout the spell and are only 5%

[3] Net liquid assets are defined as the difference between liquid assets and short-term debt, like credit card debt.

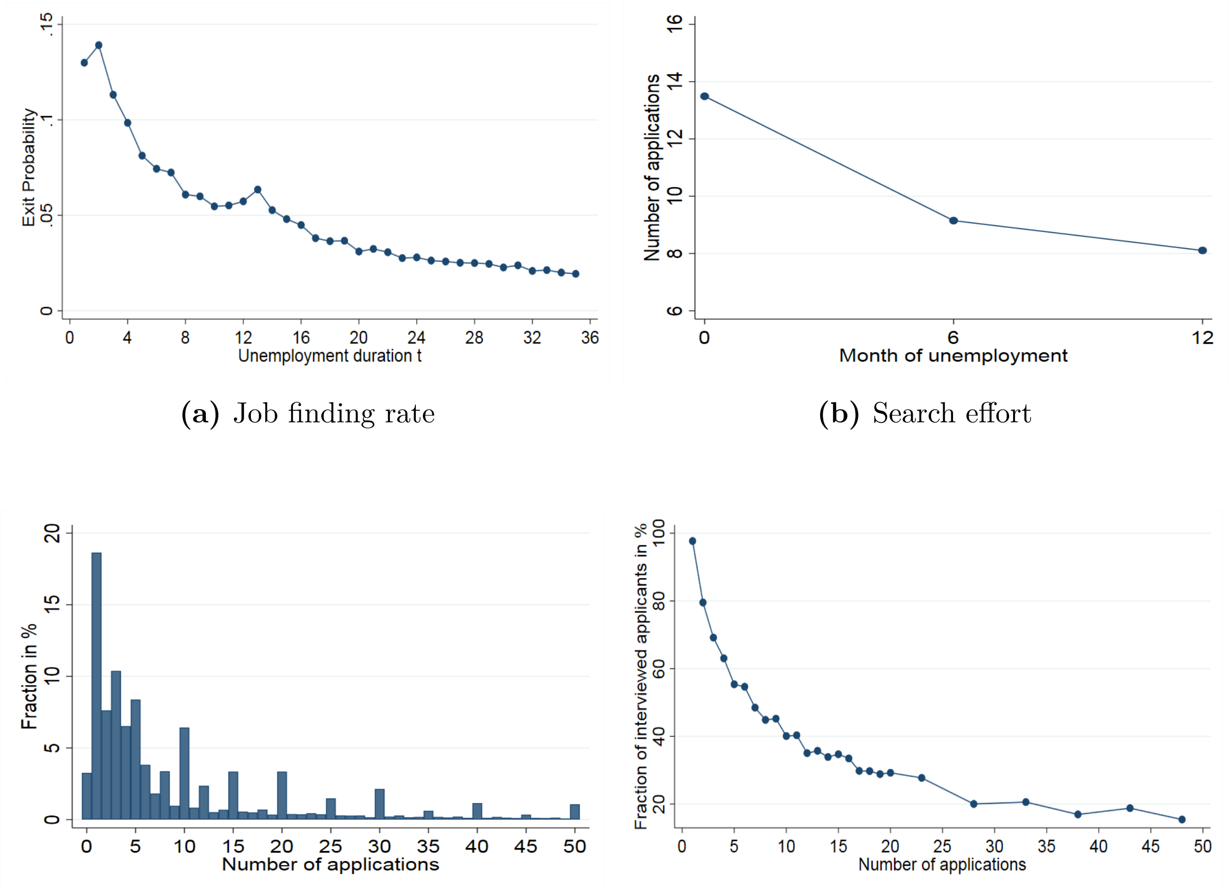
[4] This time is defined as the difference between the acceptance of a job offer by an applicant to the release of the job advertisement.

[5] See e.g. Chetty (2008), Lentz (2009), Hopenhayn and Nicolini (1997).

Table 1.1: Descriptive statistics

Variables	N	mean	s.d.
Panel A: Employment Register			
Re-employment wage (euros)	55,420	1,606.17	(1,059.95)
Unemployment duration (months)	59,793	12.57	(12.71)
Female	59,793	0.446	(0.497)
Age	59,793	30.80	(9.12)
Married	59,793	0.325	(0.468)
Children	59,793	0.302	(0.459)
College	56,727	0.096	(0.294)
Apprenticeship	56,727	0.751	(0.432)
Panel B: IZA Evaluation Dataset			
Number of applications Month 1	6,815	13.49	(14.95)
Number of applications Month 6	377	9.15	(10.09)
Number of applications Month 12	1,710	8.11	(9.78)
Search channels Month 1	6,898	4.78	(1.78)
Panel C: Panel on Household Finances (Quantiles)			
Net liquid assets (euros, p10)	295	-1,003	-
Net liquid assets (euros, p25)	295	0	-
Net liquid assets (euros, p50)	295	247	-
Net liquid assets (euros, p75)	295	4,885	-
Net liquid assets (euros, p90)	295	40,497	-
Net assets (euros, including home, p50)	295	894	-
Panel D: Job Vacancy Survey			
Number of applicants	62,904	14,79	(36.96)
Time vacancy is open (days)	76,240	56.88	(67.08)

Notes: This table shows descriptive statistics from our different data sources. Panel A shows descriptive statistics from the administrative employment registers of individuals who experience their first unemployment spell at the time the spell starts. Panel B summarizes search effort measures from the IZA evaluation dataset. Panel C uses the Bundesbank Panel on Household Finances for information on assets. In Panel D statistics on vacancies are shown, coming from the IAB Job Vacancy Survey. N denotes the number of observations behind each statistic, and s.d. the standard deviation.



(a) Job finding rate

(b) Search effort

(c) Distribution of applications

(d) Share of interviewed applicants

Figure 1.1: Descriptive Facts

Notes: Panel (a): This figure shows the job finding probability (hazard rate) of individuals on the y-axis as a function of the unemployment duration on the x-axis. Source: SIAB. Panel (b): This panel shows the mean number of applications unemployed agents send out in the first month of unemployment, the sixth month of unemployment and after one year of unemployment. Source: IZA ED. Panel (c): This figure illustrates the distribution of applications across vacancies. The y-axis denotes the fraction of vacancies that receive a certain number of applications. Source: JVS. Panel (d): This panel shows the fraction of interviewed applicants as a function of the number of applications received. Source: JVS.

after one year and 2.5% after two years of unemployment.[6] Hence, the chance to find a job becomes smaller and smaller the longer someone is unemployed. There are two explanations for this decline in the hazard rate out of unemployment: (a) selection/heterogeneity, or (b) (true) duration dependence. Heterogeneity can enter in the form of productivity differences of job seekers. Duration dependence describes declining job prospects for individuals given their type. Most likely, both, selection and duration dependence, contribute to falling hazard rates.

Search effort. Since we are interested in dynamic UI policies it is important how individuals' search effort throughout their unemployment spell reacts, because search effort responses are a main determinant of the moral hazard costs associated with unemployment insurance. Figure 1.1 panel (b) illustrates the number of applications that agents write per month as a function of their unemployment duration. At the beginning of the spell they send out more than 13 applications per month, after six months around nine applications are sent out and after twelve months only eight applications are sent out on average. Hence, the average search effort seems to decrease over the spell.[7] The graphs look very similar when restricting the sample to individuals who are unemployed for 12 months and tracking their search effort over time (see the appendix for details). Note, we have ignored other measures of search effort for now, e.g. the number of search channels or time used for job search. Our choice is motivated by the fact that our model explicitly allows agents to send out applications.[8]

Multiple applications per vacancy. A very important factor that determines job search outcomes is how many other applicants are searching for a similar job. Hence, depending on the number of applications per vacancy the job finding rate might be higher or lower for a given search effort. The importance of these *crowding out effects* depend on the number of competitors of an applicant for a job. Intuitively, if there are many applicants per vacancy some job searchers will get no offer for the job and need to continue their search. Figure 1.1 panel (c) plots the histogram of the number of applications an open vacancy receives. The average number of applications is around 15, with a median of 5 applications per vacancy. This panel suggests that firms have considerable levy to pick the best applicant and that the *outside option* of a firm is to screen or hire alternative applicants.

Employer screening. Employer screening by vacancies takes usually place by restricting first to a subset of applicants that get invited to an interview. In panel (d) of Figure 1.1 we show that the share of applicants that receive an interview invitation depend on the number of applications a vacancy receives. One can see that the more applications there

[6]The small spike at 12 months is due to the benefit exhaustion which leads more people to exit unemployment. See DellaVigna et al. (2017) for a detailed exploration of the benefit exhaustion spike.

[7]Declining search effort over the UI spell was also documented for the US by Krueger and Mueller (2011).

[8]Lichter (2016) also uses the number of applications as a search measure and discusses this choice in more detail.

are, the less likely it is to get invited to an interview. The interview shares are around 50% for vacancies with 5 applications, i.e. at the median, and only 30% for vacancies with 15 applications, i.e. at the mean. In the job vacancy survey, employers are also asked whether they consider unemployed applicants depending on the unemployment duration of the applicant. Conditional on considering unemployed applicants at all only 75% of firms consider applicants with more than a few months of unemployment duration and only 60% of firms consider applicants with more than twelve months of unemployment duration. Hence, only 60% of firms that are in principle willing to consider unemployed applicants are willing to accept long-term unemployed. Figure A.2 in the appendix illustrates this graphically. Complementary to our survey evidence, the importance of employer screening for true duration dependence was also studied by Kroft et al. (2013) in an experimental audit study. They find that the callback rate (interview invitation) of an application that was sent out to open vacancies strongly depends on the unemployment duration presented in the CV of the applicant. In fact, the probability to receive a callback from an employer declines by roughly 50% over the unemployment spell. Note that declining callback rates can in principle also be generated by models of human capital depreciation. However, Kroft et al. (2013) demonstrate that the decline of the callback rate is much weaker when the unemployment-to-vacancy ratio is high. This finding is hard to rationalize with human capital depreciation, since human capital would depreciate independently of labor market conditions. Employer screening, on the other hand, predicts that unemployment duration is less informative about productivity under adverse labor market conditions, since then individuals with high productivity also stay unemployed longer. This is in line with the evidence provided by Kroft et al. (2013).[9]

1.3 Model

We extend a standard search model with risk aversion, endogenous search effort and savings, that has been used to study optimal UI, by incorporating firms' hiring decision to account for the empirical patterns described in the previous section. The key feature of our model is that workers are heterogeneous in productivity and firms have to select candidates from a pool of multiple applications. Since productivity is only observed by workers, firms base hiring decisions on the expected productivity of each worker, taking unemployment duration and a noisy signal about worker quality into account.

[9] In addition, note that they find that the callback rate declines strongly within the first six months of unemployment and is essentially flat afterwards. If the decline in callback rates would mostly be about human capital depreciation, one would expected a more gradual decline that also affects the long-term unemployed.

1.3.1 Workers

Time is discrete and each period corresponds to a month. We follow the literature on optimal unemployment insurance by assuming that workers are born unemployed (Chetty (2006), Shimer and Werning (2008)) and that there is no job destruction, so that finding a job is an absorbing state. Workers live for T periods and in every period of the model, a unit mass of newly unemployed workers is born. Workers who have been unemployed for t periods get UI benefits that depend on t:

$$b_t = \begin{cases} b_1 & \text{if t } \leq D \\ b_2 & \text{if t } > D \end{cases}$$

Thus, workers can get an initial level b_1 for up to D months and a level of b_2 afterwards.[10] Workers differ in their productivity π_j and each generation of workers contains a share α_j of type $j = 1, ..., J$. In addition, each type has an exogenous initial level of assets, denoted as $k_{0,j}$.

Employed workers only decide on the optimal level of consumption and savings and the corresponding value function and budget constraint for duration $t < T$ are:

$$\begin{aligned} V^e(k,t) &= \max_{k_{t+1} \geq 0} \left\{ u(c_t) + \beta V^e(k_{t+1}, t+1) \right\} \\ c_t &= Rk_t + (1-\tau)w - k_{t+1} \end{aligned}$$

k_t and k_{t+1} are the asset levels in each period. Workers are risk-averse and discount the future at rate β and the interest rate is given by R. There are no separations and employment is an absorbing state.[11] In addition, note that all workers face a no-borrowing constraint ($k_{t+1} \geq 0$).[12]

Unemployed workers decide on both consumption and savings and their search intensity. Searching with intensity s has a cost $\psi(s)$, but leads to a match probability $p(s) = s$, which can be interpreted as sending an application to a firm.[13] Importantly, the probability of exiting unemployment - the hazard rate - contains both the probability of meeting a firm and of actually being hired by the firm:

[10]Note that in practice, the amount of unemployment benefits is often tied to the pre-unemployment wage. Because our model abstracts from wage heterogeneity the pre-unemployment wage is conceptually indistinguishable from the post-unemployment wage.

[11]Allowing for separations is in principle possible but would complicate the model by generating an endogenous initial asset distribution. Hence, for simplicity we assume that jobs last forever.

[12]The no-borrowing assumption is standard in the literature, see e.g. Chetty (2006), and creates an insurance motive for the government in the first place. Without borrowing constraints, individuals would just take a loan and there would be no need for the government to provide insurance to the unemployed.

[13]For simplicity, we focus on the case where workers may send out a single application, as is also done in Fernandez-Blanco and Preugschat (2018) or Villena-Roldan (2012). The implications of multiple applications per worker are discussed in Section 1.6.

$$h_{j,t} = s_{j,t} \cdot g_j(t) \tag{1.1}$$

$g_j(t)$ is the expected hiring probability and is determined in equilibrium, as will be discussed in the next sections.[14] Jobs start in the next period. The survival rate in unemployment, i.e. the probability of still being unemployed after t periods, is then defined as

$$S_{j,t} = \prod_{t'=0}^{t-1}(1 - h_{j,t'})$$

Taken together, the value function for unemployed workers is given by:

$$\begin{aligned} V^u(k,t) = \max_{s,k_{t+1}\geq 0} \Big\{ u(c_t) - \psi(s) \quad &+ \quad \beta h_{j,t}(s) V^e(k_{t+1}, t+1) \\ &+ \quad \beta(1 - h_{j,t}(s)) V^u(k_{t+1}, t+1) \Big\} \end{aligned}$$

The budget constraint is $c_t = Rk_t + b_t - k_{t+1}$. Note that changes to the benefit system influence the value of unemployment relative to employment and therefore affect workers' search decisions.

In each period of the model, there is a pool of unemployed workers that consists of the new generation and workers from previous generations that did not find a job in previous periods. While further details will be discussed in the equilibrium section, it is useful to note that the number of workers of type j and duration t that are matched with firms in each period is given by:

$$a_{j,t} = \alpha_j \cdot S_{j,t} \cdot s_{j,t} \tag{1.2}$$

Here, α_j is the unconditional type share, $S_{j,t}$ is the survival rate until duration t and $s_{j,t}$ is the search effort at that duration. Aggregating over types and duration, this leads to a mass of matched workers that will be considered by firms, which we will refer to as the pool of applications.

1.3.2 Firms

When workers are matched with a firm, the match-specific productivity $q \in \{0, 1\}$ is drawn and the probability that it takes the value 1 is given by worker productivity π_j. Thus, high-productivity workers have a high chance of being productive in any match. We refer to the case of $q = 1$ as the worker being qualified for a vacancy.[15] Firms produce an output y

[14] Note that we use the term *hiring probability* for the probability of being hired conditional on being matched (as also e.g. Lehr (2017)), while similar terms are also often used in the literature to describe to number of new hires by firms over total employment.

[15] This is similar to the set-up of Fernandez-Blanco and Preugschat (2018), who also assume that workers differ in their probability of being qualified for vacancies. In a similar spirit, Jarosch and Pilossoph (2018) assume that both workers and firms differ in their (deterministic) productivity and production only takes place when worker productivity is higher than firm productivity.

when employing a qualified worker and zero otherwise. Thus, note that conditional on being qualified, workers produce the same output.[16]

Workers are matched to firms according to an urn-ball matching technology, where each matched worker randomly arrives at a firm. From the point of view of the firm, the number of applications it receives follows a Poisson distribution with parameter $\mu = \frac{a}{v}$, where a is the mass of matched workers and v is the mass of vacancies. For each candidate, firms do not observe if they are qualified, but only their unemployment duration and a noisy signal about the type of the worker. The signals sent by type j are drawn from a normal distribution, where we normalize the mean to j and estimate the variance σ to match the data. Thus, high types on average send better signals. Firms can interview applicants and thereby perfectly reveal their productivity. We restrict firms to pay the exogenous wage.[17] Firms rank applicants by their expected productivity and sequentially interview applicants until one applicant turns out to be qualified.[18]. The other applicants are not hired. Since the firm always has to pay the wage, it will never hire an unqualified worker. A key feature of this framework is that firms rank applicants not only based on unemployment duration, but also take the signal into account.[19] Note that ranking is justified as long as there is a positive screening cost.[20].

Thus, a firm first computes the expected type probabilities of each applicant. Firms know the composition of the overall pool of applications, i.e. the mass of applications $a_{j,t}$ sent by agents of type j and duration t. Firms also know the distributions of the signals. Conditional on the realized signal ϕ and unemployment duration t, the probability of an applicant being type j follows from Bayes' rule:

$$P(j|\phi, t) = \frac{f_j(\phi) \cdot a_{j,t}}{\sum_k f_k(\phi) \cdot a_{k,t}} \tag{1.3}$$

This probability corresponds to the share of applications of type j in the overall pool of

[16]Allowing the output to differ between low and high types would in principle be feasible in our framework and an interesting extension because it would allow to investigate the trade-off between providing information about the quality of applicants for firms and veiling information to protect unproductive types from statistical discrimination. In our setup, the planner would like to eliminate statistical discrimination because it reduces the job prospects of the long-term unemployed. In contrast, when productivity differs the planner also has an incentive to *provide* information to firms to maximize production. Note, however, that in the current framework, reducing screening can also have an adverse effect on firms if it is achieved by increasing the search effort of low types and increasing the effort of high types, which would reduce vacancy creation.

[17]The implications of endogenous wages are discussed in Section 1.6. Assuming a fixed wage is broadly in line with evidence about constant reservation wages over the spell and a moderate decline in re-employment wages by duration.

[18]An alternative approach that would give similar outcomes is to assume that firms choose which share of applicants they screen, while discarding the others. This second approach to recruitment selection is used e.g. in Villena-Roldan (2012) or Wolthoff (2017).

[19]In other ranking models in the literature (Blanchard and Diamond (1994), Fernandez-Blanco and Preugschat (2018)), the ranking is only based on duration.

[20]In the main part of the analysis, we focus on the case of a screening cost $C \to 0$

applications from agents with duration j, weighted by the density of the signal. Since the mass of applications is given by $a_{j,t} = \alpha_j S_{j,t} s_{j,t}$, a high duration of unemployment is a negative signal about productivity when a large share of applicants with duration t has a low productivity. Note that this does not only depend on the relative survival rates, but also on the relative search effort. For example, if there are many more low types than high types, but low types do not search. Firms will takes this into account and infer that the applicant must be a high type. Finally, note that in the limit case $\sigma \to 0$, the signal perfectly reveals workers' type and there is no reason to take the duration into account. Conversely, when $\sigma \to \infty$, the signal contains no information and firms only rank applicants based on duration. For intermediate cases with $\sigma \in (0, \infty)$, firms weigh the information contained in both components and their relative importance is endogenous. When the benefit system keeps productive types in the pool longer, duration can become less informative about productivity and the ranking order depends more strongly on the signal.

To arrive at the expected hiring rate, we first define the expected profit based on the conditional type probabilities:

$$\Pi(\phi, t) \quad = \quad \sum_j P(j|\phi, t) \; \pi_j y - w \tag{1.4}$$

It is useful to first focus on the case of an applicant i with fixed (ϕ, t, j), with j being the type, who is matched with a vacancy that has just one randomly drawn other applicant $\tilde{i}$ with characteristics $(\tilde{\phi}, \tilde{t}, \tilde{j})$. Applicant $\tilde{i}$ is interviewed before applicant i whenever $\Pi(\tilde{\phi}, \tilde{t}) \geq \Pi(\phi, t)$ and hired if also being qualified for the job, which happens with probability $\pi_{\tilde{j}}$. We define $p(\phi, t)$ as the probability that given ϕ and t, agent i is *not* interviewed, because the firm interviews and hires worker $\tilde{i}$ before, integrating over $(\tilde{\phi}, \tilde{t}, \tilde{j})$:

$$p(t, \phi) = \sum_{\tilde{j}=1}^{J} \frac{a_{\tilde{j}}}{a} \cdot \pi_{\tilde{j}} \cdot P\Big(\Pi(\tilde{\phi}, \tilde{t}) \geq \Pi(\phi, t) \;—\; \tilde{j}, t, \phi \;\Big) \tag{1.5}$$

$\frac{a_{\tilde{j}}}{a}$ is the probability of drawing type $\tilde{j}$ from the pool of all applications, with a being the total number of applications and $a_{\tilde{j}}$ the number of applications sent by type $\tilde{j}$. This is multiplied with the probability that type $\tilde{j}$ is hired according to the intuition described above.[21] The probability $p(\phi, t)$ describes how likely it is to not be invited for the interview when there is one other applicant. In general, the number of other applicants follows a Poisson distribution, where the mean μ is the mean number of applications per vacancy. In addition, the signal ϕ that the agent sends is stochastic. Integrating over both the number of other applicants and the signal, we get the following expression for the expected hiring rate:[22]

[21] In the appendix, we describe how the probability that the competitor sends a better signal is computed.

[22] This expression follows from the fact that the number of other applicants for a vacancy is Poisson distributed. The Poisson probability density function is $f(k) = \exp(-\mu)\frac{\mu^k}{k!}$. The probability that agent (j, t) with signal ϕ is the best applicant is $\sum_{a=0}^{\infty}(1 - p(t, \phi))^a f(a)$, since given a other applicants $(1 - p(\cdot))^a$ is the probability that none of them is hired first. This can be simplified to to the expression used for $g_j(t)$.

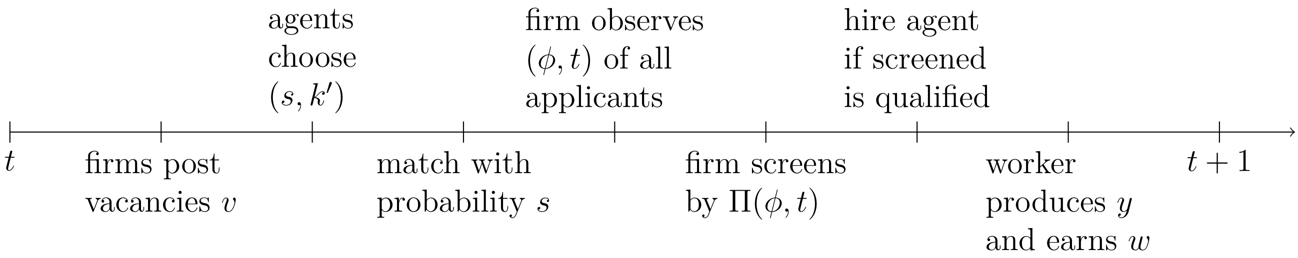


Figure 1.2: Timing of the model

$$g_j(t) = \pi_j \int_\phi \exp\Big(- p(\phi, t) \cdot \mu \Big) dF_j(\phi) \tag{1.6}$$

The expected hiring rate of worker i consist of the integral, which is the probability that no other applicant is screened and hired before, and the probability π_j that the worker is qualified for the job. The integral can be interpreted as a *callback curve*: it represents the probability of being contacted and screened by an employer. Thus, it is the model analogue to recent audit studies which measure the decline in the callback rate (e.g. Kroft et al. (2013)). Callback rates map into hiring rates by pre-multiplying the probability of being qualified for the vacancy. Note that there are two components that lead to a decline in the callback curve with duration. First, for a given agent with a high duration, $p(\phi, t)$ tends to be high, which means that the firm is likely to first interview and potentially hire one other randomly drawn applicant. This depends on how informative duration is about types and on the composition of the pool of applications - if the short-term unemployed search a lot, it is more likely that a random other applicant has a short duration and is potentially considered first. Second, this effect is scaled by the mean number of applications per vacancy, which is given by μ. In the limit case of no competition ($\mu = 0$), the hiring rate is flat and equal to π_j. In the case of a large applications-per-vacancy ratio μ the competition for jobs is large and callback rates are lower.

The mass of vacancies is pinned down by a free-entry condition. As in Lise and Robin (2017), firms can pay $c(v)$ to advertise v vacancies. Vacancies last for one period. The value of an additional vacancy is the net output multiplied by the probability of receiving at least one qualified application:[23]

$$J^v = \frac{y - w}{1 - \beta}\Big(1 - \exp\Big(- \frac{\sum \pi_j a_j}{v}\Big)\Big)$$

In equilibrium, the marginal vacancy costs are equal to the expected value of an additional

[23]Note that we assume that vacancies survive forever and that after the vacancy is filled it stays filled forever. This is a helpful approximation especially when T is large enough.

vacancy:[24]

$$c'(v) = J^v$$

Conceptually, free entry ensures that firms punish redistribution towards workers by exiting. Hence, vacancies might negatively or positively react to changes in unemployment policies. In our framework, different benefit schemes can reduce firm profits by either reducing overall search effort or by reducing the applications of high types relative to low types, because each case makes it less likely that vacancies receive at least one qualified candidate. As a result, firms would reduce the amount of vacancies being posted. Later, when we discuss optimal policy, these incentives for vacancies must be taken into account. In Figure 1.2 we summarize the timing of our model graphically.

1.3.3 Equilibrium

The equilibrium of the model consists of

- Policy functions for search effort s_{j,t,k_t} and savings $k_{t+1} = g_u(k_t, t, j)$ for the unemployed and $k_{t+1} = g_e(k_t)$ for the employed, for each type j, duration t
- Survival functions $S_{j,t}$
- Expected hiring rates $g_{j,t}$
- A mass of vacancies v

such that the policy functions of workers solve the problems described by the value functions for the employed and unemployed, and such that the expected hiring rates are optimal according to equation (1.6) given the implied survival rates.[25]

1.3.4 Optimal Policy

The governments' set of policy instruments $P = (b_1, b_2, D, \tau)$ consists of the benefits b_1 that are paid from period $t = 1$ until period $t = D$. D denotes the last month until benefits b_1 are received and represents the potential benefit duration. From period $t = D + 1$ until period T agents receive benefits b_2. This defines the policy schedule b_t, where $b_t = b_1$ if $t \leq D$ and $b_t = b_2$ if $t > D$. The proportional income tax τ is collected from the employed to finance the expenditures. The tax has also the interpretation of an actuarial fair insurance premium here. We restrict the analysis to this class of schedules because it facilitates numerical optimization

[24]Depending on the functional form of $c'(v)$ vacancy creation rents accrue to firms if vacancy costs are not constant. However, it is not obvious how to interpret these rents and we ignore them throughout the rest of the paper.

[25]While uniqueness of the equilibrium cannot be proved analytically, we checked for the possibility of multiple equilibria, especially around the estimated parameter values, and always converge to the same equilibrium.

over the policy space. In addition, these schedules are fairly close to the policy instruments that are used in practice.[26]

The objective of the planner is to maximize the value of a newly born generation of unemployed. We assume that every unemployed individual has the same welfare weight when born, which amounts to a standard utilitarian welfare criterion as in Chetty (2006):

$$W(P) = \int_j V_j^u(P)\alpha_j dj \tag{1.7}$$

However, the government can only maximize the welfare of agents subject to the following budget constraint, that balances expected revenue and expenditure from a cohort:

$$G(P) = \int_j \Big(\underbrace{\sum_{t=0}^{T} R^{-t}(1 - S_{j,t})w\tau}_{\text{expected revenue}} - \underbrace{\sum_{t=0}^{T} R^{-t} S_{j,t} b_t}_{\text{exp. expenditure}} \Big)\alpha_j dj \tag{1.8}$$

Note that revenues and expenditures are weighted by the survival rates, because individuals receive only benefits if they are still unemployed in period t and only pay taxes ($w\tau$) if they work in period t. The budget constraint implies that expected revenue generated with the employment tax must equal expected expenditures. As in Kolsrud et al. (2018) we assume that the budget must be balanced within a certain generation and therefore benefits and revenues are discounted by the interest rate.[27]

Discussion. In this framework, the screening mechanism matters for optimal policy through various channels. First, there is the classical trade-off between providing insurance to risk-averse individuals and distorting their search incentives (see e.g. Chetty (2006)). Insurance is valued because agents are credit constrained and cannot borrow. Hence, agents deplete their assets throughout the unemployment spell until they become hand-to-mouth consumers. Depending on the initial asset position, agents move closer to becoming hand-to-mouth if they stay unemployed for longer. The key measure of moral hazard is the elasticity of search effort with respect to UI benefits. Note that introducing screening changes the extent of moral hazard: forward-looking individuals will anticipate that they will have lower job prospects if they become long-term unemployed and search more intensively in the beginning, which can reduce their responsiveness to benefits.

Second, the presence of screening gives rise to equilibrium effects: the UI system changes not only search decisions, but also the expected probabilities of being hired. On the one hand, this is due to the fact that UI policy changes the selection of types over the unemployment

[26]See Section 1.5 for a discussion of the shape of more flexible classes of schedules.

[27]Alternatively, one could remove the discounting and collect taxes from the steady state distribution of employed and pay benefits to the steady state distribution of unemployed. We prefer our specification because then the tax τ has the interpretation of an actuarial fair insurance premium assuming that agents do not know their type ex-ante or that insurance pricing by type is not feasible.

spell. For example, consider the case of raising benefits at each duration. This will lead high types to stay in the unemployment pool longer and this makes being unemployed for a certain time less informative about productivity, as the relative survival rates change. This channel is also theoretically discussed in Kolsrud et al. (2018) and Lehr (2017). On the other hand, in our framework, the size and composition of the pool of applications that firms get matters for the determination of hiring rates. If policy changes search effort, this impacts the applications-per-vacancy ratio and a higher mean number of applications reduces the job chances of the long-term unemployed. In addition, if the short-term unemployed search a lot, this reduces hiring rates for the long-term unemployed. In a similar spirit, if low types search a lot, this decreases job chances of the high types who are unemployed for the while. Furthermore, vacancies adjust in equilibrium and optimal policy must take into account that different benefit schemes might lead to a different vacancy posting behavior because the value of a vacancy might be affected, through a change in the composition of applicants or their search effort. Finally, since agents are heterogeneous, a utilitarian planner potentially redistributes between them.

Combining these channels, the shape of the optimal schedule is theoretically open. Without duration dependence or heterogeneity, moral hazard considerations typically lead to lower benefits for the long-term unemployed than for the short-term unemployed (see e.g. Hopenhayn and Nicolini (1997)). However, benefits for the long-term unemployed could also be higher because the unemployed run down their assets during the spell or because duration dependence reduces the moral hazard costs of providing benefits for the long-term unemployed. In addition, the equilibrium effects have to be taken into account and it is not clear if introducing screening matters mostly because of its influence on workers' search incentives or because of the equilibrium effects. These questions are addressed in our quantitative analysis in Section 1.5.

1.4 Estimation

So far we have described the data and some empirical facts followed by a discussion of the model and the mechanisms. In this section we will connect both by connecting our model to the data. We will first present the estimation setup and will then discuss the estimation results.

1.4.1 Setup

Specification. To estimate the model that we formulated in Section 1.3, we impose the following functional forms on the instantaneous utility function and the search cost function:

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}$$
$$\psi(s) = \frac{s^{1+\frac{1}{\lambda}}}{1+\frac{1}{\lambda}}$$

where λ denotes the elasticity of search effort with respect to the value of employment. The functional form is a common assumption and used in DellaVigna et al. (2017) or Lentz (2009). The instantaneous utility function is a standard CRRA utility function where γ is the risk aversion parameter and at the same time the inverse of the intertemporal elasticity of substitution.[28]

In our model agents are heterogeneous in two dimensions: (a) their probability of being qualified and (b) their initial assets. In our baseline version of the model we allow for two different productivity types π and three different initial asset types k_0, which in total leaves us with $J = 6$ types.[29] Signals are drawn from normal distributions with mean 0 for the low type and mean 1 for the high type.[30] We set initial assets for the unemployed to be uniformly distributed with 0, 500 and 3,000 euros. These values are set in order to match roughly the liquid assets of unemployed individuals in the PHF dataset. Every qualified type generates a profit $y > w$ for the firm in case he is qualified. y can be normalized because only the wedge between the vacancy cost and the $y - w$ gap is relevant for the determination of the vacancies. High types differ in their idiosyncratic match productivity. High types are qualified in π_H cases, while low types are qualified in π_L cases only. Unqualified applicants are always rejected. Hence, firms have an incentive to screen types with respect to their productivity in order to gain a higher expected profit. Since we do not aim to make any statements about production one can see these profits as a normalization. The wage agents receive during employment is fixed and we set $w = 1,606$ euros, which matches the mean re-employment wage in our sample of unemployed. The estimation is based on the current schedule, so that benefits b_t are set to a replacement rate of 63.5% within the first year and social assistance is equal to 40% after one year.[31] These numbers capture closely benefits

[28]Alternatively, one could think about a CARA utility specification. The constant relative risk aversion choice is motivated by the possibility of wealth effects, which implies different attitudes toward gambles with respect to wealth, i.e. individuals who have less savings will search more. Shimer and Werning (2008) compare the implications of CARA and CRRA to optimal UI and find only minor differences, because wealth effects are quantitatively very small in a search model like ours.

[29]Allowing for more types in both dimensions is easily possible but does not add any conceptual insights. Productivity and initial assets are uncorrelated, however, this can also easily be relaxed but has only negligible quantitative impacts.

[30]This is a pure normalization because we estimate the standard deviation of the normal distribution.

[31]UA is means-tested and a fixed amount. Hence we choose a value for the replacement rate that roughly amounts to the replacement rate that a typical UA recipient would receive.

paid to unemployed in our sample period. The vacancy posting costs are quadratic in the number of vacancies and we calibrate ex-ante the marginal cost of a vacancy to be equal to $\kappa = 100$. The functional form for the vacancy posting costs we use is $c(v) = \kappa v^{1+\rho}$, where we set $\rho = 1$ to obtain quadratic vacancy costs. The time horizon in our model is $T = 96$, which amounts to eight years. By choosing this relatively large time horizon we avoid that agents' search behavior is influenced by *end-of-life* effects.[32]

Estimation. Some additional parameters are set prior to estimation to standard values from the literature. We set the monthly time discount parameter equal to $\beta = 0.995$, which leaves us with an annual discount factor of roughly 5%. Risk aversion is equal to $\gamma = 2$ as in Chetty (2008) and Kolsrud et al. (2018). The interest rate is set to $R = \frac{1}{\beta}$ as in Chetty (2008), Lentz (2009), or Shimer and Werning (2008). This leaves us with the following parameters to be estimated:

$$\theta = \{\lambda,\ \pi_H,\ \pi_L,\ \alpha_L,\ \sigma\} \tag{1.9}$$

Thus the parameter vector contains the search effort elasticity λ, the productivity probability of the productive type π_H, the productivity probability of the unproductive type π_L, the unconditional type probability α_L and the variance of the signal σ.

In order to estimate the parameter vector θ, we apply a classical minimum distance (CMD) estimator as it is also applied by DellaVigna et al. (2017):

$$\min_{\theta} \ (m(\theta) - \hat{m})' W (m(\theta) - \hat{m}) \tag{1.10}$$

where $m(\theta)$ is the vector of model-implied moments, $\hat{m}$ is the vector of empirical moments, and W is the weighting matrix which we set to be equal to the identity matrix. The theoretical moments are simulated from the model and the reduced form moments are estimated as described in Section 1.2.2. The CMD criterion essentially chooses parameters in such a way, that the distance between the model-implied moments and the observed empirical moments becomes smallest.[33] For the estimation of the parameters we use a genetic algorithm, which is a global optimization routine.[34] Standard errors are then given by the diagonal elements of $(H'WH)^{-1}(H'W\Lambda WH)(H'WH)^{-1}/N$, where W is the weighting matrix, H is the Jacobian of the objective function evaluated at the estimated parameter values and Λ is a matrix with the inverse of the empirical moment variances on the diagonal.

[32] Mechanically, in $T = 96$ agents stop to search because it only provides disutility to them. This *end-of-life* effect also influences search effort in the periods before. However, in our specification these effects become small very quickly and do not influence search in a quantitatively important manner in the first years of unemployment.

[33] Note that in the estimation we use percent deviations instead of levels to give all moments the same weight.

[34] Global optimization routines are helpful for possibly non-differentiable problems and problems with local minima.

Moments. First, our moment vector includes the hazard moments from the first 24 months. Next, we include the average change in the search effort in month six and twelve relative to the first survey interview conditional on staying unemployed for one year. We also include the unconditional change in the search effort in month six and twelve relative to the first survey interview. Then we add the average number of acceptable applications that a vacancy receives as can be seen in Figure 1.1.[35] Finally, we add six multiple spell moments where we use the mean unemployment duration in spell two conditional on unemployment duration in spell one. Note that we mimic the multiple spell sample in our model by simulating two unemployment spells for workers with the same type and the identical level of initial assets. This preserves the intuition of the length of the first unemployment spell being informative about the second spell of a certain type, while avoiding to explicitly model job destruction and keeping our framework more in line with standard UI frameworks.[36] Figure A.4 shows this non-parametrically. The Figure shows that the longer an individual's UI duration is in the first spell the longer is the UI duration in the second spell. As discussed in Alvarez et al. (2016), the idea is that the stronger the correlation between the unemployment durations in the two spells, the more important heterogeneity must be. The relatively small slope of the curve suggests that duration dependence might be important and that heterogeneity is not the sole driver of the declining hazard. This leaves us with a total amount of 35 moments to match. Minimizing (1) with respect to θ gives us the estimated parameter vector.

Identification. The parameters are jointly identified if any parameter vector θ has distinct predictions for the behavior of agents. Intuitively, changing a certain parameter needs to have different implications for the moment vector $m(\theta)$ than changing another parameter. In our model, the level and slope of the hazard curve are closely aligned with the idiosyncratic productivity parameters π_j and the unconditional distribution of high types α_L. The search effort over the unemployment duration and especially the change in the search effort is informative about the search cost elasticity λ. The multiple spell moments deliver additional information on the unobserved heterogeneity in the model. The higher the slope of the curve of the mean durations, the more heterogeneity in job finding rates there should be. The intuition here is that the observation of two spells allows in principle to estimate a fixed-effect for individuals. If the correlation between UI duration in spell one is strongly correlated with UI duration in spell two, this hints towards sizeable heterogeneity (Alvarez et al. (2016)), and vice versa. This information is particularly helpful to estimate σ since the variance of the signal determines the importance of duration dependence in the model.

[35]To be very precise, we truncate the moment at 250 applications. However, only a handful of firms report that many acceptable applications.

[36]Empirically, we extent our sample to the period from 1983 until 2011 such that we have a sufficiently large sample of individuals with two unemployment spells.

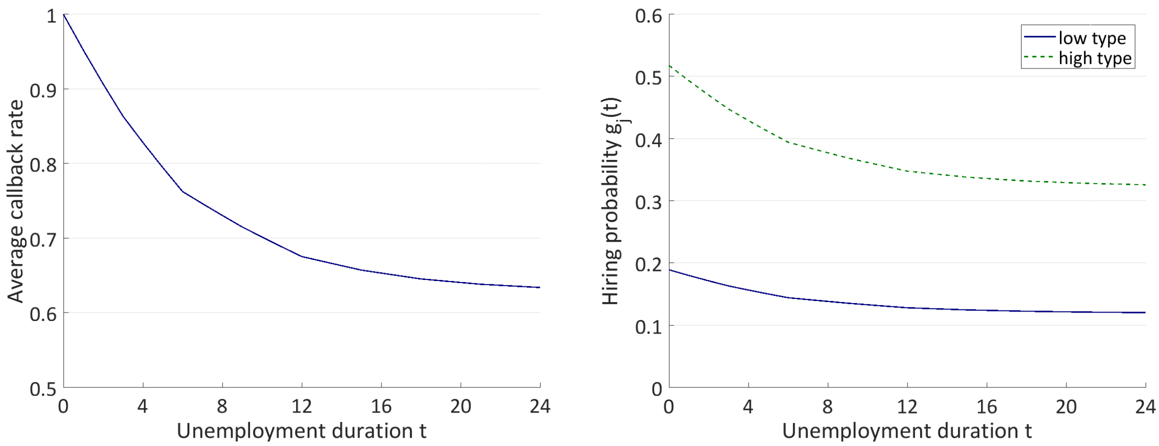
Table 1.2: Estimation results

Parameter	Estimate	s.e.
λ	2.539	(0.001)
π_L	0.213	(0.000)
π_H	0.576	(0.001)
α_L	0.648	(0.001)
σ	6.850	(0.003)

Notes: This table summarizes the estimation results of our parameters. Column two shows the estimated parameters and column three the respective standard error.

1.4.2 Estimation Results

In Table 1.2 we show the estimated parameters and the respective standard errors. We estimate the search cost elasticity λ to be 2.5, which is a relatively large elasticity of search effort with respect to the value of employment. This implies that agents will react relatively strong to benefit changes because a large responsiveness in search effort translates into large responses to benefit changes. The productivity probabilities and unconditional type probability suggest that the majority of individuals are of the low type ($\alpha_L = 0.685$), and that low types fulfill the requirements of the firm in roughly 20% of all matches, while high types fulfill the requirements of the firm in 58% of all matches. The heterogeneity in the productivity will translate into a heterogeneity in hiring rates as shown in panel (b) of Figure 1.3. We estimate the variance of the signal to be equal to $\sigma = 6.85$ which implies that the productivity is relatively noisy. In other words, signals are relatively informative and firms have a relatively strong incentive to screen applicants according to their unemployment duration because more high types are alive when an agent with a short duration is screened. To get a feeling for the importance of the signal versus the importance of the duration consider the case where only the duration is taken into account. Then the probability that an applicant with a shorter duration is interviewed is one. In our estimated model, the probability that a candidate with an unemployment duration of six months is screened versus a candidate who is unemployed for five months is between 0.31-0.38, depending on the agent type combination. Alternatively, the probability that a candidate with twelve months is screened relative to an applicant with eleven months of UI duration is between 0.28-0.35, depending on the type. If duration would be uninformative, the probabilities would be equal to 0.5. In panel (a) and (b) of Figure 1.3 we illustrate the screening and hiring behavior of firms that the model implies. Panel (a) shows the average decline in the callback rate of an application relative to period one. Our model suggests that the probability to get screened by a firm, i.e. the probability of a callback, declines throughout the unemployment spell and is only around 70% after one year and goes towards 60% after two years of unemployment. Note that callback rates for both types are very similar due to the large magnitude of σ. Hence, our model suggests only a small heterogeneity in



(a) Average normalized callback rate

(b) Hiring rates

Figure 1.3: Model-implied callback and hiring rates

Notes: The left panel shows the model-implied average callback rate of an application normalized to one in period $t = 1$. The right panel shows the type-specific hiring rates for unemployed that the model generates. The solid line corresponds to the low type and the dashed line to the high type.

the callback rate. This screening behavior translates directly into hiring rates since the hiring probability equals the callback probability times the productivity of the type, as shown in panel (b). For both types, hiring rates decline because the screening probability declines. However, the hiring probability per application of a high type is around 50% in the beginning because he is more qualified for firms than the low type. The low type has a hiring rate of 20% in the beginning which also declines the longer he is unemployed. Hence, we find considerable heterogeneity in productivity as well as important duration dependence in the hiring rate. The estimated heterogeneity and duration dependence in hiring rates then maps into job finding rates of agents. The job finding rate is the product of the hiring rate and the probability to send out an application, namely the search effort of the individual. The dashed line in Figure 1.4 shows the model-implied job finding rate of our model.

Model Fit. How well does our model fit the targeted data moments and how well does our model describe non-targeted empirical patterns? In terms of targeted moments the fit is extremely good. Figure 1.4 shows the fit of the hazard rate where the solid line is the data hazard and the dashed line the model-implied hazard. We are able to fit the hazard curve in basically every month except the time around the benefit exhaustion.[37] Table 1.3 shows the additional targeted data moments and the model implied moments. We can fit the unconditional and conditional changes in the search effort very well and also the second spell moments by capturing a positive slope. Finally, we slightly over-predict the mean number

[37]Here, other factors might be important, e.g. that people exit registered unemployment because they are not eligible for social assistance. Because we do not model these features we disregard the spike at benefit exhaustion. See DellaVigna et al. (2017) for an exploration with present-biased and reference-dependent agents.

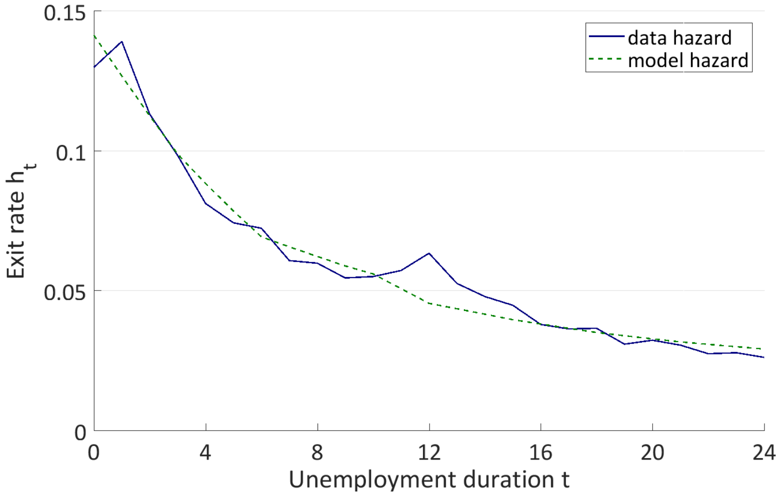


Figure 1.4: Model fit: Hazard rates

Notes: This figure illustrates the model fit of the job finding rate. The solid line corresponds to the data hazard and the dashed line corresponds to the model-implied job finding rate.

Table 1.3: Data moments versus model moments (excluding hazard)

Moment	Data	Model
Unconditional change in search effort $t = 6$	0.710	0.763
Unconditional change in search effort $t = 12$	0.601	0.618
Conditional change in search effort $t = 6$	0.740	0.751
Conditional change in search effort $t = 12$	0.730	0.599
Mean duration second spell bin [1,4]	0.118	0.108
Mean duration second spell bin [5,8]	0.129	0.116
Mean duration second spell bin [9,12]	0.139	0.123
Mean duration second spell bin [13,16]	0.136	0.132
Mean duration second spell bin [17,20]	0.138	0.140
Mean duration second spell bin [21,24]	0.134	0.148
Mean acceptable applications	4.302	5.760

Notes: This table shows the fitted moments from our model. In the second column one can see the data moments and in the third column the model-implied moments. The 24 hazard moments are excluded from the table and can be seen in Figure 1.4. The second spell moments are divided by 100.

of applications a firm receives. Indeed, the data moment is equal to 4.3 while the model implied mean number of applications is 5.8.

These are two important pieces of evidence that we did not directly included in our estimation: (a) callback rates and (b) duration elasticities with respect to potential benefit durations. Kroft et al. (2013) find in an experimental audit study that the callback rate from an application declines by about 40 percentage points after one year. In addition, the JVS data suggest that 40 percentage points of firms are not willing to consider unemployed applicants with an unemployment duration of one year or more as shown in Figure A.2. Our model indeed implies a very similar pattern in terms of callback probabilities. As discussed above our estimated model predicts a very similar average decline in callback rates. This makes us confident that the magnitude of the estimated screening channel in our model is plausible, since it compares well to the empirical findings on firm-induced duration dependence.

In Schmieder et al. (2012) the authors exploit quasi-experimental variation in age cutoffs of potential benefit durations in Germany. If one looses his job above a specific age cutoff the maximal potential benefit duration increases from 12 to 18 months. In their paper they implement a regression discontinuity design and find that additional six months of benefits increase the mean non-employment duration by 0.78 months. In our model, we can perform this simulation and we find that a benefit extension of six months implies an increase in the mean duration by 0.81 months. This is extremely close to the causal estimate from the data and makes us confident that our estimate of the search elasticity λ is reasonable. It ensures that the model-implied responsiveness to benefits is realistic. Since we are finally interested in optimal unemployment insurance we want to have plausible behavioral patterns with respect to benefit payments.

Robustness. Our model is estimated using a genetic algorithm routine. The advantage of this approach is a solution that can better handle non-differentiable objective functions and is better suited to find the global solution in a problem with possibly many local minima. However, the drawback is that it is a stochastic optimizer and possibly delivers different estimates in each estimation. Therefore we were running a bunch of estimations with different bounds on the parameter spaces and different initial population spaces. The estimates were always very similar to the reported ones above. We have chosen to report the set of parameters that attained the smallest value of the criterion function. We also tried to use different moments for the estimation including 12 or 35 hazard moments, dropping search moments, dropping multiple spell moments and different definitions of the mean number of applications. In all cases, the estimates were close to the reported ones. We also have tried different functional forms and specifications of the pre-determined parameters. There the estimated parameters naturally differ by more, however the qualitative features and conceptual predictions stay the same. Note that two particular specifications are important for the results: (a) the risk aversion parameter γ and (b) the curvature of the vacancy cost

ρ, which we assume to be quadratic. The higher the risk aversion γ the larger demand for insurance and the higher optimal UI benefits. Second, the larger the curvature of the vacancy cost function the less responsive are vacancies in equilibrium. This can then determine the sign and magnitude of the applications-per-vacancy channel which translates into either increasing or decreasing hiring rates. For our baseline specification we have used parameters that are either in line with previous literature as discussed above or deliver the best fit to our data moments.

So far, we did not allow for observables like gender, education and other observables from our model. One might suspect that job finding rates differ for these groups and that there is sorting along the unemployment spell on observables which might affect our findings. Therefore, we have computed observable-adjusted hazard rates which were extremely similar to the average hazard rate that we report. We tried restricting the sample to men and different time periods. Again, the hazard rates, the search behavior of agents and other data moments were very similar. It might be that less educated individuals or older individuals survive longer in unemployment and that this creates heterogeneity that our model wrongly attributes to heterogeneity in unobservables. We have therefore created samples for observable education, age and gender cells and compared job finding rates. Besides minor differences in the level there was basically no difference in the decline in the hazard. This is a consequence of only little sorting along the unemployment spell in terms of observables. In Figure A.5 and A.6 in the appendix we have plotted the mean education of the unemployed sample along the unemployment duration and the fraction of female along the unemployment duration. We see that the curves are pretty flat and that there is not much sorting in terms of observables. This makes us confident that ignoring observables in our model is a good approximation in our setting and allows us to work with a more parsimonious model.[38]

1.5 Welfare Analysis

In this section we use the estimated model for welfare analysis by solving for the optimal policy problem discussed in Section 1.3. Afterwards we compare the optimal policy to different counterfactual policy simulations, followed by a discussion of more flexible UI schedules.

1.5.1 Optimal Policy Results

To solve the optimal policy problem outlined in Section 1.3, we solve the model on a grid for the policy parameters b_1 and b_2 and for each potential benefit duration D, using 1 percentage point steps for the benefit levels. The tax is automatically calculated via the

[38] To save space, we do not report figures and tables on the discussed robustness checks. All of the robustness checks and alternative specifications are available on request from the authors.

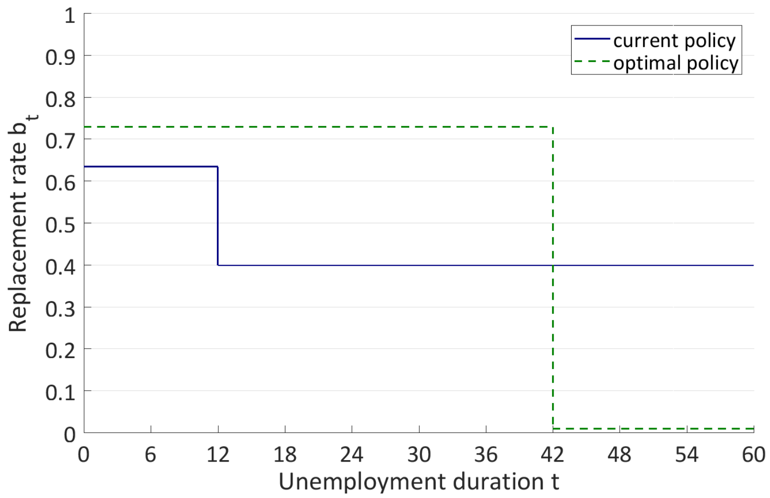


Figure 1.5: Optimal UI versus current UI

Notes: In this graph we compare the current UI policy in Germany (solid line) to the optimal policy suggested by the estimated model (dashed line). The x-axis shows the unemployment duration in months and the y-axis the replacement rate of benefits in terms of the past wage.

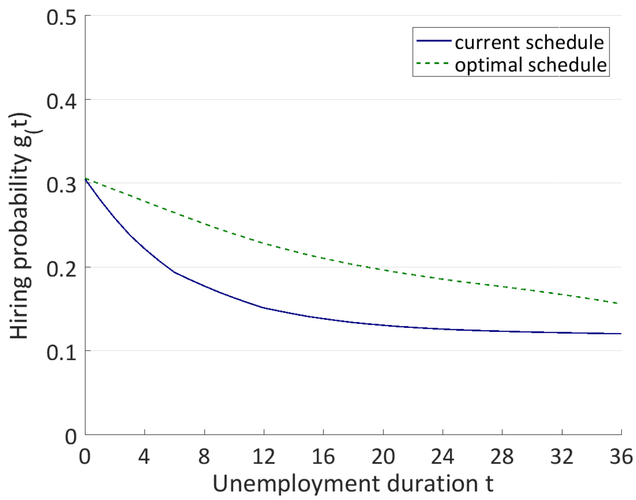
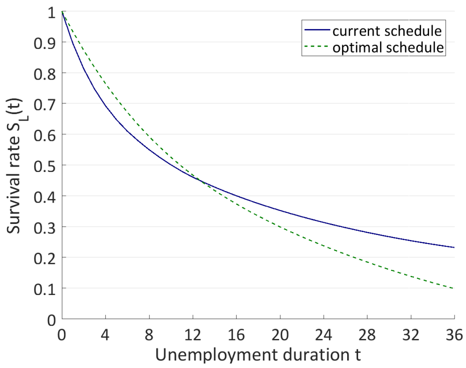
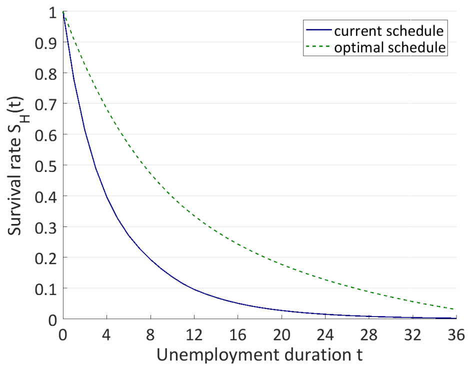


Figure 1.6: Optimal UI versus current UI

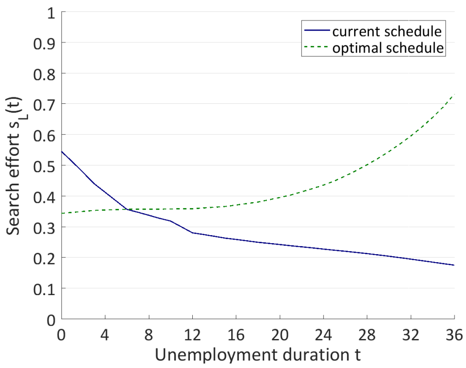
Notes: In this graph we compare the current UI policy in Germany (solid line) to the optimal policy suggested by the estimated model (dashed line). The x-axis shows the unemployment duration in months and the y-axis the replacement rate of benefits in terms of the past wage.



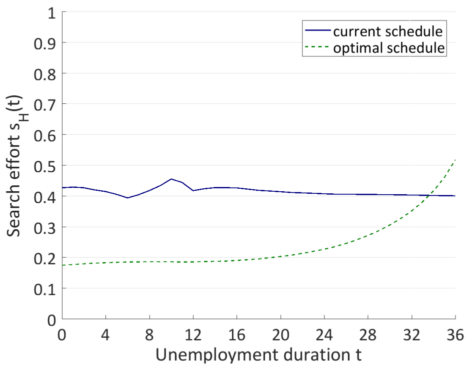
(a) Survival of unproductive type



(b) Survival of productive type



(c) Search effort of unproductive type



(d) Search effort of productive type

Figure 1.7: Counterfactual model simulations

Notes: The above panels show counterfactual model simulations of the search effort of unemployed and the survival probability in unemployment. Panel (a) shows the survival in unemployment of the unproductive type under the current policy (solid line) and the optimal policy ((dashed line) as a function of the unemployment duration in months on the x-axis. Panel (b) shows the same for the productive type. Panel (c) shows the search effort of the unproductive type under the current policy (solid line) and the optimal policy (dashed line) as a function of the unemployment duration in months on the x-axis. Panel (d) shows the same for the productive type.

budget constraint. This gives us the global optimum of the welfare problem. The dashed line in Figure 1.5 shows the optimal policy schedule implied by our model. To have a meaningful benchmark we compare the optimal schedule to the current UI schedule in Germany as shown in the solid line in Figure 1.5. The current policy pays benefits for one year and offers social assistance thereafter. We find that the optimal policy should pay 73% of the wage in the first 42 months and a 1% replacement rate afterwards.[39] As one can see the optimal schedule differs substantially from actual policies. Our main finding is that benefits should be (a) higher in the first years, (b) paid for around three and a half years and (c) be very low afterwards. The resulting optimal schedule is a combination of incentivizing agents to search enough, providing insurance to budget constraint agents and to account for firms hiring, screening and vacancy responses.

To build intuition for the relevance of equilibrium effects for the optimal policy result in Figure 1.5 consider the average hiring rate of unemployed in Figure 1.6. Figure 1.6 plots the average hiring rate of unemployed job seekers as a function of the unemployment duration. The solid line shows the hiring probability under the current policy, i.e. at the estimated level. In contrast, the dashed line shows that the hiring probability is less declining with unemployment duration under the optimal policy. Hence, the planner reduces the importance of screening by duration and shifts the hiring probability of agents upwards. A higher hiring probability suggests that firms are more willing to hire the long-term unemployed. Panel (a) and (b) of Figure 1.7 illustrates why this happens when the optimal policy is implemented. Panel (a) and (b) show the survival probability of the unproductive and productive type, respectively. Note that in the long term, under the current schedule some unproductive types stay unemployed for very long, while under the optimal policy after four years almost all unproductive types are working. In both panels the solid line shows the survival probability at the estimated level and the dashed line under the optimal policy. One can see that the optimal policy considerably alters the dynamic composition of the unemployment pool. As panel (a) and (b) suggest, at any point of the unemployment duration the relative composition changes towards the productive type, i.e. at any point there are relatively more good types unemployed compared to the current setting. This in turn implies that firms are more likely to consider long-term unemployed because the pool of applicants is of a better quality under the optimal policy. The changed composition of unemployed is a result of the change in search incentives for the two types as illustrated in panel (c) and (d) of Figure 1.7. Again, the dashed line shows the search effort under the optimal policy and we compare it to the setting with the current policy. Under the optimal policy the unproductive type is incentivized to search more while the productive types searches less on average. Hence, the composition of unemployed will move towards the productive types because now relatively more unproductive types exit early in their spell. Hence, the planner considerably alters

[39] We solve for the optimal policy on a discrete grid and can therefore evaluate the welfare for each policy. We find that the optimal policy is unique because no other policy schedule leads to the same welfare.

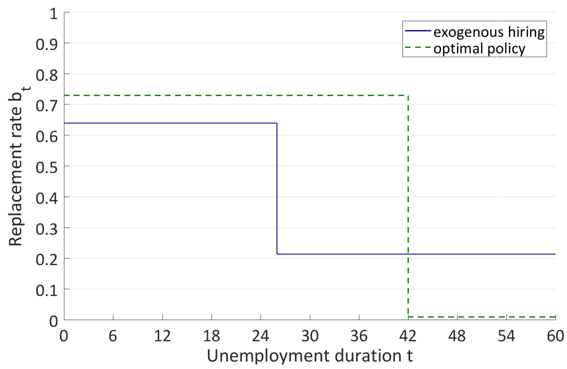
search behavior of agents and hence the hiring and screening behavior of firms. However, Figure 1.5 does not allow to distinguish how important these endogenous firm responses are in terms of changing the optimal policy, relative to a setting without endogenous firm responses where only search incentives and insurance motives are at work. We will discuss the relevance of this firm adjustments and how they shape optimal policy in equilibrium in the next subsection.

How large is the welfare gain of moving from the current policy to the optimal policy for the unemployed? In other words, how much cash-on-hand would we need to pay an unemployed individual under the current regime such that he is as well off as with the optimal policy? When we implement this experiment we find that the gain of moving to the optimal policy amounts to a lump-sum payment of nearly $5,500$ euros to an unemployed at the beginning of his spell. This is a fairly large amount and moving to the optimal policy implies a large welfare gain in our model.

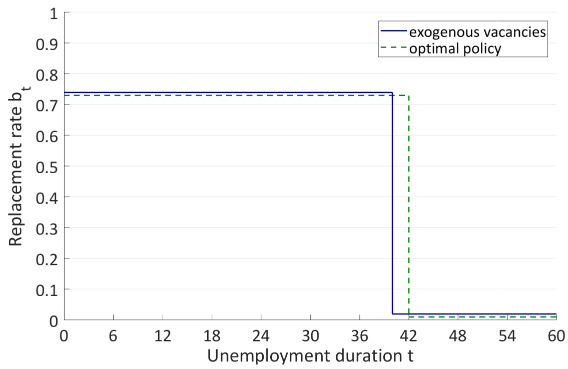
1.5.2 Discussion

To show the quantitative importance of firm responses for the baseline result presented in the last subsection, we perform various counterfactual simulations to decompose the importance of firm responses for the optimal policy design problem.

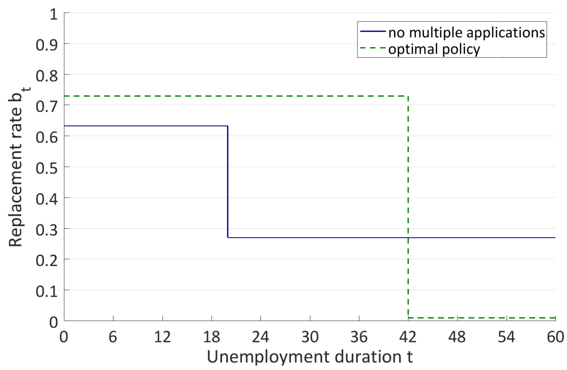
Exogenous hiring rates. In the baseline model hiring rates for job seekers are endogenous to UI policies. As we have illustrated in Section 1.3 higher benefits can lead to higher hiring rates through the adjustment of firm beliefs about the pool of applicants and through changes in the applications-per-vacancy ratio. If we fix hiring rates for job seekers at the level of the estimated model under the actual policy in place and then re-solve the planner problem we can decompose the component of the optimal UI policy that can be attributed to the endogenous firm responses, namely hiring rates and vacancy creation. In panel (a) figure 1.8 we compare the optimal UI policy in the baseline model (dashed line) with the optimal policy when hiring rates are exogenously set at the level of the estimated model (solid line). We find that the schedules substantially differ and that UI with exogenous hiring is less generous and paid for a shorter amount of time. Benefits after two years are however higher with exogenous hiring rates, which is because with exogenous hiring more agents survive longer in unemployment and the insurance motive becomes stronger. To be more precise, endogenous hiring rates allow the planner to lift up these hiring rates by providing different incentives for job seekers and firms. By implementing the optimal schedule the planner increases the value of search and therefore reduces long-term unemployment. However, this is an equilibrium effect, because more search effort of job seekers increases the value of a vacancy and the expected profit of hiring. These equilibrium adjustments are absent in partial equilibrium models. What panel (a) shows is that a large part of the benefit extension compared to the



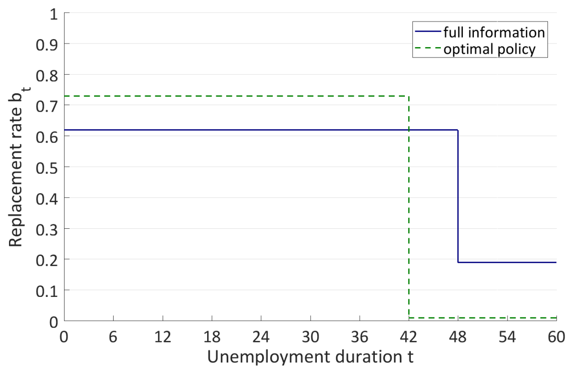
(a) Exogenous hiring rates



(b) Exogenous vacancies



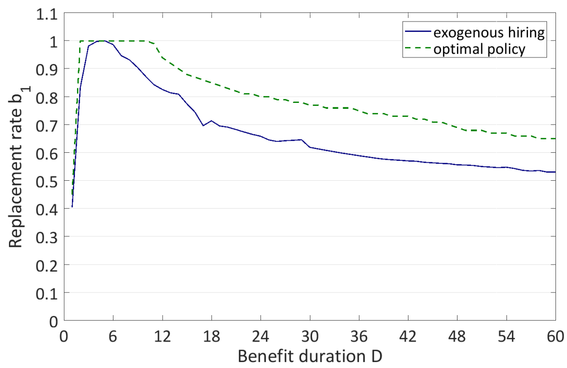
(c) No multiple applications



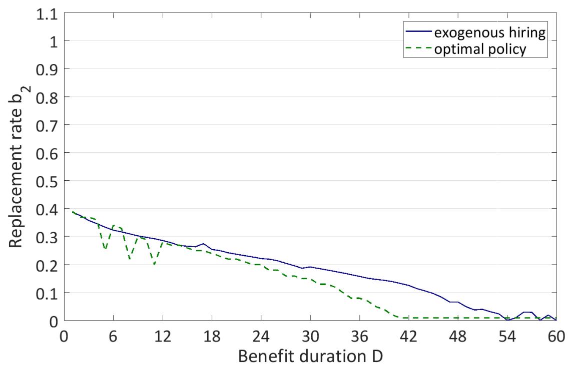
(d) Full information

Figure 1.8: Counterfactual policy results

Notes: This figure compares the optimal policy of our baseline model (dashed line in both panels) with different counterfactuals. Panel (a) compares to a setting where hiring rates are policy invariant at the level of the estimated model (solid line in panel (a)). Panel (b) to a setting where the mass of vacancies is policy invariant at the level of the estimated model (solid line in panel (b)). Panel (c) to a setting without multiple applications and without screening (solid line in panel (c)). Panel (d) to a setting where firms observe agents' productivity, i.e. $\sigma = 0$ and signals are informative (solid line in panel (d)). In all panels, the x-axis shows the unemployment duration in months and the y-axis the replacement rate of benefits in terms of the past wage.



(a) Optimal b_1 benefits



(b) Optimal b_2 benefits

Figure 1.9: Non-linear optimal UI

Notes: Panel (a) illustrates benefits b_1 for the baseline case (dashed line) and the case with exogenous hiring rates (solid line) as a function of the potential benefit duration D on the x-axis. The y-axis denotes the optimal benefit level for b_1. Panel (b) shows the same two cases for optimal benefits b_2 as a function of the benefit duration D.

actually implemented schedules is driven by endogenous firm responses. The reason for this finding is that even small changes of hiring rates can create large changes in search effort and survival rates. This shows that incorporating endogenous hiring decisions is quantitatively very important for welfare conclusions in terms of optimal UI policies, because it changes the optimal benefit level and benefit duration in a non-negligible manner.

Exogenous mass of vacancies. The above finding in panel (a) of Figure 1.8 is a mix between vacancy responses and hiring rate responses of firms. Therefore, in panel (b) of figure 1.8 we exogenously fix the amount of vacancies in the economy at the level of the estimated model and allow hiring rates to be endogenous. This experiment allows us to decompose the importance of the hiring response, i.e. the applications-per-vacancy channel and the firm beliefs, holding fixed the number of open positions. We find that the vacancy channel is quantitatively very small and optimal benefits are similar to our baseline policy where vacancies are allowed to adjust in equilibrium. Hence, the longer potential duration and the generosity of benefits is mainly driven be the endogeneity of hiring rates, not vacancies.

No multiple applications. This naturally leads to the question how optimal UI would look like in our model if there was only one application per vacancy, i.e. there are infinitely many vacancies and no crowding-out among applicants. This limiting case where vacancy costs κ are equal to zero is an important benchmark for our model, because it shuts down the employer screening channel. This implies that every applicant gets screened and hired in case he is qualified. The difference to the exogenous hiring case is that the callback rate is flat and that there is no duration dependence in the model. The single applications per vacancy limit is equivalent to a standard partial equilibrium search model with heterogeneity in job arrival rates. Figure 1.8 panel (c) illustrates the optimal policy in this setting compared to our screening model with multiple applications per vacancies. The solid line shows how optimal UI should look like in the absence of employer screening. Interestingly, the optimal schedule is close to the actually implemented schedule. The only difference is that benefits are paid a few months longer and that b_2 is somewhat smaller. The optimal schedule is similar to the case with exogenous hiring, however D is smaller when there are no multiple applications. This is because if agents do not need to compete with other applicants their job finding rates are higher and the demand for insurance is lower.

Full information of vacancies. One additional interesting comparison on the importance of screening is the full information case where there are multiple applications but firms perfectly observe agents' types and productivity. In this case hiring rates become flat and true duration dependence disappears, but the applications-per-vacancy ratio is endogenous and not equal to one. The solid line in Figure 1.8 in panel (d) shows how optimal UI looks like if there is full information about the productivity of applicants. Because this implies lower job finding rates of bad types the demand for insurance, even in the long term increases. Hence, optimal UI is paid for longer (48 months) and b_2 is at a higher level.

Fully dynamic UI schedules. So far we have restricted to optimal UI schedules with four policy parameters. This is for two reasons: (a) our optimal schedules mimic current policies and (b) solving the government problem with more flexible parametrizations is numerically not feasible. However, we can illustrate fully flexible optimal UI policies with a distinct b_t for each unemployment duration t by calculating the optimal b_1 and b_2 level for each potential duration D. This gives some indication about the shape of a more flexible schedule. In Figure 2.1 panel (a) the dashed line shows how the optimal b_1 level in the baseline is set as a function of the potential duration D on the x-axis. In panel (b) the dashed line shows how the optimal b_2 level is set as a function of the potential benefit duration D. Panel (a) suggests that optimal benefits should follow a hump-shaped pattern and that UI benefits should be increasing in the first months of unemployment and be decreasing thereafter. To see this, note that if only paid for 1 month, the optimal level of b_1 is only about 0.4. If paid for two months, however, this level is higher, which can only be the case if the optimal schedule is increasing at first.

The solid lines in the two figures allow us to compare the optimal shape to the setting with exogenous hiring rates at the level of the estimated model. One can see that under screening b_1 is increasing faster and stays at a high level for longer than in the setting without screening.[40] Hence, fully dynamic optimal UI schedules under screening should follow a more

[40]Note that we restrict to policies with $b_t \leq 1$ because benefit levels above the wage are not of practical interest. However, this restriction leads to numerical fluctuations in panel (b) as one can see with the spikes in the optimal b_2 level at durations where b_1 hits the upper bound. The spikes disappear if the upper bound is set to a higher level.

pronounced hump-shape and be more generous and paid for a longer time than in a setting with exogenous hiring rates, which is perfectly in line with our more restrictive policy results in the baseline case with four policy parameters.

Alternative Parametrizations. In the appendix Figures A.7, A.8, A.9 and A.10 we show some additional alternative parametrizations of the model to check how important various parameters and assumptions are for the optimal policy outcomes. Naturally, the risk aversion of agents matters for the generosity of benefits. The more risk averse agents are the longer the potential benefit duration and vice versa. As DellaVigna et al. (2017) suggests agents seem to have large discount factors or behave as if they are present biased. Therefore, as an alternative we use $\beta = 0.95$ which amounts to an annual discount factor of 0.54. If agents discount the future at the higher rate, benefits are higher early on but lower later in the spell, which is exactly what one would expect if agents value the present more relative to the future. The elasticity of the vacancy creation channel seems to be not very important for optimal UI schedules as we show in the appendix. Finally, if we assume that all agents start without assets to the unemployment spell, then optimal UI policy hardly changes compared to the optimal schedule.

1.6 Extensions

In this final section we will discuss three extensions of our model and how they would alter our findings: multiple applications of the unemployed, screening costs of the firm and endogenous wages.

Multiple applications per worker. While we focused on the case of each worker sending out at most one application, it is also possible to consider the general case where workers can send out more applications. The main advantage of this extension is that it allows the model to replicate the observed facts about the number of applications individuals send (see Figure 1.1) more directly.

Following Kaas (2010) and Shimer (2004), a convenient way to include multiple applications is to allow workers to search with continuous search intensity s and stochastically send out a number of applications that follows a Poisson distribution with mean s. In this case, the hazard rate is the expected probability of at least one application resulting in an offer, $h_j(t) = 1 - \exp\Big(-g_j(t)s\Big)$, and $g_j(t)$ has the interpretation of being the endogenous *success probability* of each application, while s is the expected number of applications sent.[41] Introducing multiple applications in this way does not change the rest of the model.

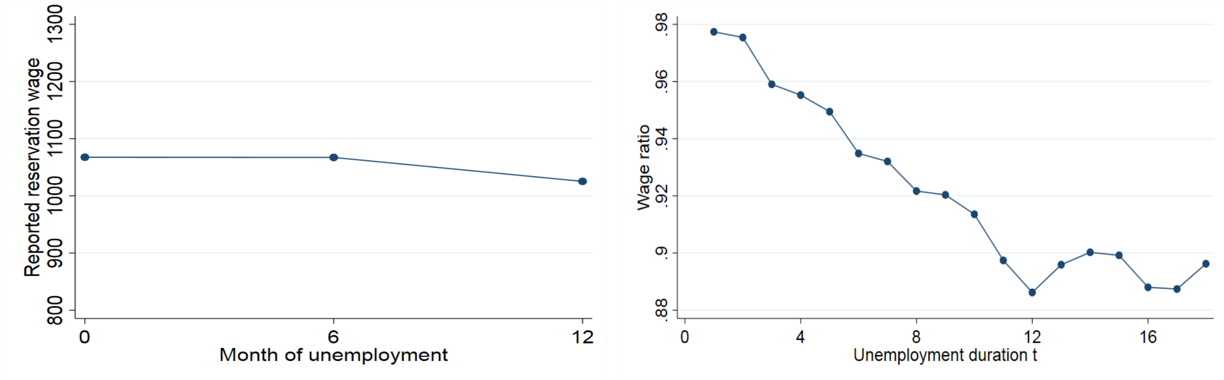
We experimented with this version of the model and the results are qualitatively similar. A main difference is that multiple applications, in principle, introduce another coordination friction, since agents get multiple offers and can accept only one. As a result, some vacancies make offers that are rejected. This gives rise to the question if these firms should be allowed to contact other applicants, if their first offer gets rejected. Otherwise, the coordination friction reduces firm profits and therefore the number of vacancies. There are different approaches to this issue in the literature. Some recent paper allow for *recalls*, i.e. the possibility to contact other applicants (see e.g. Kircher (2009)), while others do not (Kaas (2010), Gautier et al. (2016), Albrecht et al. (2006)). Without recall, it can be desirable to make workers search less, since this makes the additional coordination friction less severe and increases entry. For simplicity, and since we do not want to focus on this additional coordination friction, we restrict ourselves to the case of one application per worker, as is also done in Fernandez-Blanco and Preugschat (2018) or Villena-Roldan (2012).

Screening costs. Another possible extension is to make screening costly for firms, rather than assuming that screening costs are tiny. In our setting, firms would still screen all applicants for most realistic values of the screening cost (since the lower bound of the expected profit is $\pi_L y$, which is the expected profit of the low type).[42] While one could argue that the screening costs are included in the vacancy posting costs, an interesting feature of introducing screening costs is that it would make the vacancy cost partially endogenous: when unemployment duration or signals are not informative, firms on average have to screen more applicants before finding a qualified one and would have less incentives to create vacancies. From a policy perspective, screening costs may provide a rationale for trying to make duration informative, since this would make hiring easier for firms. In the current version of the model, the potential welfare gains from a decrease in screening already have to be weighted against the potential decline in the number of vacancies. Screening costs would amplify the latter effect.

Endogenous wages. A further extension would be to depart from the assumption of a fixed wage. Our main motivation for this assumption is that it is a reasonable approximation of the empirical evidence, which is discussed below, and that introducing endogenous wages in our framework likely makes the analysis much less tractable. In standard matching models with just one applicant per vacancy, wages are often assumed to be determined by

[41] A worker who sends a applications gets at least one offer with probability $1-(1-g_j(t))^a$ and the expression results from taking the expectation over a, which follows a Poisson distribution with mean s. It is interesting to note that this setting provides a micro-foundation for using $1-\exp(-\lambda s)$ as a functional form for the arrival rate, which is commonly used in partial equilibrium models.

[42] See Jarosch and Pilossoph (2018) for a discussion of how to calibrate a parameter for screening costs.



(a) Reservation Wage **(b)** Wage ratio

Figure 1.10: Reservation wages and realized wages by unemployment duration

Notes: The left panel shows the mean reported reservation wage (from the IZA ED) of individuals who have been unemployed for zero, six or twelve months. The right panel shows the (realized) ratio of the wage before and after unemployment (based on SIAB data) by unemployment duration.

Nash bargaining. However, this is more problematic when there are multiple applicants per vacancy, since firms would have to simultaneously bargain with each of the applicants. With wage posting, on the other hand, characterizing the equilibrium becomes challenging, especially in our context of endogenous search effort and savings, both of which are important for the analysis of optimal UI.[43]

From an empirical point of view, there is increasing evidence to support the assumption of a fixed wage, conditional on worker characteristics. For example, Krueger and Mueller (2016) find that reservation wages stay remarkably constant over the unemployment spells. Hall and Mueller (2015) show that individuals often accept the first job offer they get. Their evidence also suggests that relatively few individuals have the opportunity to bargain about their wages, but rather face the option to accept fixed offers. Our datasets support these findings for reservation wages as can be seen in Figure 1.10 panel (a). There one can see that self-reported reservation wages are essentially flat throughout the unemployment spell. In addition, in the JVS data employers report whether the hiring process included some form of wage bargaining with the applicant and only 34% of firms report that this was the case. Looking at realized wages, Figure 1.10 shows that the average ratio between the post- and pre-unemployment wage drops fairly moderately from 98% to 90% after one year, even without controlling for selection on observables throughout the spell.[44]

[43] Fernandez-Blanco and Preugschat (2018) consider the case of wage posting with directed search, but assume that workers do not know their type and that there is no effort or savings choice.

[44] See also Schmieder et al. (2016) for a more detailed analysis of the wage effects throughout the unemployment spell.

1.7 Conclusion

This chapter has analyzed a dynamic search model where firms can choose from a pool of applicants and have incomplete information about their quality. Firms rank applicants by their expected productivity, which makes it less likely that the long-term unemployed are invited for interviews in the presence of other applicants. The model is estimated to match several important features of the data regarding job-finding rates, search effort and vacancies.

Our welfare analysis suggests that equilibrium effects in the form of endogenous hiring and interview decisions are quantitatively very important for the optimal design of unemployment benefits. We find that allowing for these equilibrium effects leads to benefit schemes that are more generous in the first place, benefits are paid for a longer time, but benefits are very low at longer unemployment durations. More generally, our results demonstrate that modeling the details of the hiring process can have quantitatively sizable implications for optimal UI policy and that this requires integrating features into search and matching models that have often been abstracted from - most importantly, the possibility of multiple applications per vacancy.

An interesting aspect that we have not made explicit so far is that long-term unemployment is not such a bad signal in recessions when the applications-per-vacancy ratio is high. If there are many unemployed applicants per open vacancy then screening matters more and benefits can be more generous and paid for a longer time. Hence, our findings can rationalize benefit extensions as those implemented through the Great Recession in the US. Another important question for future research is to find additional quasi-experimental evidence on the importance of competition for jobs among many applicants and employer screening. Reduced-form evidence on how hiring decisions respond to unemployment policies would nicely complement our more structural approach. Alternatively, one could think of non-standard policy instruments, for example hiring subsidies for firms that are used in some countries.[45] Such an instrument could help to mitigate screening by giving firms incentives to screen unemployed with a long unemployment duration.

[45]In 2014, the German government announced to spend 150 million euros on wage subsidies for the long-term unemployed.

Chapter 2

The Marriage Market, Inequality and the Progressivity of the Income Tax

2.1 Introduction

A progressive income tax is a common policy tool to reduce inequality in consumption and welfare. The debate on tax progressivity has largely focused on redistribution *across* families and has paid little attention to the question how the allocation of consumption and time *within* families could be affected.[1] Recent empirical work finds that there is substantial inequality in consumption and leisure within households and that intra-household allocations depend on the relative income of spouses (Lise and Seitz (2011), Lise and Yamada (2018)). This suggests that in marriages where one spouse has a higher earnings potential than the other, increasing tax rates at higher incomes and lowering them for lower incomes could improve the bargaining position of the lower-earnings spouse. This bargaining effect would reduce intra-household inequality in consumption and leisure, on top of the usual effect of redistributing from richer to poorer households. In addition, progressivity could affect inequality through its impact on marriage and divorce. Whether an individual is single or married, and whom they are married to, has a significant impact on their living standard. How does tax progressivity affect intra-household inequality and the marriage market? And how much do these channels affect inequality relative to the usual effects of progressivity?

To address these questions, I develop an equilibrium search and matching model with intra-household bargaining, labor supply and savings. Wage inequality results from differences in labor market ability and shocks over time. Couples benefit from joint consumption, can pool home production time and share labor market risk, but not all goods are shared equally between spouses. Personal consumption and leisure are private goods. Couples decide efficiently on allocations, but are unable to fully commit to future allocations (as in e.g. Mazzocco et al. (2013)). The bargaining position of a spouse is given by how well-off they

[1] See e.g. Heathcote et al. (2017) or Conesa and Krueger (2006).

would be as single. Bargaining positions are linked to the marriage market, since being single includes the possibility of marrying in the future. In the example of a high-wage husband married to a low-wage wife, the husband would be able afford a higher living standard if he was single and would have better future marriage prospects. In this case, a progressive income tax improves the outside option of the wife, since it increases the after-tax income she would earn as single. This raises her share of consumption and leisure in marriage and reduces intra-household inequality.

Tax progressivity further affects marriage and divorce, besides the sharing of resources in marriage. When meeting on the marriage market, individuals have to decide whether to get married or to stay single and wait to meet someone else. They take both the economic characteristics of their potential partner and a non-economic match quality into account. The main mechanisms through which tax progressivity affects marriage and divorce are that it influences the economic value of individuals on the marriage market and how selective they are about potential partners. An increase in progressivity makes low-income individuals richer and they become more attractive on the marriage market and can afford to be more selective. The effects on high income individuals are the opposite. In addition, from the point of view of each individual, progressivity makes all potential partners more similar in their economic characteristics and reduces the value of staying single in order to continue search.

I calibrate the model to Dutch survey data on intra-household allocations, labor market outcomes and marital histories. Before turning to the policy analysis, I first decompose cross-sectional inequality in private consumption and welfare for the status quo. I decompose the population variance into contributions due to inequality within and between married couples, and due to singles. The within-couple component captures that spouses consume different amounts of private goods, which is typically absent in studies of progressivity. In the calibrated model, the within-component accounts for 23.9% of the total variance of private consumption, singles for 21.4% and the between-couple component for 54.3%. The relative magnitude of the within- and between variance in the model compares well to the data on private consumption from the Dutch LISS panel. On top of private consumption, the calibrated model allows to study inequality in the *utility* from private and public consumption, leisure and home production. The advantage of this measure is that it allows to aggregate welfare from consumption expenditure and time use and takes all public and private goods that are available into account. The contribution of the within-couple variance to total inequality is 6.1% in this case. This reflects households spending part of their time and expenditure on public goods, which are consumed equally by both spouses. As a result, the contribution of the within-couple variance is smaller than for private consumption only.

I then study the effects of a reform that increases in progressivity and analyze the impact on inequality within and across households. The main experiment is a budget-balanced increase in progressivity, by 40%, which increase average tax rates at higher incomes and decreases them at lower incomes.[2] As the majority of OECD countries, the Netherlands has a system of individual taxation, in which spouses are taxed based on their individual earnings.[3] The focus of the analysis is to decompose the total reduction in inequality in terms of changes in within- and across-household inequality. The reform decreases the variance of private consumption by 10.9% and reduces both the within- and the between- household variance. The contribution of the within-household component to the total reduction in inequality is 24.77% for private consumption and 11.93% for the utility measure. Thus, in terms of utility from private and public goods, 11.93% of the total reduction in inequality would be ignored by abstracting from intra-household inequality.

The increase in progressivity in principle affects inequality both through its effects on bargaining positions and on marriage rates and assortative mating. Since there are economies of scale in consumption and home production, married individuals have a higher living standard than singles. Thus, changes in marriage and divorce contribute to the effect of the reform on inequality. In particular, inequality across couples could be affected by a change in the composition of couples. I further use the model to quantify the role of these composition effects. The model rationalizes the observed gender imbalance in domestic work hours with a gender asymmetry in the home production technology, which makes women more likely to reduce their market hours in marriage. This is reflected in matching patterns and low-wage men have a lower probability of getting married, relative to higher wage men. The increase in tax progressivity induces a small increase in the probability than low-wage men get married to higher-wage women. This mechanism alone would generate an decrease in assortative mating. I decompose the variance across couples further to investigate whether such composition effects have a noticeable impact on inequality and find that their role is very small. Composition effects explain between -0.3% and -2.8% of the reduction in inequality within and across couples.

In addition to cross-sectional inequality, I further study how the reform affects expected life-time welfare. I isolate the role of the marriage market channels (marriage and bargaining) by also computing the welfare effect of the reform for an alternative version of the model, in which marriage and divorce decisions and the intra-household bargaining weight are fixed to the pre-reform decision rules. This can be interpreted as the impact of the reform if marriage market decisions were exogenous and would not react to the reform. I then compare the welfare effect of the policy between the full and the alternative model. The contribution of the marriage market is the difference between these two cases. Without the marriage market adjustments, the reform leads to a welfare reduction by 1.22%, in terms of consumption equivalents relative to the status quo, whereas welfare decreases by only 0.9% when taking the endogenous marriage market responses into account.

[2] The change in progressivity refers to the progressivity parameter of the tax function used in Heathcote et al. (2017) or Holter et al. (2019). For details, see section 2.3.3.

[3] Exceptions are Germany, the United States and France where tax liabilities are computed based on the income of both spouses.

The literature on tax progressivity has mostly focused on inequality across households and abstracted from intra-household inequality and endogenous marriage and divorce. For example, in the macroeconomic literature, Conesa and Krueger (2006) or Heathcote et al. (2017) quantitatively characterize the optimal level of tax progressivity. Similarly, a large literature in public economics (recently surveyed by Piketty and Saez (2013)) has studied optimal non-linear tax schedule in models without intra-household inequality. A few papers have studied the effects of tax policy in two-earner models (e.g. Guner et al. (2011, 2012), Bick and Fuchs-Schündeln (2017)), but typically take intra-household allocations and marriage outcomes as given. Notable exceptions, where taxes are allowed to affect intra-household allocations, are Alesina et al. (2011) and Bastani (2013), who study gender-based taxation and thus a different policy than tax progressivity. Gayle and Shephard (2019) study the optimal taxation of couples in a static matching model, in which taxes can affect intra-household allocations, and focus on the optimal degree of jointness of the tax schedule. This chapter contributes by focusing on progressivity and the relative importance of inequality within and across households. Thus, this chapter sheds additional light on the role of intra-household inequality. In addition, I analyze a dynamic marriage market, which leads to different mechanisms relative to a static matching model as in Gayle and Shephard (2019).

My research is further related to papers which have used relatively similar frameworks - dynamic bargaining models - to study different aspects of the tax and transfer system, such as for how long individuals are eligible to receive social assistance benefits (Low et al. (2018)), the EITC (Mazzocco et al. (2013)) or a comparison of joint and individual taxation (Bronson and Mazzocco (2019)). Relative to these papers, this chapter adds the marriage market equilibrium, so that the distributions of potential partners are endogenously determined in equilibrium. Equilibrium models of marriage and intra-household allocations are still relatively rare (see e.g. Shephard (2019), Ciscato (2019), Goussé et al. (2017), Chiappori et al. (2018) or Reynoso (2019) for some recent work). Chade and Ventura (2002, 2005) study the impact of the differential tax treatment between singles and couples on the marriage market. In my analysis, the tax code is formally neutral with respect to marriage, due to individual taxation, but tax progressivity affects marriage rates by reducing wage inequality. The focus of this chapter on intra-household bargaining is also related to Knowles (2012), who studies the role of bargaining for explaining the time trend in relative leisure between men and women. Finally, several papers have highlighted the role of the marriage market for inequality (e.g. Greenwood et al. (2016), Fernandez et al. (2005) or Fernández and Rogerson (2001)). These papers abstract from bargaining and intra-household inequality, and thus focus on different aspects of the marriage market. In addition, they do not consider the impact of policies.

In the empirical literature, a growing number of studies has investigated the role of intra-household inequality. Lise and Yamada (2018) analyze Japanese panel data and relate allocations to differences in relative wages. Santaeulàlia-Llopis and Zheng (2017) study

Chinese data and find that standard equivalence scales understate inequality in food consumption substantially and emphasize the role of luxury goods. Other papers have estimated collective household models to infer private consumption from micro-data on allocations (e.g. Lise and Seitz (2011), Cherchye et al. (2012, 2015, 2018)). Lise and Seitz (2011) conclude that intra-household inequality accounts for about 25% of consumption inequality in recent years and that it is crucial for the assessment of the time trend. In addition, empirical studies that test unitary and collective household models often reject unitary models (e.g. Attanasio and Lechene (2014), Lundberg et al. (1997)) and find that intra-household allocations react to changes in outside options, which is consistent with models of bargaining.

The rest of this chapter is organized as follows. Section 2.2 first shows stylized facts about intra-household allocations. Section 2.3 describes the model, which is calibrated in section 2.4. Section 2.5 shows the details of the policy simulations and discusses the results. Section 2.6 concludes.

2.2 Motivating Facts

This section briefly describes stylized facts about intra-household allocations. The main data set being used is the "Time Use and Consumption" module of the Dutch LISS panel (see Cherchye et al. (2012) or Cherchye et al. (2017) for a detailed description). The module contains relatively detailed questions about household expenditures. Individuals are asked about their private consumption as well as about overall spending for different categories and kids. The survey questions refer to personal expenditure for food, clothing, cigarettes, leisure time expenditure and a few other categories. At the same time, only a part of the overall consumption expenditure is directly assignable to a household member; the larger part is non-assignable (containing for example trips, housing expenditure or utility payments). I follow Cherchye et al. (2012) in their classification of non-assignable consumption. In practice, then, it is assumed that a fraction of non-assignable consumption is public.

In the following, I highlight several aspects which motivate some of the model assumptions. First, panel (a) of figure shows the ratio of the wife's private (assignable) consumption relative to total private assignable consumption. This figure suggests that there are quite many couples in which spouses consume different amounts of private goods, which supports the notion of intra-household inequality in consumption.[4] In a within-between decomposition of the variance of private consumption, the within-couple component has 51.1% the size of the between-couple component.[5]

Second, panel (b) shows the distribution of the leisure hours of the wife relative to total

[4] Note that consumption surveys likely contain some measurement error. While the LISS panel asks individuals about their expenditure for an average month, a more detailed diary which keeps track of infrequent purchases would be useful. As a result, there might be some measurement error resulting from individuals incorrectly estimating their expenditure.

[5] The within-between decompositions are described in more detail in section 2.5.2.

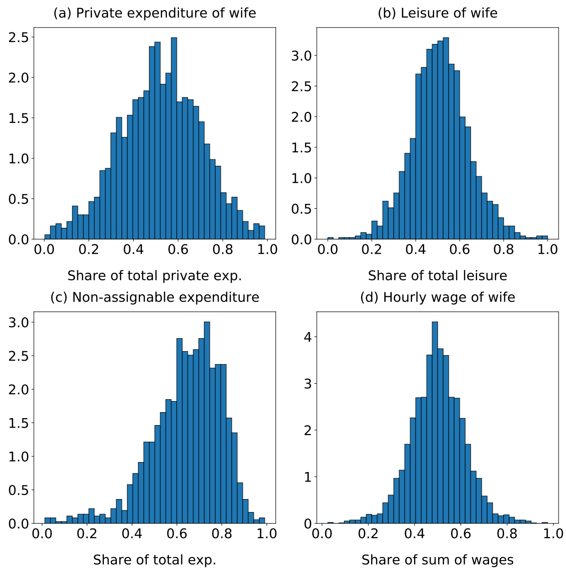


Figure 2.1: Distributions of consumption, leisure and wages

Notes: Panel (a) shows the share of private assignable consumption of the wife relative to total (assignable) private consumption in the household. Panel (b) and (d) do the same for hourly wage rate. Panel (c) depicts the share of expenditure used for non-assignable consumption. The data comes from the LISS panel.

leisure of the household. The figure indicates that it is often the case that one spouse has more leisure than their partner, which suggests that the allocation of leisure can be an important dimension of intra-household inequality.

Third, an important question regarding the potential for within-household inequality in consumption is the role of public goods. Panel (c) depicts the ratio of non-assignable expenditure, i.e. expenditure that cannot be attributed to the personal consumption of either spouse in the survey, relative to total expenditure and shows that it makes up for a relatively large fraction of consumption in many households (for example due to rent payments). This indicates that accounting for the publicness of some consumption is important.[6]

Finally, panel (d) shows the ratio of the hourly wage of the wife relative to the sum of wages. The figure is important since one would expect within-household inequality to be most relevant when spouses strongly differ in their earnings potential. Thus, the relative importance of within- and between household inequality should strongly depend on how assortative matching is in terms of wages. The figure provides an indication of how many couples there are in which either of the spouses has an economically stronger position. The focus is on *hourly wages*, rather than on income, since wages are a better indicator of the earnings *potential* of an individual, especially for married women. The data indicates that

[6]Note that classifying consumption as either private or public is not straightforwad and ultimately requires making assumptions about preferences. For example, one could regard a car as a public good, since it can to some extent be used by both spouses, or as a private good, if only one spouse cares about having a car. In addition, spouses could disagree over *which* car to buy, so that a given purchase, while technically being used by both spouses, is partly 'private' in the sense that it fulfils the preferences of one spouse more than the other.

dispersion in wages is relatively large. In a within-between decomposition of hourly wages, the within-household component accounts for about 40% of the total variation. Interestingly, the distribution is relatively symmetric, in the sense that there are also many couples in which the wife has a higher wage than the husband.

Lise and Yamada (2018) report similar graphs for Japan. The main difference between the Netherlands and Japan is that the gender difference is much less pronounced in the Netherlands, whereas Japan has a large gender wage gap and the mean share of private consumption of women is considerably lower. In the Netherlands, by contrast, there is also a noticeable fraction of couples in which the woman has the higher wage rate. Earnings also exhibit a strong gender difference in the Netherlands, since part-time employment is very common for women.

2.3 Model

2.3.1 Overview

The model is an equilibrium search model of marriage, divorce, labor supply and savings. In each period, a new cohort of men and women is born and lives for a finite time horizon T. Each model period corresponds to two years and the terminal period is $T = 30$. Individuals participate in the labor market until retirement age $T_R = 21$. Individuals are heterogeneous in their labor market ability and face shocks to their earnings potential over their lifetime. They decide on consumption and savings, as well as on splitting their time use between market work, home production and leisure. There are two consumption goods. One is private and the other one is public within the household and consumed equally by both spouses, in case an individual is married. The home good, which is produced using the time inputs of spouses, is also public within the household. The main features of the model are that it allows for within-household inequality in personal consumption and leisure, and that marriage and divorce decisions are endogenous.

Individuals start their lives single and can find a spouse on a search-based marriage market. In each period, singles may meet a potential partner from the distribution of available singles. Singles only meet potential partners who are from the same cohort and thus have the same age. At the time of a meeting, individuals take the economic characteristics of their potential partner and a non-economic 'love' component into account. Marriage takes place if both individuals prefer getting married over staying single for another period. Couples exogenously have children according to a stochastic process. The economic benefits from marriage are economies of scale in consumption and home production, since the public consumption good and home production time are public goods within the household, and the possibility to share labor market risk. However, not all goods are shared equally by both spouses. Personal consumption and leisure are consumed privately.

Decision-making in marriage is modeled as a limited commitment contract (Mazzocco et al. (2013)). This means that couples decide cooperatively on allocations. The Pareto weight of a spouse determines the weight in the objective function of the household and thus how favorable the allocation is towards this spouse. The Pareto weight is initially determined by Nash bargaining in the period when the couple gets married. The *outside option* in this bargaining situation is the value of staying single and meeting another potential partner in the next period. After marriage, the Pareto weight can be renegotiated when one partner would prefer to divorce given the current Pareto weight. The non-economic match quality of the couple can deteriorate over time, which, together with the wage shocks, may lead to divorce. Divorced individuals re-enter the marriage market and can remarry in future periods.

The distributions of potential partners that singles can meet are endogenously determined in equilibrium. For each age group, these distributions depends on which individuals got married or divorced in previous periods and thus leave or reenter the pool of available singles. In a steady-state equilibrium, individuals have correct expectations over the future distributions of singles. Thus, finding an equilibrium corresponds to finding a fixed point between the expectations of agents and the implied distributions of singles.

In the following, the model will be described in more detail. Section 2.3.8 then shows illustrations about how intra-household allocations are determined and discusses how this is affected by the tax system.

2.3.2 Preferences and Home Production

Individuals differ in gender $g \in \{f, m\}$ and have preferences over consumption, leisure and a home good, which is produced with time as an input. There is a private and a public consumption good and both goods are available to singles and married individuals. The public good C is consumed equally by both spouses when an individual is married, as opposed to the private consumption good c. The home good is denoted as D and is a public good within the household. Individuals have a total time budget of 1, which can be divided between market work (h_g), domestic work (d_g) and leisure (l_g):

$$h_g + d_g + l_g = 1$$

The per-period utility function from consumption, leisure and home production is given by:

$$u(c, C, l, D) = \alpha_c \frac{c^{1-\gamma}}{1-\gamma} + (1-\alpha_c) \frac{C^{1-\gamma}}{1-\gamma} + \alpha_l \frac{l^{1-\gamma}}{1-\gamma} + \alpha_d(b) \frac{D^{1-\gamma}}{1-\gamma}$$

Married individuals further get utility from the match quality of their marriage which will be discussed in more detail in section 2.3.5. The preference parameter for the home good D is allowed to depend on the presence of kids (b), which captures that couples with children typically spend more time on home production. The home good is produced according to the following home production technology, that takes the domestic work time of each spouse as the inputs.[7]

$$D = \begin{cases} \eta d_g & \text{when single} \\ (\eta_f(b) d_f^z + \eta_m(b) d_m^z)^{\frac{1}{z}} & \text{when married} \end{cases}$$

The parameter z controls the substitutability between the two time inputs. I assume that the productivity parameter of home time is the same for single men an single women, but allowed to differ between genders in marriage, depending on whether there are children ($\eta_g(b)$). When $\eta_f(b) > \eta_m(b)$, married women spend more time in home production than men do, even in couples where both partners have the same wage rate. This captures the empirical observation that women take a disproportionate share of total home production time. The specification could for example represent gender norms and does not necessarily reflect actual differences in 'productivity'.

The productivity parameter for married men is set to $\eta_m(b) = 1 - \eta_f(b)$. Individuals discount the future at rate β and have access to a risk-free asset that yields an exogenous interest rate R.

2.3.3 The Labor Market and the Government

Labor supply choices belong to a discrete set of hours $\mathbb{H} = \{h_1, ..., h_k\}$. The (log) wage of an individual is given by:

$$\log w_{i,t} = \kappa_i + \epsilon_{i,t} + \kappa_g^G + u_{it}$$

κ_i represents the ability of the individual. κ_g^G refers to the gender wage gap.[8] $\epsilon_{i,t}$ is a Markov chain that represents *persistent* changes in earnings potential over the life-cycle. u_{it} is a *temporary* (i.i.d.) shock and follows a normal distribution with mean 0 and variance σ_T.[9] In retirement, individuals get a fixed replacement rate of their full-time earnings.

The government collects revenue through a tax on labor income. The tax system is represented

[7]The advantage of this specification of preferences and home production is that is allows for closed form solutions for consumption (c_f, c_m and C) conditional on total household expenditure and the Pareto weight, which makes the model more tractable. It would be an interesting extension to consider a home production function where the public good enters as a substitute to time (as in Knowles (2012) or Lise and Yamada (2018)). For example, the interpretation could be that households can partially outsource their domestic work. However, it would increase the computational burden. Similarly, allowing for non-separability between consumption and leisure would be an interesting extension.

[8]Note that the model abstracts from human capital accumulation/learning by doing. This was included in an earlier version of the paper, but replaced by a richer distribution of permanent types to model the income distribution and assortative mating in a more realistic way, which is important for the study of inequality. The current specification can be thought of as frontloading the earnings potential of an individual into the ability type. Adding learning-by-doing on top of the heterogeneity and persistent earnings risk would lead to a very large state space.

[9]The distinction between temporary and transitory shocks is often made in the literature (see e.g. Krueger and Perri (2006)). In addition, the temporary shock is helpful for computational reasons, since it makes the problem more smooth.

by a parametric tax function, using the functional form from Heathcote et al. (2017):

$$T(y) = \max\Big((1 - \psi_1 y^{-\psi_2}) \cdot y, 0\Big)$$

This tax function implies that the after-tax income (conditional on being positive) is given by:

$$y - T(y) = \psi_1 y^{1-\psi_2}$$

The parameter ψ_1 determines the level of the tax, since it proportionally shifts the after-tax income. ψ_2 determines the progressivity of the tax system. Progressivity can be defined as the average tax rate ($\frac{T(y)}{y}$) being increasing in income and is commonly measured by the progressivity tax wedge (see e.g. Holter et al. (2019) for details). When y_1 and y_2 ($y_2 > y_1$) are both in the region of positive tax payments, this can be computed as follows:

$$PW(y_1, y_2) = 1 - \frac{1 - T'(y_2)}{1 - T'(y_1)} = 1 - \left(\frac{y_1}{y_2}\right)^{\psi_2}$$

Thus, given the functional form for the tax function, the progressivity wedge only depends on ψ_2. Married couples are taxed individually in the Netherlands, so that the tax liability is calculated for each spouse separately.[10] For simplicity, this specification abstracts from aspects like transferable deductions between spouses, which would introduce a small degree of dependence between partners.[11] Note that the tax function $T(y)$ is restricted to be non-negative. This has the interpretation that $T(y)$ models only the tax and not the transfer system.[12]

Transfers are modeled as a simple income floor depending on *family income*, which prevents households from experiencing very low consumption levels. Households with an income below B_{min} receive a transfer to reach this level, that is financed out of tax revenue. The budget constraint of the government equates the expected life-time revenue from a newly born cohort to an exogenous spending requirement $\bar{G}$:

$$\text{Rev}(\psi_1, \psi_2) = \mathbb{E}\Big[\sum_t \beta^t \Big((T(y_{i,t}) - B(y_{i,t})\Big)\Big] = \bar{G}$$

Here, B indicates whether an individual gets the social assistance benefit. The revenue $\bar{G}$ is used to finance non-valued government expenditure.[13]

[10]The actual Dutch tax schedule consists of several steps, within which the marginal tax rate is constant. Approximating the schedule with a tax function with few parameters makes policy experiments more straightforward. In the calibration section, I will discuss the choice of parameters and how it compares to the actual schedule in more detail.

[11]See de Boer et al. (2018) for a detailed overview of the institutional setting in the Netherlands.

[12]Excluding transfers from $T(y)$ is required since $T(y)$ is applied to the *individual* labor income of spouses, whereas transfer are typically computed on the basis of *family* income.

[13]This can also be interpreted as the government producing an economy-wide public good P, which provides each individual with an additive utility $u^P(P)$ and does not affect choices. An alternative assumption would be to rebate the tax revenue to workers, which would provide an additional minimum income.

Note that in addition to all family-related channels, the model features the standard policy trade-off of progressive taxation (as in e.g. Conesa and Krueger (2006) or Heathcote et al. (2017)). An increase in progressivity (ψ_2) is potentially welfare-improving because it allows lowering tax rates for lower incomes (through ψ_1) while keeping the government budget constant. On the one hand, this can be desirable since it redistributes to low-income individuals who have a higher marginal utility from income. In addition, a budget-balanced increase in progressivity provides more insurance against wage fluctuations. On the other hand, the increase in progressivity can also reduce welfare because of the labor supply disincentives through high marginal tax rates.

2.3.4 Singles

Singles decide on consumption, savings and the time spent for home production, leisure and labor supply. They are characterized by their gender, ability, the productivity shock, the presence of children, if there are any from a previous marriage, and assets. Thus, the state vector of a single is:

$$\omega^S_{i,t} = (g, \alpha_{i,t}, \epsilon_{i,t}, b_{i,t}, A_{i,t})$$

In the beginning of each period, individuals observe the realization of the persistent productivity shock and whether the children grow up. Their decision problem is described by the following value function and budget constraint:

$$\begin{aligned} V^S_{i,t} &= \max_{c,C,h,l,d,A'} \left\{ u(c, C, l, D) + \beta \mathrm{EV}^S_{i,t}(A') \right\} \\ c + C + A' &= w_{i,t}h - T(w_{i,t}h) + RA_{i,t} \\ d + h + l &= 1 \end{aligned}$$

The current asset level is denoted as $A_{i,t}$, while A' is the amount of savings for the next period. The continuation value $\mathrm{EV}^S_{i,t}(A')$ includes the expected utility from being single in the future, as well as from future marriage, divorce and remarriage. State variables, or objects that depend on state variables, are indexed by the individual and period.

In addition to their ability, the initial condition of singles also contains a realization of the productivity shock. To ensure that the initial draw is consistent with the earnings process, it is obtained by drawing from the distribution of the process after two periods, starting with the median wage shock. In addition, individuals draw a small initial amount of assets, that could e.g. be interpreted as a parental transfer. The assets are drawn from a uniform distribution between 0 and 10% of the full-time earnings of the individuals.[14]

[14] The initial asset endowment is useful for computational reasons, since it makes the (equilibrium) asset distributions in the first periods more smooth.

2.3.5 Couples

Before turning to matching and marriage decisions, it is useful to discuss the decision problem of married couples and, in particular, the role of the Pareto weight in more detail. Married couples decide on consumption, labor supply and home production time. The utility of a married individual ($g \in \{f, m\}$), conditional on these choices is:

$$U^g_{i,t}(c, C, l, D, A') = u(c, C, l, D) + \theta_{i,t} + \beta \mathrm{EU}^g_{i,t}(\lambda_{i,t}, A')$$

On top of the economic utility $u(c, C, l, D)$, married individuals also experience the additive 'love shock' $\theta_{i,t}$, which represents the non-economic quality of the current relationship. Match quality evolves stochastically according to an $AR(1)$ process with persistence ρ_θ and variance σ_θ:

$$\theta_{i,t} = \rho_\theta \theta_{i,t-1} + \epsilon^\theta_{i,t} + \bar{\theta}_{\alpha_f,\alpha_m}$$

$\epsilon^\theta_{i,t}$ represents the innovation of the match quality process.[15] $\bar{\theta}_{\alpha_i,\alpha_{j(i)}}$ is a preference term that generates additional utility depending on the combination of ability types (α_f and α_m) of the couple. In practice, I assume that

$$\bar{\theta}_{\alpha_f,\alpha_m} = \begin{cases} \bar{\theta} & \alpha_f \text{ and } \alpha_m \text{ in same ability quartile} \\ 0 & \text{else} \end{cases}$$

The interpretation of this specification is that individuals enjoy additional utility when they are married to someone with a broadly similar ability type, for example because of shared interests or social conventions. Thus, $\bar{\theta}$ will be referred to as the homophily parameter. This helps the model to generate a realistic degree of assortative mating.[16]

$\mathrm{EU}_{i,t}(\lambda_{i,t}, A')$ is the continuation value of marriage that includes the possibility of divorce in the next period. Married couples exogenously have children with a probability $p^b(t)$, that declines with age and is zero once the couple reaches the age of 40. Having children is a binary state ($b \in \{0, 1\}$), which can be thought of as all couples having two kids. Children leave the household with probability $p^{b,g}$.[17] The state space of married couples includes the current Pareto weight λ, the ability level (α), and productivity shock (ϵ) of each spouse as well as the assets, the current level of match quality (θ) and the presence of children (b):

$$\Omega^C = \{(\lambda, \alpha_f, \alpha_m, \epsilon_f, \epsilon_m, A, \theta, b)\}$$

[15] In practice, the match quality process is approximated with a discrete process using the Rouwenhorst method.

[16] This is comparable to Greenwood et al. (2016), who have a similar preference term for college and non-college individuals. In my context, some alternative speficiations would be possible, for example to let the preference term depend more continuously on the type difference, for example as some function of $|\alpha_f - \alpha_m|$.

[17] This assumption avoids keeping track of the age of kids, which would greatly enlarge the state space.

The most important variable that determines the allocation within the couple is the Pareto weight $\lambda_{i,t}$. The weights are normalized such that $\lambda_{i,t}$ is the weight of the wife and $1-\lambda_{i,t}$ is the weight of the husband. The Pareto weight is endogenously determined. For the moment, focus on a couple with a given weight Pareto weight $\lambda_{i,t}$, which was obtained in a previous period. The household problem is to maximize the weighted sum of utilities of wife and husband, subject to the budget and time constraints:

$$\begin{aligned}
\max_{\mathscr{C}=(c_f,c_m,C,A',h_f,h_m,d_f,d_m,l_f,l_m)} & \quad \lambda_{i,t}U^f_{i,t}(\mathscr{C})+(1-\lambda_{i,t})U^m_{i,t}(\mathscr{C}) \\
c_f+c_m+C+A' &= w_{f,t}h_f+w_{m,t}h_m-T(w_{f,t}h_f)-T(w_{m,t}h_m)+RA_{i,t} \\
h_f+d_f+l_f &= 1 \\
h_m+d_m+l_m &= 1
\end{aligned}$$

Thus, a high value of $\lambda_{i,t}$ corresponds to a high weight placed on the utility of the wife and a large amount of private consumption and leisure for her, whereas a low value of $\lambda_{i,t}$ leads to high utility for the husband. The utility levels evaluated for the optimal choices are denoted as $U^*_{f,t}(\lambda_{i,t},\omega^M_{i,t})$ and $U^*_{m,t}(\lambda_{i,t},\omega^M_{i,t})$, given λ and the state variables of the couple ω_c.

Since spouses can unilaterally file for divorce in the beginning of a period, allocations have to make sure each spouse prefers staying married over divorcing. If this is not the case, the Pareto weight can be adjusted to ensure that the spouse who would want to divorce stays in marriage (limited commitment).[18] The participation constraint requires that each spouse is better off in marriage than as a single:

$$U^*_{g,t}(\lambda_{i,t},\omega_c)\geq V^S_{g,t}\Big(A_{i,t}/2,\alpha_{g,t},\epsilon_{g,t},b_{i,t}\Big)$$

Assets are split equally in divorce and children stay with the mother, who also takes them into the next marriage.[19] The Pareto weight of the couple stays constant until a participation constraint is violated. In case one spouse would want to divorce, it may be possible to increase their Pareto weight to make them stay in marriage. If this is feasible and the other spouse also still prefers staying in marriage, the Pareto weight is adjusted in favor of the spouse who wants to leave otherwise. Efficiency requires making the *minimal* adjustment relative to the previous Pareto weight, so that the spouse is just indifferent between divorcing and staying in the marriage. If such an adjustment is not possible, divorce occurs.

The interpretation of limited commitment is that spouses behave as under a constrained efficient contract, given the constraint that the contract can be renegotiated if the partic-

[18] See Chiappori and Mazzocco (2017) for a theoretical overview of limited commitment models.

[19] The equal division of assets is a simplification that is often made in the literature (see e.g. Fernandez and Wong (2017); Mazzocco et al. (2013)) and close to the legal regime in the Netherlands. Importantly, this assumption rules out that couples can write prenuptial contracts to e.g. keep separate property, which is rare in practice. Note that allowing couples to keep separate property would *increase* the importance of bargaining, since then the spouse with the higher bargaining power would be able to keep more property in their name, increasing the share of assets in divorce.

ipation constraint of a spouse is violated in a future period. As a result, risk-sharing in marriage is imperfect. Shocks to wages and match quality can trigger changes in the Pareto weight of the couple. This is the more likely the lower the match quality of the couple, since rebargaining requires that one individual would prefer to get divorced given the current bargaining weight, which is less likely when match quality is high.[20]

2.3.6 Meetings and Marriage

Singles participate in the marriage market with probability $m(t)$. This probability declines with age to capture that it is harder to meet potential partners later in life, when fewer people are single.[21] Singles who enter the marriage market randomly meet a potential partner from the (endogenous) pool of available singles. For tractability, I assume that singles only meet others from the same cohort, so that spouses have the same age.[22] The characteristics of the potential spouse are their ability, productivity, assets and children. The probability of meeting a spouse with a state vector (ϵ, α, A, b) is given by:

$$M_{t,g}(\epsilon, \alpha, A, b) = m(t) \cdot \Lambda_{t,g}(\epsilon, \alpha, A, b)$$

$\Lambda_{t,g}$ reflects the likelihood of meeting a potential spouse with certain characteristics, given the distributions of singles at each age. $\Lambda_{t,g}$ is endogenously determined in equilibrium, since it depends on which individuals got married or divorced at previous younger ages and on how much they saved and worked. The characteristics of each potential spouse are denoted as $\omega_{i,t}$. In addition, an initial realization of match quality $\theta_{i,t}$ is drawn from a $N(\mu_{I,\sigma}, \sigma_{I,\theta})$ distribution for each match, which captures the non-economic quality of the potential marriage.

At the time of a meeting, have to decide whether to get married and, if so, set an initial value for the Pareto weight $\lambda_{i,t}$. They observe the match quality draw and the persistent characteristics of the potential partner. For technical reasons, it is helpful to assume that the temporary wage shocks are realized *after* individuals have decided on whether to get married. This ensures that the *expected* utility from marriage $\mathbb{E}_w V^M_{g,t}(\lambda, \theta_{i,t}, \omega_{f,t}, \omega_{m,t})$ is a smooth function of the Pareto weight, even though labor supply choices are discrete.[23] The expectation $\mathbb{E}_w$ refers to the expectation over the temporary wage shocks. Note that the temporary shocks are i.i.d., so that the timing assumption should have a small impact on

[20]Evidence from Lise and Yamada (2018) supports these models of decision-making, as they find that decision power changes infrequently and more often before divorce and for big shocks.

[21]A richer specification would be to model $M(t)$ endogenously as a function of the fraction of singles at each age group, as in Guvenen and Rendall (2015).

[22]Otherwise, one would need to keep track of the age of both spouses, which would greatly enlarge the state space.

[23]The smoothness of the Pareto frontier is helpful when computing the bargaining solutions.

marriage decisions or bargaining. Importantly, individuals observe the *persistent* wage shock of the potential partner before making the marriage decision.

When deciding to get married, individuals compare the value from staying single for one more period and drawing another potential partner in the next period ($\mathbb{E}_w V^S_{g,t}$) to the value of getting married to the current potential partner ($\mathbb{E}_w V^M_{g,t}$). Given a value for the Pareto weight λ, the *gain from marriage* of individual g is:

$$G_g(\lambda, \omega_{f,t}, \omega_{m,t}, \theta_{i,t}) = \mathbb{E}_w V^M_{g,t}(\lambda, \omega_{f,t}, \omega_{m,t}, \theta_{i,t}) - \mathbb{E}_w V^S_{g,t}(\omega_{g,t})$$

For marriage to take place, there must be a Pareto weight $\lambda \in (0,1)$ such that both individuals prefer marriage over staying single:

$$G_g(\lambda, \omega_{f,t}, \omega_{m,t}) \geq 0 \qquad \forall\ g \in \{f, m\}$$

If there is no Pareto weight such that both individuals prefer marriage, the match is rejected and the individuals stay single for another period. In other cases, there is an interval of weights $[\underline{\lambda}, \bar{\lambda}] \subset (0,1)$ for which both individuals prefer marriage over staying single. The Pareto weight also has the interpretation of a transfer, since spouses can reduce their own utility in order to increase the utility of their partner.[24] The Pareto weight is determined by Nash bargaining, which chooses the weight λ which maximizes the Nash product:[25]

$$\tilde{\lambda} = \text{argmax}_\lambda G_f(\lambda, \omega_{f,t}, \omega_{m,t}, \theta_{i,t}) \cdot G_m(\lambda, \omega_{f,t}, \omega_{m,t}, \theta_{i,t})$$

Note that this is the most essential model ingredient that determines intra-household allocations. Section 2.3.8 provides a more intuitive discussion of how Pareto weights are determined and how this can be influenced by the tax system. Nash bargaining relates the allocation in marriage to the relative 'outside options' of each individual. The outside option is the utility of staying single for another period, as previously defined:

$$V^S_{i,t} \;\; = \;\; \max_{c,C,h,l,d,A'} \left\{ u(c, C, l, D) + \beta \text{EV}^S_{i,t}(A') \right\}$$

With the notation from this section, we can state the continuation value of singlehood as the expected utility from future periods of singlehood and the expectation over all potential marriages:

$$\text{EV}^S_{g,t}(A') = \int_{\omega^M = (\omega_{f,t+1}, \omega_{m,t+1}, \theta)} M(\omega^M_{t+1}) \cdot \mathbb{E}_w V^m_{g,t+1}(\lambda(\omega^M_{t+1})) + (1 - M(\omega^M_{t+1})) \cdot \mathbb{E}_w V^S_{g,t+1} \ \text{d}F(\omega^M_{t+1})$$

[24]In technical terms, utility is *imperfectly transferable*, since giving up one util does not increase the utility of the partner by an equal amount. See Galichon et al. (2018) for a theoretical discussion.

[25]Note that the general formulation of Nash bargaining contains a parameter for 'bargaining strength', which corresponds to the exponents of the Nash product. Here, it is assumed that this parameter is 0.5.

ω_t^M is a vector containing the characteristics of both individuals and the match quality realization. $M(\omega_t^M)$ is an indicator variable that takes the value of 1 if marriage occurs given these match characteristics and $\lambda(\omega_t^M)$ is the Pareto weight that is chosen in this case. In particular, the expression for $\mathrm{EV}_{g,t}^S(A')$ highlights that outside options are determined both by the living standard while single and by the full distribution of potential marriages in the future.

2.3.7 Marriage Market Equilibrium

2.3.7.1 Definition of Equilibrium

A stationary equilibrium consists of distributions of singles, policy functions for singles and couples and matching rules such that

1. the policy functions $(A, c, C, h, l, d) = P_{g,t}^S(\omega^S)$ solve the problem of singles

2. the policy functions $(A, c_f, c_m, C, h_f, h_m, l_f, l_m, d_f, d_m) = P_t^M(\omega^M)$ solve the problem of married couples

3. separation and rebargaining $(D, \tilde{\lambda})$ occur according to the limited commitment procedure

4. the matching rule (m, λ) satisfies the participation constraints and the bargaining solution, where m is an indicator for getting married and λ the initial Pareto weight

5. the implied distributions of singles, $\Lambda_{t,g}(\epsilon, H, \alpha, A)$, are consistent with the distributions that are used to determine the optimal choices and value functions from (1) - (4)

2.3.7.2 Discussion and Computation

The main concern of the marriage market equilibrium is to find distributions of singles of each gender and age group $(\Lambda_{t,g}(\epsilon, \alpha, A, b))$. Since these distributions determine which potential partners an individual can meet at each age, they partly determine the forward-looking decisions about marriage, divorce and savings. At the same time, these decisions feed back into the distributions of available singles at each age. Thus, equilibria need to be computed via fixed-point iteration. Starting with a guess for the distributions of singles at each age, one can solve the life-cycle problem recursively to determine the value functions. Then, the actual distributions of singles, given the guess of the distribution, are computed from a simulated panel of 1 million individuals and the guess is updated. This is repeated until convergence. Since there are relatively many state and choice variables, computing the equilibrium is fairly time-consuming - the parametrization used for the policy experiments runs on a cluster using 448 cores.[26] The computational details are described in the appendix.

2.3.8 How does Tax Progressivity affect Intra-Household Inequality?

To illustrate the mechanisms of the model, this section discusses an example to show how Pareto weights are determined and highlights the aspects that are relevant for the effects of progressivity. First, consider a simple static example, in which the economy consists of two individuals. Individuals consume the public and the private good and labor supply is fixed. The utility function over private and public consumption is given by:

$$u(c, C) = \alpha_c \frac{c^{1-\gamma}}{1-\gamma} + (1-\alpha_c)\frac{C^{1-\gamma}}{1-\gamma}$$

Figure 2.2 illustrates the bargaining situation. The individuals need to determine the weight of the woman λ, the weight of the man being $1-\lambda$. The decision weight determines consumption and leisure in marriage. Each value of λ corresponds to a point on the Pareto frontier, which is a combination of utility levels for the individuals. The higher λ, the higher the utility level of the woman. The black (vertical) line is the value of singlehood for the woman and the green (horizontal) line is the one for the man. In the graph on the left side, wages are unequal and the value of being single is lower for the woman than for the man. The final allocation must lead to each utility level being higher than the value of singlehood. This is the case for a range of values of the Pareto weight, which corresponds to the line segment between the points where the Pareto frontier intersects the values of singlehood. Whenever this is the case, marriage takes place. The bargaining solution picks one particular point on the Pareto frontier, which is marked by the dot. In the case of unequal wages, the Pareto weight is more favorable for the man, who receives more consumption, reflecting the difference in outside options.

The graph on the right side shows the case where wages are equal. As a result, the value of being single is now equal for the two individuals. The new allocation is marked by the dot and assigns equal utility in marriage to each partner. In this example, total household income and thus the Pareto frontier were unchanged. Such bargaining effects are in line with a range of empirical studies that tests unitary and collective household models and finds that allocations react to changes in outside options (e.g. Attanasio and Lechene (2014), Lundberg et al. (1997)).

In addition to the outside options, intra-household allocations are also determined by the shape and location of the Pareto frontier (which is unchanged in the example in figure 2.2).

[26] The state space for married couples includes 48960000 elements (the match quality realization, the permanent wage shocks and ability types of each spouse, whether there are children, the Pareto weight and the asset level).

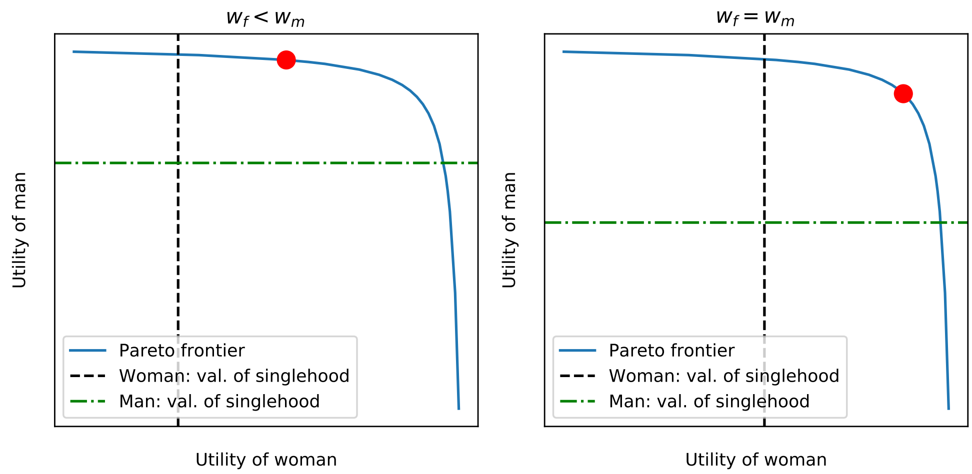


Figure 2.2: Illustration - Determination of Pareto weight

Notes: The graphs illustrate the impact of changing relative wages on the bargaining outcome in a simple example. The blue line is the Pareto frontier. Each point on the frontier corresponds to a Pareto weight. The dot shows the bargaining outcome given the outside options.

With Nash bargaining, the initial Pareto weight is chosen such that the product of the gains from marriage is maximized.[27] The main determinants of the Pareto weight are best illustrated by the first-order condition of the bargaining solution:

$$\frac{V_{f,t}^M(\lambda) - V_{f,t}^S}{V_{m,t}^M(\lambda) - V_{m,t}^S} = \frac{\partial V_{m,t}^M}{\partial V_{f,t}^M}(\lambda)$$

The terms $V_{g,t}^M(\lambda) - V_{g,t}^S$ represent the gains from marriage for each individual $g \in \{f, m\}$, given a Pareto weight λ. Allocations are determined by the relative gains from marriage of each individual and the slope of the Pareto frontier, which indicates how transferable utility is between spouses. When both individuals have a relatively high gain from marriage irrespective of the Pareto weight, which is for example the case when the match quality of the marriage $(\theta_{i,t})$ is high, the left-hand side of the equation is close to unity. This shifts the Pareto weight towards 0.5, where the slope of the frontier is also close to unity. Besides public consumption, home production and match quality, the gains from marriage also include risk-sharing in this model, which affects the Pareto weight.

Influencing the intra-household decision weight can improve utilitarian welfare, since the

[27]Note that there are other bargaining solutions. A common alternative to Nash bargaining is the 'Egalitarian' bargaining solution (see Kalai (1977) or Knowles (2012)), which chooses the Pareto weight to equalize the gains from marriage (the surplus) of each spouse. With linear utility, Nash and Kalai bargaining are equivalent. With concave utilities, Nash bargaining also takes the slope of the Pareto frontier into account. The main difference is that Kalai bargaining leads to a higher elasticity of the Pareto weight with respect to the utilities. Thus, with Kalai bargaining, there would be more intra-household inequality and a stronger reaction of Pareto weights to the tax system. As discussed in section 2.4.2, the implied elasticities of Nash bargaining seem quite in line with empirical evidence. A second difference is that utility terms which are shared equally by both spouses (such as the non-economic 'love' component $\theta_{i,t}$) do not influence the Kalai bargaining outcome, while the Nash outcome does depend on $\theta_{i,t}$.

gain of one person usually offsets the loss of the other (utility is imperfectly transferable). In the example discussed above, equalizing wages, such that the decision weight is set to 0.5 is the optimal policy for a utilitarian government that places equal weights on both individuals.

The discussion so far focused on static situations, in which the outside option is to stay single for the period. In the full model, the outside option for each individual is not to stay single for the rest of the lifetime, but includes the possibility of future (re-)marriage. As a result, outside options depend on the distributions of potential partners that will be available in the future, the probability of getting married in the future, and future Pareto weights. Progressivity affects outside options by equalizing living standards while single and also by changing future marriage market outcomes. Note that the marriage market equilibrium implies that the distributions of potential partners that individuals meet are endogenous and changes due to policy reforms. Changes in progressivity can lead to changes in selection into marriage, which is reflected by the equilibrium distributions of types. In addition, progressivity changes how much different individuals can save and therefore the asset distributions of singles.

Before turning to the calibration and results, I briefly discuss alternative policy instruments which could be used to reduce intra-household inequality. Since the source of intra-household inequality in this model is differences in wages (which make the values of single life unequal), tax progressivity is a fairly direct instrument to reduce these wage differences. To equalize the threat points of spouses, one could in principle also tax only singles very progressively. However, a highly progressive income tax for singles only would create large marriage bonuses and penalties and is less likely to be implemented in practice. At the very low end of the income distribution, the transfer system could also help to equalize bargaining positions. While this is a very interesting issue, the transfer system affects only a part of the population and the income tax would be more suitable to address intra-household inequality along the whole distribution.

2.4 Calibration

2.4.1 Calibrated Parameters

A number of parameters are set before the calibration routine. The coefficient of risk aversion is set to 1.5, which is a standard value in the literature. The discount factor is 0.9409, which implies a yearly discount factor of 0.97 (one period are two years). The interest rate is set to $R = 1/\beta = 1.063$. To set the preference parameter for the private consumption good, α_c, I use the fact that the preference specification implies a closed form solution for the share of public consumption of total expenditure, which only depends on the Pareto weight and the coefficient of risk aversion. Setting the preference parameter α_c to 0.277 implies that households spend roughly 61% of their expenditure on public goods.[28] Since wages are exogenous in the model, the wage process can also be set externally. The parameters that need to be set are the variance and persistence of the wage shock. This is done using data from the Dutch Socioeconomic Panel, due to its larger sample size. The wage process and the ability distribution are calibrated to match the cross-sectional dispersion of wages, as well as within-individual wage changes. Thus, in particular, the calibration targets the empirical degree of wage inequality. The fit of the wage moments and the corresponding parameters are shown in the appendix. The remaining parameters are calibrated to match a set of data moments. The calibrated parameters are the preference parameters for leisure and the home good, the variance and persistence of the match quality shock, the meeting rate in the last period and the homophily parameter for assortative matching.[29] Regarding the moments, I match the share of individuals that is currently married or cohabiting at age 20 and 36, the share that ever married or cohabiting by age 36, the share that ever experienced a divorce or separation by ages 36 and 42.[30] These moments are constructed based on the Dutch Kinship Survey, which provides retrospective data on life histories. In addition, I match women's average hours spent on domestic and market work. These moments are based on the LISS panel and included separately by the number of kids, since the preference for home production is allowed to differ for couples with kids. Finally, to get a realistic degree of assortative matching, I also target the share of the total variance of hourly wages that is due to the variation within households. This moment is primarily matched by the preference term for homophily. Intuitively, if couples were randomly matched, the share of the within-variation would be high since there would be many couples with unequal wages. Interestingly, this measure is relatively high in the data (0.4), suggesting that wage variation within couples is quite substantial.

The model is calibrated by minimizing the distance between model and data moments. Solving for the equilibrium distributions for each trial set of parameters would be computationally very expensive. Thus, I experimented with different approaches to reduce this computational burden. In practice, I start with a reasonable initial guess for the distributions, solve and simulate the model only once for each trial parameter vector, then update the distributions with a very small weight on the new distributions and proceed with the next trial parameter vector. In particular, this avoids solving for the full distributions for parameter vectors

[28]The exact share varies a little depending on the Pareto weight. The motivation for the data target is the procedure described in Cherchye et al. (2017): since only a part of the consumption expenditure is assignable in the LISS data, the authors assume that 50% of the remaining expenditure is public and the rest is private. With this assumption, I obtain 61% as the mean share of expenditure on public goods.

[29]Recall that the meeting rate is assumed to be linear, so that only the rate in the last period needs to be determined. The meeting rate in the first period is set to 1. The decline in the meeting rate helps the model to rationalize that the fraction of married individuals is lower than one at older ages.

[30]In the following, the term 'married' will refer to both married and cohabiting couples in the context of the data. Cohabitation is included in the data moments to accurately target the share of individuals in long-term partnerships, which would otherwise be understated. Note that there are some legal differences between married and cohabiting couples that are not captured by the model (relating e.g. to the division of assets upon divorce).

Table 2.1: Calibrated parameters

Parameter	Value
Scale of leisure, women (α_l)	1.20
Scale of home production, no kids ($\alpha_D(0)$)	0.21
Scale of home production, kids ($\alpha_D(1)$)	0.68
Home productivity, women (η_f)	0.62
Homophily term ($\bar{\theta}$)	0.73
Variance of match quality, married (σ_l)	1.29
Autocorrelation of match quality (ρ_l)	0.89
Mean of initial match quality ($\mu_{I,\theta}$)	-0.21
Variance of initial match quality ($\sigma_{I,\theta}$)	0.74
Meeting rate in last period	0.01

which would deliver a bad fit to the data anyway. After many iterations of the optimization algorithm, the distributions are close to the equilibrium distributions. The final parameters and distributions from this procedure are then used as starting points for the usual fixed point iteration, which converges after a few steps and delivers a good fit to the data. The calibrated parameter values are shown in table 2.1. Table 2.2 compares the model moments to the data moments that are targeted in the calibration.

2.4.2 Model Implications and Comparison to Data

Before turning to the results, this section first shows how the model compares to the data, in terms of inequality within and between households. Figure 2.3 shows the relative private consumption and leisure within couples, expressed as the ratio of female consumption and leisure relative to the sum of both household members. Thus, a value of 0.5 corresponds to both spouses having equal amounts. Relative consumption and leisure are untargeted in the calibration and are determined by the bargaining solution and the marriage market equilibrium. Both in the model and in the data, there is more dispersion in consumption than in leisure. In addition, the figure shows that the model generates less dispersion in consumption and leisure than present in the data, leaving room for alternative explanations of intra-household inequality (other than bargaining) that are not captured by the model, such as preference heterogeneity, which could be introduced as a residual.[31] To assess whether the bargaining solution gives empirically plausible outcomes, one can compare the model-implied

[31] Preference heterogeneity would be an interesting extension, partly because it could *increase* the potential for intra-household inequality: when the preferences of spouses differ, bargaining power becomes more important.

Table 2.2: Model fit

	Model	Data
Work hours, women, couples without kids	0.37	0.35
Work hours, women, couples with kids	0.20	0.20
Home hours, women, couples without kids	0.17	0.21
Home hours, women, couples with kids	0.42	0.48
Leisure, women, couples without kids	0.46	0.44
Leisure, women, couples with kids	0.39	0.32
Mean share of housework (women)	0.61	0.60
Currently married, age 20	0.25	0.24
Currently married, age 36	0.82	0.79
Ever married, age 36	0.93	0.92
Ever divorced, age 36	0.25	0.31
Ever divorced, age 42	0.32	0.31
Share of within-couple wage variance	0.37	0.40

Notes: This table summarizes the fit of the model. The time-use moments are expressed as the fraction of total non-sleeping time which are used for the corresponding activity.

elasticity of the Pareto weight with respect to relative wages to empirical estimates. Lise and Yamada (2018) estimate that on average increasing the difference in wages at the time of marriage by 10% translates into a 2.3% difference in the Pareto weight. In the model, the corresponding change in the Pareto weight is 1.9%, indicating that the implied elasticity of the Pareto weight seems reasonable and compares well to empirical studies.[32] In particular, this elasticity of the Pareto weight is an important statistic for policy analysis. Finally, the figure also contains the distribution of household income of couples in the model and in the data, indicating that the model does a reasonable job at capturing inequality across households.

[32]Note that the model-implied elasticity of the Pareto weight could be increased by using Kalai bargaining instead of Nash bargaining. However, in that case, the elasticity would be higher than the value found by Lise and Yamada (2018).

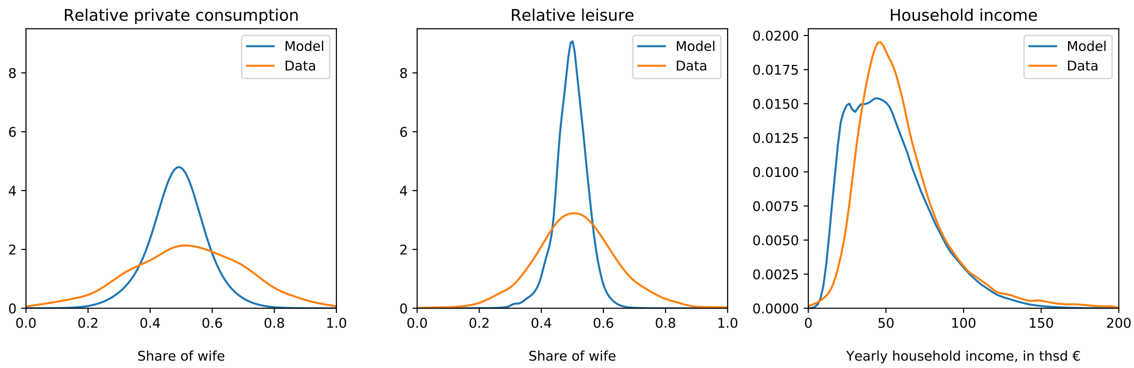


Figure 2.3: Inequality within and across couples

Notes: The figure shows the distribution of relative (private) consumption and leisure within households and the corresponding distributions in the data. Relative allocations are computed as the amount of the wife relative to the sum within the household (for example, $\frac{c_f}{c_f+c_m}$*). In addition, the figure shows the distribution of (gross) household income.*

To illustrate the Pareto weights underlying the intra-household allocations, table 2.3 shows the mean Pareto weight for each combination of ability group. Each group contains three ability types. The Pareto weights vary substantially across couples. In cases in which husband and wife are in the same ability group, the weights are relatively even on average. In unequal marriages, the individual with the better outside option gets a higher share of the surplus. For example, in couples which the wife is in the lowest ability group (L) and the husband in the highest (H), the mean Pareto weight is 0.26.

Table 2.3: Pareto weights by ability group of wife and husband

	Men - L	Men - LM	Men - UM	Men - H
Women - L	0.51	0.44	0.36	0.26
Women - LM	0.58	0.50	0.43	0.32
Women - UM	0.65	0.59	0.49	0.39
Women - H	0.74	0.68	0.60	0.48

Notes: The table shows the mean Pareto weights of couples across ability groups. ”L” is the lowest group, ”H” the highest. For example, group ”L” contains types 1, 2 and 3.

2.5 Results

In this section, I first use the calibrated model to analyze inequality given the current tax schedule in more detail. How much of inequality is due to inequality *within* couples and *across* couples? How much is due to singles and due to the difference in means between singles and married individuals? I then describe the policy experiment - an increase in tax progressivity - and show how the reform affects inequality through its impact on within-family allocations, marriage and divorce. To complement the analysis of (cross-sectional) *inequality*, I finally also turn to the effect of the reform on (ex-ante) *welfare*.

2.5.1 Inequality of What?

The model allows to go a step further than studying income inequality and consider inequality in consumption and utility more generally. Consumption inequality has been widely studied (see for instance Krueger and Perri (2006), Blundell and Preston (1998), Blundell et al. (2008)). As demonstrated by Lise and Seitz (2011), consumption inequality is potentially understated by ignoring within-household inequality and assuming that consumption is shared equally within the household instead.

To get a complete picture of inequality, one would ultimately be interested in comparing *utility* in the population, since this captures that individuals differ not only in consumption, but also in their leisure and home production. This has recently been highlighted by Chiappori and Meghir (2015) and Cherchye et al. (2018), who propose to compare the economic well-being across individuals in terms of money-metric welfare indices, which take the full basket of goods into account. Thus, in addition to consumption inequality, I also consider inequality in utility ($u(c_{it}, C_{it}, l_{it}, D_{it})$). Note that this requires a structural model, since utility depends on the calibrated preference parameters, which attach weights to each good and are needed to be able to summarize all goods in a single index. While the main analysis focuses on inequality in utility directly, I also converted utility levels to a money-metric index, the values of which are more easily interpretable, and obtained similar results for inequality. These robustness checks are relegated to section 2.5.5.

Finally, I also consider inequality in expected *life-time* utility. The motivation for this is that this is the broadest utility index, which does not only capture the current economic well-being, that may be influenced by short-term shocks, but also the expectation about the future. The expected *remaining* life-time utility of an individual with age t is denoted as $\mathbb{E}u_{i,t}$:

$$\mathbb{E}u_{it} = u(c_{it}, C_{it}, l_{it}, D_{it}) + \beta V_{it}$$

Here, V_{it} denotes the continuation value, conditional on being single or married and the corresponding state variables. To focus on economic inequality, I subtract the expected utility from 'love' (θ) from the continuation value, so that $\mathbb{E}u_{i,t}$ only captures expected streams of consumption, leisure and home production.[33] Note that differences in (remaining) life-time utility are partly due to a difference in the remaining number of periods when individuals do not have the same age (for example, consider comparing $\mathbb{E}u_{it}$ of a 20-year-old and a 40-year-old). To make utility levels comparable across age groups, I compute the *per-period* value u_{it}^{LT} that would be required over the remaining life-time to achieve utility $\mathbb{E}u_{it}$:

$$\sum_{t'=t}^{T} \beta^{t'-t} u_{it}^{LT} = \mathbb{E}u_{it}$$

Thus, u_{it}^{LT} converts the expected life-time utility to a per-period average and can be interpreted as a measure of expected economic well-being.

2.5.2 Decomposition of Inequality

The variance of these outcomes, denoted as X_i, can be decomposed into components due to inequality within and between households and due to singles.[34] The variances within and between couples are defined as follows (h is the index of the married household):

$$\begin{aligned} V^W(X_i) &= E\Big(V\Big(X_i|i \in h\Big)\Big) \\ V^B(X_i) &= V\Big(E(X_i|i \in h)\Big) \end{aligned}$$

The interpretation of the within-couple variance is that it measures how much allocations differ between spouses and takes the expectation of this variance over all couples. The between-household component is the variance of the household means, measuring inequality across households.

To decompose inequality of the population, one further needs to take singles into account. The shares of singles and married individuals in the population are p_s and p_m. Singles enter through two separate terms. The first is the variance of X_i in the group of singles, $V^S(X_i)$. The second term reflects that singles and married individuals differ in their mean outcome.[35]

[33]Since θ enters utility additively, one can compute the expected value of future realizations of θ for each state and subtract this value from the continuation values of singles and married individuals as defined previously. Note that the match quality shock θ would in principle introduce some interesting considerations about inequality. Low-wage individuals might be more willing to accept and remain in 'bad' marriages for economic reasons, so that their poverty would be reflected in the non-economic match quality component.

[34]In the appendix, I also consider other inequality measures, such as the Theil index, for which the results are very similar to the results from using the variance.

[35]This term is the between-variance that results from applying a within-between decomposition to the two groups of singles and married individuals:

$$\mathrm{V}(X_i) = p_m V^M(X_i) + p_s V^S(X_i) + V^{BMS}(X_i)$$

Here, $V^M(X_i)$ is the variance across married individuals.

This is denoted as the variance 'between married and single individuals':

$$
\begin{aligned}
V^{BMS}(X_i) &= p_s\Big(E(X_i|\text{ singles}) - E(X_i)\Big)^2 \\
&+ p_m\Big(E(X_i|\text{ married}) - E(X_i)\Big)^2
\end{aligned}
$$

V^{BMS} partly captures economies of scale. Married individuals can share the public consumption good with their spouse and are able to afford a higher living standard than singles. In addition, spouses can pool home production time and enjoy more leisure. Finally, the term can also reflect selection into marriage - if, for example, high ability individuals were more likely to get married, this would lead to a mean difference between singles and married individuals even when there are no economies of scale.

Using these expression leads to the following decomposition of the population variance:

$$
V(X_i) = p_m V^W(X_i) + p_m V^B(X_i) + p_s V^S(X_i) + V^{BMS}(X_i)
$$

Note that the components are weighted by p_s and p_m. The higher the fraction of singles or married, the higher is the contribution of the group to the population variance.

Table 2.4 shows the results of the decomposition based on the calibrated model. The decomposition captures cross-sectional inequality, since it pools all age groups and reflects the population in steady state. The first column reports the decomposition of log private consumption. The single component accounts for 21.4% of the total variance. Note that this value takes the population fraction of singles into account, which reduces the size of the component relative to the variance of singles. The within household variance accounts for 23.9% of inequality in private consumption, while the between household component accounts for about 54.3%. The variance between married couples and singles is small (0.4%), indicating that the mean difference in private consumption between the two groups is small.

The second column shows the decomposition for the per-period utility from private and public consumption, leisure and home production. Thus, this measure takes all private and public goods into account and weights them according to the calibrated preference parameters. Reflecting the importance of public goods (public consumption and home production), the within-couple component now accounts for 6.1% of total inequality. Interestingly, the variance between singles and married couples gets more important in this case (8.9%), which shows that the economies of scale in terms of public consumption and home production, which couples enjoy relative to singles, are substantial.

The third column finally shows the decomposition for expected life-time utility (u_{it}^{LT}). At any point in time, some individuals may be worse (or better) off due to randomness in wages, marital status and fertility, but this is mitigated when taking the expectation over the future. Note that the presence of divorce creates additional differences in the well-being of spouses. For example, consider a couple of a low-wage man and a high-wage woman,

Table 2.4: Variance decomposition

	$\log(\mathbf{c_{it}})$	$\mathbf{u(c_{it}, C_{it}, l_{it}, D_{it})}$	$\mathbf{u_{it}^{LT}}$
Variance	0.1	1.26	0.64
Within couples (% of total)	23.9	6.1	9.5
Between couples (% of total)	54.3	65	72.9
Singles (% of total)	21.4	20	17.2
Betw. sin. and mar (% of total)	0.4	8.9	0.4

Notes: The table shows the variance decomposition for private consumption, the per-period utility and (per-period) remaining life-time utility based on the model. Each row reports the fraction of the total fraction due to this component.

which is likely to divorce soon. If marriages on average are substantially assortative, the man would be worse off than the woman in terms of expected future utility, since he would expect to get remarried to a woman with a lower wage than his current wife. This reasoning also applies to inequality *across* couples, since a 'stable' couple is better off than a couple where divorce is impending. Overall, the within-household components accounts for 9.5% of total inequality for the life-time utility measure, which is relatively similar to the case of looking at per-period utility only. Note that the importance of inequality between singles and married individuals drops significantly (from 8.9% to 0.4%). This reflects singlehood being a transitory phenomenon for most individuals. While singles have a lower living standard than married individuals, they mostly expect to get married soon.

For comparison and robustness, I also conducted the decompositions for alternative measures, including measures that convert utility levels into money-metric indices. These cases are discussed in more detail in section 2.5.5. The results from the alternative measures are very similar to considering inequality in utility.

2.5.3 Experiment: Increase in Tax Progressivity

The policy experiment studies the effects of a hypothetical increase in progressivity, by varying the progressivity parameter (ψ_2) of the tax function

$$T(y) = \max\Big((1 - \psi_1 y^{\psi_2}) \cdot y, 0\Big)$$

The level parameter ψ_1 is adjusted to keep the budget of the government balanced. The policy experiment thus raises tax rates for higher incomes, while lowering them for lower income. The current system is approximated by a progressivity parameter of 0.15 and a level parameter of 0.61, which results from fitting the tax function to income tax rates reported by the OECD. When the progressivity parameter (ψ_2) is changed, the level parameter (ψ_1) is adjusted to ensure that the budget constraint of the government is balanced. The target level of expenditure is set to the revenue that the government obtains for the current tax system in the calibrated model. To study the effects of these reforms, I compare steady states, which can be interpreted as analyzing the long-run impact of the reform.[36]

The hypothetical reform increases the progressivity parameter by 0.06. To illustrate the magnitude of the tax change, figure 2.4 shows the change of the average tax rate due to the policy change. The reform reduces the average tax rate at lower incomes by up to 4%, while it increases the average tax rate for high income levels (around 150.000 €) by $\approx 5\%$. The motivation for this particular reform is that it is within the range of the cross-country variation in progressivity estimated e.g. by Holter et al. (2019). The focus of the policy analysis is mainly to compare the relative magnitude of the different components of inequality. I also conducted the analysis for different potential reforms and the relative importance of the different components (in particular, the role of intra-household inequality) is similar.

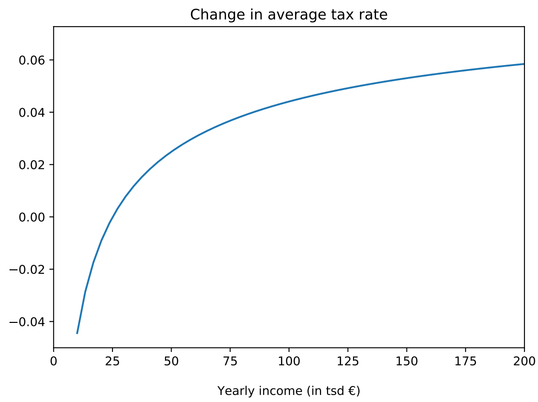


Figure 2.4: Illustration of policy experiment

Notes: The figure shows the change in the average tax rate due to the reform.

2.5.4 How does the Reform affect Inequality?

To analyze how the reform affects inequality along the dimensions of the marriage market, I decompose the *change* in inequality. The variance before and after the reform are denoted as V_0 and V_1:

$$V_k(X_i) \quad = \quad p_{m,k}V_k^W(X_i) + p_{m,k}V_k^B(X_i) + p_{s,k}V_k^S(X_i) + V_k^{BMS}(X_i)$$

[36]Computing the transition path would raise some complications regarding the marriage market equilibrium. Following the reform, individuals would have to forecast how the reform affects the marriage and divorce decisions of others, and thereby the pool of singles they will meet. Intuitively, the main difference between the transition and the new steady state concerns the decision weight in existing couples: changes in the weight would require that a participation constraint binds, which is more likely for larger reforms.

The goal is to express the percentage change in the variance ($\hat{V} = \frac{V_1 - V_0}{V_0}$) as a weighted sum of the change in each of the variances from the decomposition. The *change* in the variance between the two policy regimes can be decomposed in the following way:

$$\hat{V}(X_i) = \omega_1 \hat{V}^S + \omega_2 \hat{V}^W + \omega_3 \hat{V}^B + \hat{C}^{BMS} + \omega_0 \hat{p}_m + \hat{R}$$

$\hat{V}_S$ is the percentage change in the variance of singles, $\hat{V}^B$ and $\hat{V}^W$ are the changes of the variances between and within married couples. These changes are weighted by the importance of the component before the reform ($\omega_1, \omega_2, \omega_3$), to reflect that each of the percentage changes acts on a different base. For example, consider a case where the within-couple component declines by 20% and accounts for 10% of the total variance before the reform. This, in isolation, would lead to a 2% ($= 0.1 \cdot 0.2$) decrease in the total variance.[37] I further separate the effect of changes in the fraction of the population that is married ($\omega_0 \hat{p}_m$). p_m enters all of the components of the decomposition and has two effects. First, an increase in p_m reduces the weight placed on the variance across singles and increases it for the variances within and across couples. Second, p_m enters V^{BMS}. Separating the effect of p_m introduces a residual into the decomposition.[38] Finally, $\hat{C}^{BMS}$ is based on the variance between couples and singles. The expressions for this term and each of the weights are described in more detail in the appendix.

Table 2.5 shows the results of the decomposition for private consumption and the utility measures. The first row contains the reduction in inequality due to the reform (for example, 10.9% for private consumption). The other rows decompose the total reduction into parts due to each of the components. For private consumption, 24.77% of the total reduction is due to a reduction of within-household inequality. Thus, a fairly substantial part of the total reduction in private consumption would be missed by assuming equal sharing within households. For the utility measures, the corresponding fractions are 11.43% and 9.84%, which is smaller than for private consumption only, but still a noteworthy contribution to the total reduction in inequality. The role of the between-couple component is largest and ranges between 54.29% and 70.49%. The decomposition further allows to quantify the role of changes in the fraction of singles. The fraction of singles shifts the weights of the variance components in the decomposition. This, in itself, can generate a change in inequality. For example, when the variance of singles is higher than the variance of married individuals, the lower fraction of singles can decrease inequality. Based on the calibrated parameter values, the fraction of singles declines by a very modest amount (0.2 percentage points) in the more progressive tax system, which leads to very small implications for inequality (up to 2.86% of the total reduction).

The decompositions from tables 2.5 and focused on the contributions of each component

[37] Formally, this corresponds to the formula $\widehat{x+y} = \hat{x}\frac{x}{x+y} + \hat{y}\frac{y}{x+y}$.

[38] The residual results from applying the approximation $\widehat{xy} = \hat{x} + \hat{y} + \hat{x}\hat{y} \approx \hat{x} + \hat{y}$.

Table 2.5: Variance decomposition - Policy change

	$\log(\mathbf{c_i})$	$\mathbf{u}(\mathbf{c_i}, \mathbf{C_i}, \mathbf{l_i}, \mathbf{D_i})$	$\mathbf{u_{it}^{LT}}$
Change (%)	-10.9	-3.5	-6.1
Decomposition of change			
Within couples (%)	24.77	11.43	9.84
Between couples (%)	63.3	54.29	70.49
Singles (%)	11.93	31.43	21.31
Betw. sin. and mar. (%)	-0.92	0	0
Single probability (%)	0	2.86	-1.64

Notes: This table shows the effect of the policy change on the variance of private consumption, the per-period utility and life-time utility. The first row shows the total change. The other rows decompose this total change in inequality and add up to 100%.

to the total change in variance. This depends both on how important a component is under the status quo and how *reactive* it is to the tax system. Recall the example where the within-couple component declines by 20% (this is the 'reactiveness') due to the reform and accounts for 10% of inequality under the status quo , resulting in a 2% decrease in inequality. Table 2.6 focuses only the *reactiveness*, in terms of how strongly the within- and between-couple variance decline. This is interesting because it could be that the reactiveness of the components differs, which could be interpreted as progressivity being more 'effective' at reducing either the within- or between couple variance.

Table 2.6 shows that the reactiveness of both components is quite similar for private consumption. The between-couple component declines a bit more strongly than the within-couple component (-12.7% vs -11.4%). Thus, the relative rate of decline is 0.89. Interestingly, the within-couple variance declines *more strongly* than the between-couple component for the per-period utility (-6.1% vs -2.7%), so that in this case the within-couple component is more reactive. Note that differences in reactiveness imply that the *composition* of inequality is endogenous to the tax system. For example, for the utility measure, the fact that the within-household component declines more strongly than the between-household component suggests that a smaller fraction of the *remaining* inequality should be due to within-household inequality.[39] In practice, however, these composition changes are very

[39] An extreme example for this effect would be a case, in which a tax reform leaves intra-household inequality unchanged, but completely eliminates inequality across households. Then, the 'remaining' inequality would entirely be due to intra-household inequality.

small, which is shown in appendix table B.5. The share of the within-household component declines only from 6.2 to 6.1%. Finally, table B.6 replicates the decomposition of the effect of increasing progressivity for the case of *decreasing* it to a level of $\psi_2 = 0$, which replaces the tax function by a proportional tax. The relative importance of the changes in within- and across-couple inequality are similar to the previous results. Introducing a flat tax increases inequality and within-household inequality accounts for 25.4% of this increase for private consumption and 8.33% of the increase for the per-period utility.

Table 2.6: Unweighted change in variance

	Var. within couples	Var. between couples	Within rel. to between
(a) $\log(\mathbf{c_i})$			
Variance ($\psi_2 = 0.15$)	0.029	0.066	0.439
Change ($\psi_2 = 0.21$)	-0.114	-0.127	0.898
(b) $\mathbf{u}(\mathbf{c_i}, \mathbf{C_i}, \mathbf{l_i}, \mathbf{D_i})$			
Variance ($\psi_2 = 0.15$)	0.095	1.033	0.092
Change ($\psi_2 = 0.21$)	-0.061	-0.027	2.259

Notes: This table shows the variance within and between couples and of singles. The first row in each section shows the level of the variance for the calibrated model and the other rows show the change (in percent) due to policy reforms. The final column shows the ratio between the within and the between component of each row.

2.5.5 Inequality Decompositions: Robustness

For further comparison and robustness, table 2.7 reports the results of the variance decomposition for three additional measures of well-being. The first column reports the decompositon results for inequality in total consumption expenditure ($c_{it} + C_{it}$). The first noteworthy feature is that the fraction due to the within-couple components drops significantly relative to the case of private consumption only and becomes small (1.3%). This reflects household expenditure being to an important extent public and indicates that there is a limited role for intra-household inequality in total consumption, unless one makes further assumptions about preference heterogeneity or reduces the share of consumption that is considered public. The importance of the within-household component here is lower than found in Lise and Seitz (2011), partly reflecting that they assume a lower fraction of public consumption. The second interesting feature is that the variance between singles and married individuals becomes much more important for total expenditure than for private expenditure only, which is due to economies of scale in consumption.

The motivation for the remaining three columns is that the main analysis has focused on inequality in *utility*, the cardinal value of which is hard to interpret. Thus, one might wonder if the results would differ for money-metric measures of well-being. The second column computes the variance decomposition for 'full consumption', as in Lise and Yamada (2018). Full consumption is defined as the sum of consumption expenditure and the *market value* of leisure and home production time. This is defined as follows:

$$I_i = c_i + C_i + \tilde{w}_i l_i + \tilde{w}_i d_i + \tilde{w}_j d_j$$

Here, $\tilde{w}_i$ is the net hourly full-time wage of the individuals, which captures the opportunity cost of leisure and home hours, and the index j refers to the home hours of the partner. Note that tax reforms have a direct impact on net wages and therefore influence this measure through their impact on $\tilde{w}_i$ and $\tilde{w}_j$ in addition to the impact on allocations. Still, the measure can be considered as a useful reference and comparison point, especially since it can in principle be measured in the data more easily than measures which require estimates of preferences.

The third and fourth column is an equivalence scale in the spirit of Chiappori and Meghir (2015) or Pendakur (2018). These measures are based on comparing utility levels across people, but convert the utility level of each individual into the monetary amount that a reference person would need to achieve this utility level. For the utility level of each person ($u(c_i, C_i, l_i, D_i)$), one can compute the equivalent amount of resources the reference person would need in order to obtain this utility level. The reference person is set to be a childless man, who does not work and receives the equivalent amount as a transfer, while optimally chosing leisure and home time. The fourth column applies this approach to the expected life-time utility and computes the amount that the reference person would need to obtain the expected per-period utility u_{it}^{LT}.

The overall take-away from considering these measures in table 2.7 is that the results are very similar to considering inequality in utility ($u(c_{it}, C_{it}, l_{it}, D_{it})$ and u_{it}^{LT}) directly. In particular, this holds for the role of within-household inequality. Recall that for the per-period utility, 6.1% of inequality is due to the within-component. For the equivalence scale and full expenditure, this value is 6.7% and 5.6%. Similarly, the equivalence scale for life-time utility leads to a very similar result as previously. Table 2.8 further shows the decomposition of the reduction in inequality due to these measures, which are also very similar to the previous decompositions. For example, for the per-period utility, 11.43% of the total reduction is due to the within component, while the corresponding number is 11.1% for the equivalence scale.

A final question for robustness is whether the choice of the inequality measure matters, which

Table 2.7: Variance decomposition - alternative measures

	$\log(c_i + C_i)$	Log full exp.	Log eq. sc.	Log eq. sc. (LT)
Singles (%)	20.4	22.6	14.2	17.3
Within couples (%)	1.3	5.6	6.7	9.0
Between couples (%)	52.0	59.5	70.2	73.5
Betw. sin. and mar (%)	26.3	12.4	9.0	0.3

Notes: This table computes the variance decomposition for the status quo for the alternative measures and reproduces table 2.4 for these measures. The columns contain log total consumption, log full expenditure and the logs of the equivalence scale values corresponding to $u(c_{it}, C_{it}, l_{it}, D_{it})$ and u_{it}^{LT}.

Table 2.8: Policy change - alternative measures

	$\log(\mathbf{c_i} + \mathbf{C_i})$	Full expenditure	Log eq. sc.	Log eq. sc. (LT)
Total (%)	-8.4	-10.9	-4.5	-7.7
Decomposition of change				
Within couples (%)	1.2	4.6	11.1	10.4
Between couples (%)	81	60.6	62.2	68.8
Singles (%)	15.5	28.4	24.4	20.8
Betw. sin. and mar. (%)	-1.2	4.6	0	0
Single probability (%)	3.6	1.8	0	-1.3

Notes: This table shows the effect of the policy change for different outcomes. The columns contain log total consumption, log full expenditure and the logs of the equivalence scale values corresponding to $u(c_{it}, C_{it}, l_{it}, D_{it})$ and u_{it}^{LT}. The first row shows the total change. The sum of the components from other rows is equal to the total change in the first row.

so far was the variance. Similar decompostions can be applied to the Mean Logarithmic Deviation and the Theil Index, since they are subgroup-decomposable. The results are qualitatively similar and relegated to the appendix.

2.5.6 Marriage Outcomes

In this section, I discuss marriage outcomes and how they are affected by the increase in tax progressivity in more detail, focusing on differences by ability and gender asymmetries. To facilitate the interpretation, the analysis groups individuals into 'ability groups' ranging from low (L) to high (H). Each of these groups contains averages over three types of the actual type distribution.[40] Table 2.9 shows marriage rates and the mean Pareto weight for women and men from each of these groups. The table is based on considering all meetings between individuals below age 30, which is the period in which most marriages occur, and calculating the probability that the meeting results in a marriage, and the mean Pareto weight conditional on marriage. For men, the table reports *their* Pareto weight (i.e. $1 - \lambda_{it}$).

Table 2.9: Marriage rates and Pareto weights

	Women - L	W - LM	W - UM	W - H	Men - L	M - LM	M - UM	M - H
(a) Status quo								
Marriage probability	0.289	0.274	0.254	0.223	0.227	0.251	0.271	0.289
Own Pareto weight	0.453	0.472	0.496	0.526	0.451	0.487	0.528	0.584
(b) Reform								
Change in mar. prob. (in pp)	-0.22	-0.26	0.29*	0.7***	0.38**	0.43**	-0.29*	0.05
Change in PW (in pp)	0.31***	0.45***	0.52***	0.62***	-0.11*	-0.38***	-0.43***	-0.94***

Notes: The table reports the marriage market outcomes under the status quo and after the reform. The asterisks indicate whether the difference of each outcome between the two simulations is statistically significant at the 1%, 5% or 10% level.

The table shows an interesting gender asymmetry with respect to marriage rates. Low ability women are *more* likely to get married than high-ability women, whereas low-ability men are *less* likely to get married than high ability men. Quantitatively, these effects are modest and the marriage probability varies by up to 6 percentage points across the ability groups. The reason for this effect is that marriages in which men are from a higher ability group than the women have a higher surplus than the opposite combination, since women are more likely to reduce their market hours in marriage. For both men and women, the Pareto weight is increasing in the own ability type and the average ranges between 0.45 and 0.58.

The second part of the table shows how these outcomes differ in the new steady state under the more progressive tax system. The changes are expressed in percentage points. Since some

[40]For example, group L contains individuals of type 1, 2 and 3.

of these are quite small, the table also reports whether the difference is statistically significant from zero, rather than representing simulation noise. The changes in the Pareto weight are mostly significant on the 1% level. Women on average receive higher Pareto weights, whereas the opposite is the case for men.[41] Regarding marriage rates, only the rate of the highest two groups of women and the lowest two groups of men increase in a statistically significant way. The magnitude of these changes ranges between 0.29 and 0.7 percentage points. To further illustrate these patterns, table 2.10 shows how the marriage probability and the Pareto weight change for each *combination* of types - for example when a women from the highest group meets a men from the lowest group. While the magnitude of these changes is again modest, the table reveals some further gender asymmetries. For marriage rates, most of the significant changes occur for meetings between high ability women and low ability men, whereas there is no significant change for the high ability men and low ability women. For example, the reform increases the probability that a women from the highest group gets married to a men from the lowest group by 0.64 percentage points. For Pareto weights, the pattern is different since the effects are largest for low ability women getting married to high ability men. The Pareto weight of the woman increases by up to 1.52 percentage points (recall that Pareto weights are scaled between zero and one).

2.5.7 Relating Changes in Marriage and Divorce to Inequality

Changes in marital behavior can contribute to the effect of the reform on inequality. For example, a decrease in the assortativeness of matching would decrease inequality *across* couples, since more high-income individuals get married to low-income individuals. At the same time, this would increase inequality *within* couples, since the fraction of couples with unequal earnings potential would rise. The goal of this section is to quantify the role of such composition effects and analyze whether they contribute to the results from the variance decompositions of section 2.5.4.

I perform further decomposition of each of the components of the variance decomposition, focusing on the role of the ability types. To illustrate the general approach, I first focus on the variance across couples. The share of couples, in which the wife has type k and the husband has type j is denoted as s_{kj}. The across-couple variance can be rewritten in the following way:

$$\begin{aligned} V(\bar{X}_h) &= \sum_{kj} s_{kj} V(\bar{X}_h|k,j) + \sum_{kj} s_{kj} \Big(E(\bar{X}_h|k,j) - E(\bar{X}_h) \Big)^2 \\ E(\bar{X}_h) &= s_{kj} E(\bar{X}_h|k,j) \end{aligned}$$

Here, $\bar{X}_h$ is the mean outcome for couple h (recall that the across-couple variance is the

[41] This results from the gender wage gap, since each type of woman on average have a lower wage than the corresponding type of man.

Table 2.10: Changes in marriage rates and Pareto weights - by type combination

	Men - L	Men - LM	Men - UM	Men - H
(a) Marriage probability				
Women - L	-0.02	-0.04	0.08	-0.22
Women - LM	0.2	-0.87**	-0.21	-0.13
Women - UM	0.46**	0.78***	-0.34	0.17
Women - H	0.64***	0.83***	0.41*	0.59
(b) Pareto weight				
Women - L	-0.06	0.74***	0.96***	1.52***
Women - LM	0.28	0.17***	0.91***	1.4***
Women - UM	-0.24	0.45***	0.31***	1.45***
Women - H	-0.39*	-0.62***	0.07	0.59***

Notes: The table reports changes between the status quo and the new steady state, expressed in percentage points. In this table, the Pareto weight always refers to the weight of the woman. The asterisks indicate whether the difference of each outcome between the two simulations is statistically significant at the 1%, 5% or 10% level.

variance of the household means). Thus, the variance across couples can be expressed as a function of the variance among couples of type (k, j) $(V(\bar{X}_h|k, j))$, the mean outcome of couples of this type $(E(\bar{X}_h|k, j))$ and the type share s_{kj}. The idea of the decomposition is to use this formula to compute a counterfactual post-reform variance, in which only the type shares s_{kj} are varied to their post-reform level and all the other components (variance and mean conditional on type) remain at their pre-reform values. The interpretation of this procedure is to ask to what extent the change in the type probabilities alone can explain the total change in the variance.[42]

Similar formulas can be derived for the other variance components. The variance across singles can be analyzed completely analogousl¡ to the variance across couples, using the same formula and only replacing the household mean $(\bar{X}_h)$ by the value of the single (X_i). The corresponding formula for the within-couple variance is:

$$E\Big(V(X_i|i \in h)\Big) = \sum_{i,j} s_{k,j} E\Big(V(X_i|i \in h)|k, j\Big)$$

[42]An alternative approach which would yield very similar results would be to reweight the simulated post-reform data to have the same composition of observables (in terms of types k and j) as the pre-reform data and compute the variance decomposition based on the reweighted data, as is done in Lise and Seitz (2011).

Like for the case of the across-couple variance, changes in the within-couple variance are in principle driven by the type shares ($s_{k,j}$) and the variances conditional on the types. Finally, the variance between singles and married individuals can be further decomposed using the following formulas:

$$\begin{aligned}
V^{BMS}(X_i) &= (1-p_m)\Big(E\Big(X_i|S\Big)-E\Big(X_i\Big)\Big)^2+p_m\Big(E\Big(X_i|M\Big)-E\Big(X_i\Big)\Big)^2\\
E(X_i|S) &= \sum_k s_k^S E\Big(X_i|S,k\Big)\\
E(X_i|M) &= \sum_{kj} s_{k,j}^M E\Big(X_i|M,k,j\Big)\\
E(X_i) &= (1-p_m)E\Big(X_i|S\Big)+p_mE\Big(X_i|M\Big)
\end{aligned}$$

s_k^S is the fraction of singles of type k and $s_{k,j}^M$ is the fraction of couples with types k an j. The common feature of these formulas, for each of the components of the variance decomposition, is that they allow to vary only the type composition of the groups of married and single individuals.[43]

Table 2.11 shows the results for this decomposition for inequality in private consumption and the per-period utility. The interpretation of the first row ('fixed p_m') is that it adjusts s_{jk} to the values in the new steady state, but keeps everything else (including p_m) constant. The interpretation is that only the type composition of the pools of married and single individuals are varied. The second row also lets p_m adjust and considers changes in the *weighted* contribution to the total variance, i.e. each of the following four summands:

$$V(X_i) = p_mV^W(X_i)+p_mV^B(X_i)+p_sV^S(X_i)+V^{BMS}(X_i)$$

For example, the interpretation of the very first table entry is that changes in s_{jk} explain -1.91% of the reduction in the within-couple variance, which means that these changes by themselves would marginally increase within-couple inequality. The overall take-away from considering the 'fixed p_m' case for both outcome variables is that the effect of composition effects alone is small. Only for the variance between married and single individuals (BMS), the composition effect can generate a change of a noteworthy magnitude (relative to the change in this component). An interesting aspect of these results is that the results from the previous section suggested that the increase in tax progressivity could decrease assortative mating, since there was a small increase in the likelihood of low-ability men getting married to higher-ability women. However, this effect is not strong enough to lead to a quantitatively remarkable reduction in inequality.

[43]Note that this in an accounting decomposition, since endogenous quantities are varied independently.

Table 2.11: Fraction of variance change explained by type probabilities (in %)

	Within couples	Between couples	Singles	BMS
(a) $\log(\mathbf{c_{it}})$				
Fixed p_m	-1.91	-0.37	2.75	24.45
Flexible p_m	-3.53	-1.8	20.01	13.99
(b) $\mathbf{u}(\mathbf{c_{it}}, \mathbf{C_{it}}, \mathbf{l_{it}}, \mathbf{D_{it}})$				
Fixed p_m	-2.8	-0.29	1.34	-23.95
Flexible p_m	-5.82	-7.79	14.91	37.75

Notes: The table reports the share of the change in each of the variance components that can be explained only by the change in type probabilities.

When also letting the fraction of married individuals adjust ('flexible p_m'), the explained fraction become somewhat larger. Note that the outcome is now for example $p_m V^W(X_i)$ (instead of $V^W(X_i)$). Since the reform increases the fraction of married individuals, this in itself increases the contribution of inequality within and across couples and decreases the contribution of singles. Thus, the explained fraction is negative in the first two columns and positive in the third. The explained share is overall largest for the per-period utility and ranges between -7.8% and 14.91%.

2.5.8 Social Welfare

In this section, I show how the reform affects expected life-time welfare. While inequality is strongly linked to social welfare, the most natural social welfare function in my model is based on the expected life-time welfare of a newly born cohort. Thus, the analysis in this section is concerned with the overall welfare impact of the reform, and to what extent it is driven by marriage market adjustments.

2.5.8.1 The Comparison Case

To assess the role of the marriage market, I consider an alternative version of the model in which the decision rules for marriage, divorce and the Pareto weight are exogenous and set to exactly reproduce the decision rules of the calibrated model. Then, I simulate the reform in this alternative model. The interpretation is to ask what the effects of the reform would be if marriage and divorce decisions and the Pareto weights were exogenous and would not react to the change in tax policy. I will refer to the initial equilibrium given this policy parameter as the benchmark case/schedule. In addition, I will refer to the model with fixed decisions as the *alternative* or *fixed marriage market* model.

More concretely, suppose that the initial progressivity parameter is ψ_2^0. Solving the model leads to decision functions which map the state variables of two singles that meet, ω_f and ω_m, and the realization of the match quality shock θ into a decision whether to get married ($M \in \{0,1\}$) and an initial Pareto weight $\lambda \in [0,1]$:

$$\begin{aligned} M &= m_0(\omega_f, \omega_m, \theta) \\ \lambda &= m_1(\omega_f, \omega_m, \theta) \\ \omega_f &= (\epsilon_f, \alpha_f, b_f, A_f) \\ \omega_m &= (\epsilon_m, \alpha_m, b_m, A_m) \end{aligned}$$

The state variables ω_g comprise the productivity shock, the type, the presence of children, and assets. Similarly, there are decision functions d_0 and d_1 which map the state variables of any couple (ω_c) into the decision whether to get divorced ($D \in \{0,1\}$) and whether to renegotiate the current Pareto weight λ to a new Pareto weight $\tilde{\lambda}$:

$$\begin{aligned} D &= d_0(\omega_c) \\ \tilde{\lambda} &= d_1(\omega_c) \\ \omega_c &= (\theta, \epsilon_f, \epsilon_m, \alpha_f, \alpha_m, b, A, \lambda) \end{aligned}$$

In the comparison case, I solve the model for a new tax schedule ψ_2^1, but use the decision functions m_0, m_1, d_0, d_1 that were obtained given ψ_2^0, instead of allowing individuals to determine these outcomes endogenously. In particular, this procedure does not allow the Pareto weights to adjust following a reform. Instead, spouses are exogenously assigned the 'old' Pareto weight, which would have been the bargaining outcome given the old tax schedule. As a result, comparing the alternative and the full model allows to study the implications of the endogenous adjustment of the Pareto weights. The exercise is related to the exercise conducted in Knowles (2012), who compares a model with flexible and fixed Pareto weights in order to quantify the error one would make with a unitary version of the model relative to a model with endogenous Pareto weights. Note that in the alternative model used here, the exogenous Pareto weight depends on the characteristics of the couple and is allowed to change exogenously over time.[44]

[44]Further note that when there are no dynamic choices (savings), the alternative model always delivers exactly the same marriage market outcomes that are obtained under the benchmark schedule. In other

The alternative model has different implications for how intra-household inequality changes with progressivity. When Pareto weights are fixed, the tax system can influence intra-household inequality mainly by changing the relative leisure of spouses. A less progressive system can increase leisure of secondary earners with low Pareto weights if it leads to a reduction in their labor supply and home production hours increase less than work hours are increased. With flexible Pareto weights, progressivity directly affects relative consumption by changing the Pareto weight.

2.5.8.2 Results

Table 2.12 shows the overall welfare change for the increase in progressivity for each ability type, averaged over men and women. These are also illustrated graphically in figure 2.5. The welfare change is expressed in terms of equivalent variation. This is the percentage change in consumption in each state under the status quo, which makes individuals indifferent between the status quo and the reform.[45] The second and third columns show the equivalent variation (in percent) for the full model and for the comparison model and the fourth column contains the difference between the two. This difference can be interpreted as the effect of the endogenous marriage market adjustments: it is the equivalent variation that the full model generates relative to the case of fixed marriage market decisions. To ease interpretation of these numbers, the final column converts the difference to a percentage amount, relative to the welfare impact in the fixed marriage market case. Note that "1" is the lowest ability type and "12" is the highest.

As one would expect, the lowest ability types benefit from the increase in progressivity, whereas the highest ones lose, both for the full model and the fixed marriage market. In the middle of the distribution most types are made worse off. This can be the case because of the labor supply disincentives generated by a more progressive tax system. Note that the fact that most types lose is also reflected in the aggregate welfare change, which averages over all types and therefore corresponds to a utilitarian welfare objective. The aggregate welfare loss is -0.92% for the full model and -1.22% for the fixed marriage market model. Thus, in relative terms, the difference amounts to 24% of the impact without marriage market adjustments.

The welfare difference between the full model and the fixed marriage market is largest for the two lowest types. In particular, the lowest type experience a welfare gain of 1.98% of

words, the cross-sectional distribution of marriages and Pareto weights is *identical* to the benchmark model. With dynamic choices, policy changes can still affect marriage and divorce rates and Pareto weights even when decision functions are fixed. This is the case because decisions are fixed *conditional on state variables* and state variables (i.e. assets) are partly endogenous and affected by the reform. Thus, the interpretation of the exercise is how much the endogenous *adjustment of the decision functions* matters for welfare, rather than marriage market outcomes per se.

[45] Since the model features both private and public consumption, the equivalent variation is computed by increasing consumption of both of these goods by a certain percentage amount in each state of the world.

Table 2.12: Welfare effects - by ability type

	(1) Full model	(2) Fixed MM	Difference: (1) - (2)	Relative
1	1.98	0.86	1.12	1.31
2	1.37	0.53	0.84	1.61
3	-0.03	-0.29	0.26	-0.89
4	0.37	-0.25	0.62	-2.50
5	0.04	-0.49	0.54	-1.09
6	-0.47	-0.70	0.22	-0.32
7	-0.92	-1.29	0.36	-0.28
8	-1.46	-1.66	0.19	-0.12
9	-1.88	-1.74	-0.15	0.08
10	-2.62	-2.87	0.25	-0.09
11	-3.48	-3.27	-0.21	0.07
12	-3.94	-3.44	-0.51	0.15
Total	-0.92	-1.22	0.30	-0.24

Notes: This table shows the difference in the welfare change between the full model and the fixed marriage market case. It is based on the values shown in figure 2.5.

consumption in the full model, whereas it would only be 0.86% if marriage market decision are fixed. These differences get somewhat smaller for higher ability types. In the upper half of the ability distribution, the marriage market adjustments generate between -12 and 15% of the welfare impact with a fixed marriage market. In general, these values illustrate the importance of the marriage market relative to the 'standard' channels of tax progressivity and show that while the marriage market does not dominate the welfare calculation, it contributes in a non-negligible way.

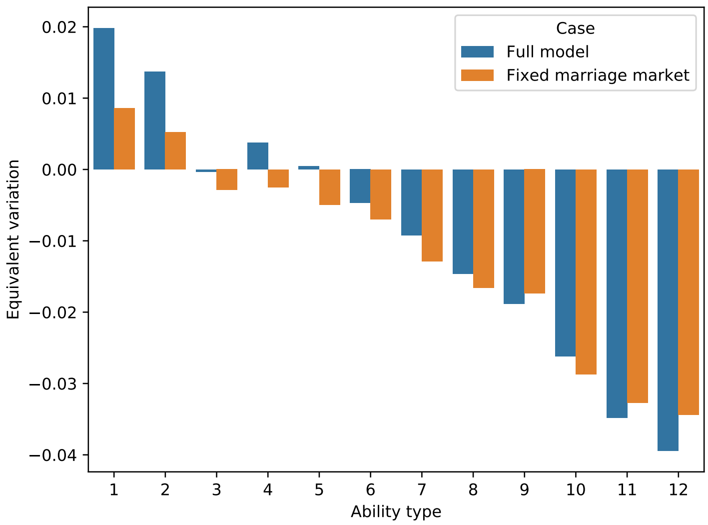


Figure 2.5: Welfare impact along the ability distribution

Notes: The figure shows the change in welfare (expressed as percent of consumption) for the full model and the fixed marriage market case.

2.6 Conclusion

This chapter studies how the progressivity of the labor income tax affects inequality through the marriage market. Progressive taxation reduces intra-household inequality because it makes the relative outside options of spouses - the values of being single - more equal. In addition, single rates, marriage and divorce are also affected since progressivity changes the economic value of individuals on the marriage market and how selective they are about potential partners. To study these effects, I calibrate an equilibrium model of marriage, divorce, labor supply and savings to data from the Netherlands. The main question that the analysis addresses is to what extent the marriage market channels - bargaining and marriage - affect inequality on top of the usual effects of progressivity.

The model is first used to decompose inequality in consumption and welfare into components due to within and between married couples. The within-couple component captures that spouses consume different amounts of private goods and has received little attention in studies of progressivity. In the calibrated model, intra-household inequality accounts for 23.9% of the cross-sectional variance in private consumption. The model further allows to study inequality in the utility from private and public consumption, leisure and home production. In this case, the intra-household component accounts for 6.1% of the total variance. The model is then used to study a hypothetical reform that increases progressivity by 40% relative to its current level. The contribution of the intra-household component to the total reduction in inequality is 24.77% for private consumption and 11.9% for the utility measure. These quantitative finding shed light on the question of how important intra-household inequality could be relative to overall inequality, which has received little attention in the literature so far. The model further suggests that the induced changes in marriage and divorce have small implications for inequality.

There are several interesting dimensions in which the present analysis could be extended. While my chapter has focused on the role of the labor income tax, it would be very interesting to model the transfer system in more detail, which has an important redistributive function at the lower end of the income distribution. Another very interesting extension would be to consider the role of endogenous skill formation. Tax progressivity can reduce the incentive the accumulate human capital in working life and to invest in skills in terms of education and both of these aspects interact with marriage and intra-household bargaining. Finally, while I have focused on intra-household inequality in personal consumption and leisure, allowing for heterogeneous preferences, for example regarding public good expenditure or risk aversion, would increase the scope for disagreement between spouses. This would amplify the importance of intra-household inequality and thereby the welfare implications of changes in bargaining positions.

Chapter 3

Fertility and Women's Occupational Choice: Germany, 1975 - 2007

3.1 Introduction

Men and women differ in many important labor market outcomes. Gender gaps in employment and especially earnings and wages have received the most attention (see Blau and Kahn (2017) for a survey). Using an event-study methodology, a number of recent papers describe how the gender gap in these outcomes widens after the birth of the first child (Fernández-Kranz et al., 2013; Angelov et al., 2016; Kleven et al., 2018, 2019). The employment costs of motherhood seem to be partly unexpected, and have probably risen in recent decades (Kuziemko et al., 2018).

This chapter goes beyond the standard extensive and intensive margins of employment and studies the relationship between motherhood and the *type* of work that women do. Our large administrative dataset allows us to observe life-cycle patterns of occupations for women born between the 1940s and the 1970s. We merge this dataset with information on the typical tasks performed in these occupations, which we aggregate into four dimensions (analytical, interactive, routine, and non-routine manual). We document differences in these patterns by cohort and by the number of children, and trace out the dynamics of the substantial child-related differences using an event-study framework.

By focusing on women's tasks on the job and their relationship to motherhood, our chapter speaks to the debate about the role of job characteristics for the last chapter of the 'great gender convergence' (Goldin, 2014). Using a similar event-study approach, Felfe (2012) shows that while the main adjustment after childbirth is in terms of working hours, mothers who switch employers also tend to favor jobs with more flexible schedules, less night work and lower levels of stress. Pertold-Gebicka et al. (2016) and Kleven et al. (2018) show that the share of public-sector jobs (already higher for women at the beginning of their careers) increases after the birth of the first child. Hardoy et al. (2017) and Kleven et al. (2018)

document that the gender gap in management opens up considerably with motherhood. Albrecht et al. (2018) find that men and women do not differ much in the rate of firm-to-firm mobility, but that men experience faster wage growth both as stayers and as switchers.

The literature on tasks and the importance of ocupations and task-specific human capital has mostly focused on men, as for instance Autor et al. (2003), Gathmann and Schönberg (2010) or Yamaguchi (2012). Relatively little attention has been devoted to gender differences and women's task content, with Cerina et al. (2018); Ngai and Petrongolo (2017); Rendall (2018) or Yamaguchi (2018) being some notable exceptions. In this chapter, we add to this literature by studying the relationship between tasks and fertility and in particular how occupations change after women have their first child.

As a first step and motivation for our analysis, we document the life-cycle patterns of employment and task intensities, focusing on gender differences and differences by completed fertility. Women from recent cohorts, such as those born between 1975 and 1979, on average work in slightly more analytical and interactive occupations than men. The picture is different for older cohorts, in which men work in more analytical occupations. Thus, there has been a remarkable gender convergence over time. However, there is still some life-cycle divergence differences. Especially women with three or more children work in occupations with lower analytical and interactive task intensities at later ages.

We then study the dynamics around childbirth in more detail. To this end, we estimate the child-related effects on labor market and task variables in an event-study framework. The results show that the effect of childbirth on employment and earnings are very large in Germany and strongly increase in the number of children. The estimates suggest that analytical and interactive task intensities fall after childbirth relative to the age trend. In other words, women's analytical task intensities on average rise less than would be predicted in the absence of children. This could contribute to the gender wage gap, since analytical tasks are typically associated with a wage premium.

These event-study estimates take both variation in tasks due to occupational switches and due to selection resulting from participation decisions into account. Thus, we also use our data to exclusively focus on the role of occupational switches. Our main results for this analysis, where we compare occupational switches of mothers and childless women, are as follows. Mothers switch their occupation more often than childless women, in particular those with a long career break. While there is little evidence for direct occupational downgrading - which would imply that on average mothers return to work in occupations with worse-paying task combinations (such as those that involve little analytical work) - mothers increase their analytical task intensity less than childless women. This can be interpreted as stagnation relative to childless women, in the sense that mothers' analytical task intensity would have grown more strongly in the absence of children. A simple back-of-the-envelope calculation suggests that this effect can explain a reduction in the hourly wage rate by 7%. In addition, we find that mothers, again in particular those with a long career break, tend to work in occupations with a bigger *distance* in the task content relative to their pre-birth occupation. A final result is that there is occupational switching towards occupations with a high fraction of part-time workers, which could be interpreted as a switch towards 'family-friendly' occupations.

Our empirical findings can be related to and interpreted in the light of some theoretical and empirical work in the literature. For example, the increased rate of occupational switches would be consistent with a model of human capital depreciation (such as Adda et al. (2017)), in which long breaks from the labor market decrease the attachment to the previous occupation, or search frictions, in which losing the attachment to the previous employer leads some women to enter employment in a different occupation. From the perspective of models of task-specific human capital (see Gathmann and Schönberg (2010)), the bigger distance between current and pre-birth tasks for mothers could also imply that they are less able to transfer their previously accumulated human capital to the new occupation, since the new tasks differ from what they did before. This could in principle also contribute to a wage loss. In addition, the results from Goldin (2014) suggest that a key characteristic of an occupation is its family-friendliness. As a result, it could be the case that having children increases the need to work in a family-friendly environment, so that women increasingly switch towards these occupations. Our results are consistent with such a story. From this perspective, changes in the task composition of work would be byproduct of the increased orientation towards family-friendliness, but the changes in tasks can still contribute to the wage cost that women accept in exchange for the higher amenity value.

The rest of this chapter is structured as follows. Section 3.2 describes the data on labor market outcomes and task intensities which we use for the analysis. In Section 3.3, we analyze the cohort profiles. Section 3.4 focuses on the dynamics around childbirth. Finally, Section 3.5 concludes.

3.2 Data

3.2.1 The BASiD data

The BASiD dataset contains information on the employment and fertility histories of a sample of the German population. The sample is based on registers maintained by the German Pension Insurance (*Versichertenkontenstichprobe*/VSKT). This dataset contains information on fertility, employment and social security claims for the period from between 1951 and 2007. In addition, the dataset is merged with information from the Integrated Employment Biographies (IEB) of the Institute for Employment Research (IAB), which is available beginning in 1975. Among other aspects, the IEB contain daily information on employment, earnings, occupation, education, whether an individual works full-time or part-time and information on establishments. The data is collected from employers, who have to report this information for all jobs subject to the Social Security system.

We restrict our sample to the time period between 1975 and 2007 to focus on the period for which the labor market data from the IEB is available.[1] Based on the data on employment spells, we construct a quarterly panel using the longest employment spell in the quarter as the main job. We restrict the sample to individuals who start their employment history in West Germany. The IEB do not contain information on self-employed individuals, public servants and employment spells abroad, since these cases are not covered by the social insurance reporting. Therefore, we treat observations from terminal employment gaps, in which no further employment spells are reported for an individuals for the rest of the life-time, as missing rather than as voluntarily non-employed. For women, we only remove terminal gaps which do not occur immediately following childbirth, since it is conceivable that they decide to leave the labor force permanently in this situation.

The information on fertility results from the fact that individuals are required to report their fertility history to the German Pension Insurance, since children are relevant for determining retirement benefits. In almost all cases, households attribute the child to the retirement account of the mother and there are only very few observations of men for whom fertility information is available. Note that there are some individuals in the sample who have either not reported a history yet or reported a history at a relatively young age, when they might still have additional children in the future. As a result, we only focus on individuals with a complete fertility history. Kreyenfeld and Mika (2008) validate the fertility information from the VSKT and find that they reflect the official population statistics well once individuals with incomplete histories are removed.

The earnings variable from the IEB is top-coded at the maximum earnings that are relevant for social insurance payments (*Beitragsbemessungsgrenze*). This affects around 12% of observations for men and 4% of observations for women. We follow the procedure from Card et al. (2013) to impute censored wages based on tobit regressions by gender, age, education and year. Since the definition of the earnings variable changed in 1984, due to the inclusion of bonuses and one-time payments, we further apply the correction approach from Fitzenberger (2012) and Dustmann et al. (2009) to correct for this structural break. Earnings are deflated so that 2010 is the reference year.

3.2.2 The BIBB Qualifications and Careers Survey

Our data on tasks that workers typically perform on the job comes from the Qualification and Career Surveys, which are conducted by the Federal Institute for Vocational Education and Training (BIBB) and have previously been used to study task intensities (see for instance

[1]While some information on employment and earnings is also included in the VSKT, the earnings variable is not directly comparable to the variable from the IEB and no information on occupation is available.

Spitz-Oener (2006), Black and Spitz-Oener (2010) or Gathmann and Schönberg (2010)). These are representative cross-sectional surveys with roughly 20000 - 35000 respondents in each wave. There are six waves available, which were conducted 1979, 1985/86, 1991/92, 1998/99, 2006 and 2012. The surveys contain detailed questions about tasks that are required in an occupation, such as how often workers repair objects or how often they have to persuade co-workers. We classify each question as representing either analytical, interactive, routine or manual tasks and assign a value of 0, $\frac{1}{3}$ or 1, depending on whether the answer is 'never', 'sometimes' or 'frequently' (or 0/1 for yes/no questions). Since the questions are not always comparable across waves, we pool all waves to compute task intensities by averaging over all responses.[2] Note that the intensities are constructed in a way so that the four dimensions do not sum to 1, which follows the approach in Spitz-Oener (2006).[3] This can be interpreted as some occupations being generally more intensive or demanding in tasks than others.

We merge the information on tasks to our BASiD sample on the occupation level, using an occupational code with 120 occupations. This classification is relatively fine, since it allows for a variety of different occupations, while typically grouping occupations which are close to each other into a single occupation (such as mathematicians, physicists and chemists).[4] In the appendix, we include a list of all occupations according to our definition, which illustrates our classification.

We compute the task intensities in each occupation based on a pooled sample of men and women.[5] To illustrate how task intensities vary across occupations, table 3.1 shows the task indices for three particular occupations.

[2]This assumes that the task intensities are stable over time within occupations.

[3]There are two cases that lead to this. First, it can be the case that individuals in one occupation state that they perform more different tasks that are classified to a certain dimension. As an example, consider the simple case with two questions about analytical tasks and individuals in one occupation doing two of these tasks and individuals in the other occupation only one. Second, it could be the case that individuals in one occupation state more often that they only 'sometimes' perform certain tasks, which also leads to a lower absolute intensity. As a robustness check, we also rescaled the task variables (t_i) to sum up to one in each occupation (by computing $\frac{t_i}{\sum_i t_i}$) and the results are qualitatively similar.

[4]With a finer classification, which allows for more different occupations, there would be more moves between occupations, but the additional moves would be smaller in terms of the task distance.

[5]This increases the sample sizes for each occupation. As a robustness check, we also computed separate task intensities for men and women and the results are qualitatively similar.

Table 3.1: Task variables: Examples

	Analytical	Interactive	Routine	Manual
Plumber	0.1	0.23	0.29	0.37
Nurse	0.25	0.34	0.28	0.41
Architect	0.49	0.38	0.18	0.14

Plumbers are an example for a highly manual occupation, which also involves some routine and interactive tasks and fewer analytical tasks. Being a nurse also involves a large extent of manual tasks, but also more analytical and interactive tasks. Architects, finally, are one of the most analytical occupations according to our measure, while doing a substantial amount of interactive tasks. Tables C.1, C.2 and C.3 in the appendix contain a complete list of the task intensities for each occupation. As examples for the 'task distance' between occupations, the two most distant occupations according to our classification are 'mechanical and motor engineers' and 'household cleaners', whereas the two closest occupations are 'carpenters' and 'plasterers'.[6] Figure C.1 further illustrates the task variables by showing their distribution across occupations. Table 3.2 reports the correlation coefficients of the task intensities across occupations. In particular, there is a fairly strong positive correlation between analytical and interactive task intensities, whereas routine and manual tasks are negatively correlated with analytical task content. Note that occupations with a higher earnings potential on average involve more analytical and interactive tasks, whereas routine and manual tasks correlate negatively with the earnings potential. The correlation coefficients between the median earnings of (childless) women and the task intensity are 0.82 for analytical, 0.66 for interactive, -0.33 for routine and -0.38 for manual tasks.

[6] We use the sum of the squared differences between tasks as a measure of occupational 'distance'. For the main analysis, we also used the absolute value of the differences in task intensities and obtained similar results.

Table 3.2: Correlation of task intensities across occupations

	Analytical	Interactive	Routine	Manual
Analytical	1.00	0.73	-0.31	-0.31
Interactive		1.00	-0.53	-0.14
Routine			1.00	0.11
Manual				1.00

To build further intuition about how task intensities vary across occupations, it is also useful to consider the mean task intensities for broader occupational groups. In table 3.3, we group the 120 occupations into 5 'sectors' and report the mean task intensities for each sector. Workers in the sector 'Manager/Professional/Technicians' (MPT) perform the most analytical tasks, followed by 'Sales/Office' and 'Care'. The across-sector differences in interactive task content are smaller and only the 'Production/Operators/Craft' (POC) sector involves much less interactive task content than the others. Routine work is most prevalent in the POC sector. Manual task content is also high in this sector, but also important in 'Services' and 'Care'.

Table 3.3: Tasks by occupational sector

	Analytical	Interactive	Routine	Manual
Manager/Professional/Technician	0.34	0.36	0.17	0.13
Sales/Office	0.23	0.42	0.22	0.14
Production/Craft	0.10	0.19	0.42	0.29
Service	0.12	0.31	0.21	0.37
Care	0.20	0.31	0.18	0.31

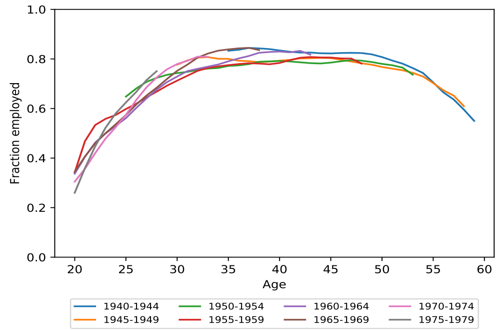
3.3 Life-Cycle Profiles of Employment and Task Intensities

We start our analysis by discussing the women's career profiles over the life-cycle and their relationship with fertility. While our main contribution is to document the patterns for task intensities, we first briefly review the life-cycle profiles of employment and hours, which provides some context for the analysis of tasks and occupations.[7] The focus of the section is to contrast men and women and document the relationship with women's completed fertility.

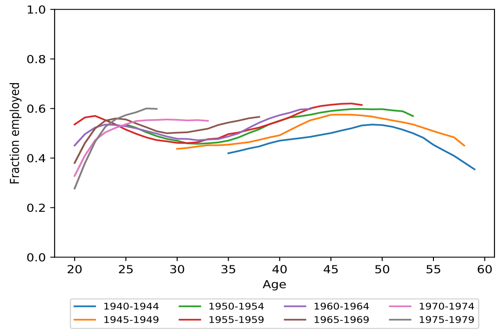
[7] In the appendix, we also show the life-cycle profiles of yearly earnings (figure C.3). These broadly reflect the patterns for employment on the extensive and intensive margin, which we discuss here.

Figure 3.1: Participation rates by cohort

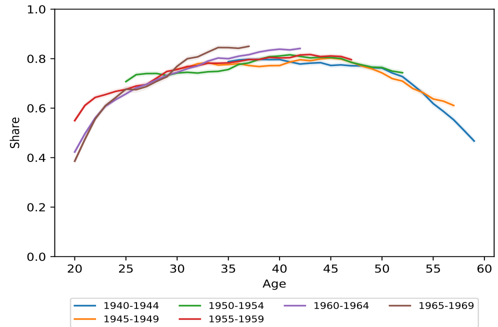
(a) Men



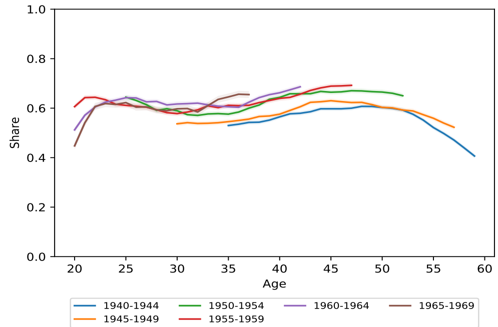
(b) Women



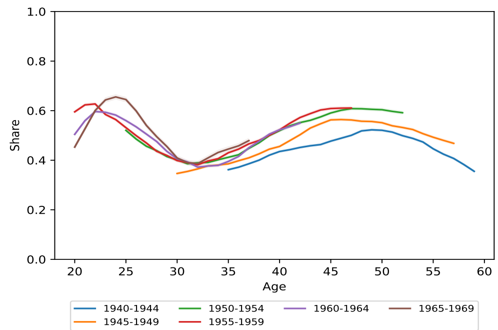
(c) Women (no children)



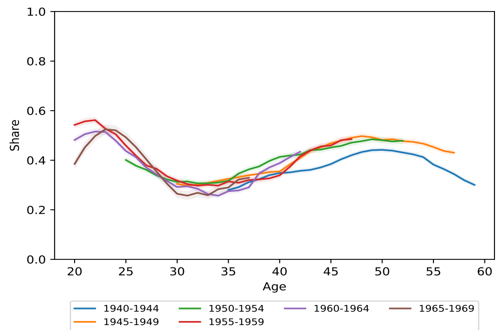
(d) Women (one child)



(e) Women (two children)



(f) Women (three or more children)



Since both women's employment and their task intensities have changed substantially over time, it is crucial to distinguish different cohorts for this comparison.

3.3.1 Employment: Extensive and Intensive Margins

Figure 3.1 shows the lifecycle of men's and women's employment for eight cohorts born during five-year intervals between 1940 and 1979. For the five youngest cohorts, we start at age 20; the oldest age that we consider is 60, for the cohort born in 1940-1944. The number of observations underlying these figures is reported in tables C.5 and C.6. For the most part, we will focus our discussion on ages 25-55, which are less affected by trends in education and early retirement. The upper left panel shows the employment rates for men. Here, there is relatively little change across cohorts, at least for the prime working years. About 80% of men in West Germany are employed between ages 30 and 50, before employment rates begin to decline.[8]

The picture looks very different for women (panel (b)). After a first peak in the participation rate when the women are in their 20s (the exact age depends on the cohort, with younger cohorts peaking later), there is a marked decline in participation which, as many studies have shown and the remaining panels confirm, is related to child-bearing. For women without children (panel (c)), the lifecycle pattern of labor market participation is almost identical to men's. Women who have one child over their lifetime exhibit only a modest reduction in the participation rate (panel (d)), while for women with two (the modal case, panel (e)) or three or more children (panel (f)) there is a strong and sustained drop in participation. For instance, women from the 1955–59 cohort with two children reach the maximum participation rate of about 60% in their early 20s and reach this level again only in their late 40s, and women from the same cohort with three or more children never return to the participation rate of their early 20s.[9]

Looking again at all women (panel (b)), a comparison across cohorts suggests that the middle cohorts (1950–1964) show a fairly similar pattern, while the oldest cohort (1940–1944) stands out for having much lower participation rates past age 35 and the 1945–1949 is in-between. This is not only a result of the decline in fertility rates (see C.2 in the appendix) and also holds for a given number of children (panels (d)–(f)). Starting with the 1965–1969 cohort, the reduction in the participation rate sets in later and becomes less pronounced; for the two youngest cohorts, no such drop is visible at all.

The gender differences are equally striking for the intensive margin of employment. As shown in Figure 3.2, the share of West German women working full-time (conditional on being employed) rapidly drops during the prime child-bearing years and remains constant at about 60% after age 35 or so. This is in stark contrast with men for whom employment nearly always takes the form of full-time work. Interestingly, comparing women of different cohorts shows that women born between 1965 and 1974 tend to work full-time more often than women from earlier cohorts during their 20s, but converge to the same low levels of full-time employment in their mid-30s as women from older cohorts. Moreover, the youngest cohort born in 1975–1979 is actually less likely to work full-time in their early 20s than women born between 1970 and 1974.

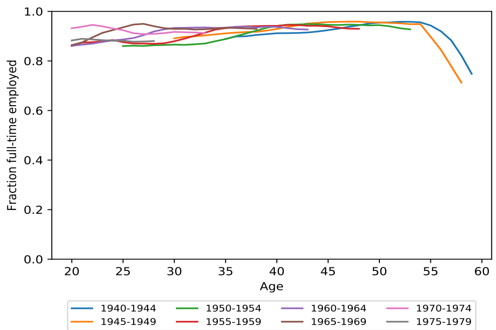
The life-cycle patterns of employment and hours have received a considerable amount of attention in the literature. McGrattan et al. (2004) focus on hours of work in the US. Since

[8]Note that the employment profiles are affected by the gaps due to self-employment, public service and living abroad, which makes it likely to underestimate employment shares.

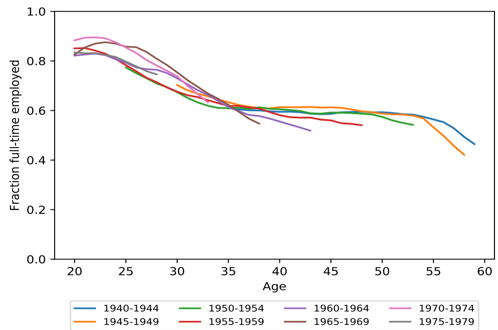
[9]Note that our data contains fertility information only for women, so that we cannot break down life-cycle employment by the number of children for men. However, since the data from the Socio-Economic Panel (SOEP) shows that the reaction of men's employment to the arrival of children is very small (see Kleven et al. (2019)), so that one would not expect noticeable differences in the life-cycle profiles by the number of children.

Figure 3.2: Full-time employment conditional on being employed

(a) Men



(b) Women



our data only contains a binary hours variable, it is easiest to compare our results to papers that focus on participation rates over the life-cycle. Goldin and Mitchell (2017) review cross-cohort differences in life-cycle employment rate in the US. While the shape of these profiles is broadly similar to the German case, the level of the participation rate tends to be lower in Germany than in the US.[10] Relative to other work on Germany (such as Fitzenberger et al. (2004)), the BASiD data provides longitudinal information on both employment and fertility and covers a large sample, which allows the breakdown by fertility and cohorts.[11]

3.3.2 Occupations and Tasks

We begin by contrasting the task content of occupations for men and women (Figure 3.3). Since women's task intensities have changed substantially across cohorts, it is important to consider different cohorts. Cohort by cohort, women have worked in occupations that were more analytical, more interactive, and had higher demands in terms of (non-routine) manual tasks. At the same time, there has been a shift away from occupations that featured a high share of routine tasks. As the comparison with men shows, this pattern is not merely driven by a general trend towards the automatization or outsourcing of routine tasks: if anything, men in West Germany work *more* often in routine occupations in the younger cohorts, and the highest values for both analytical and interactive tasks were reached for the cohorts born in the early 1950s. Overall, however, the task content of occupations has changed remarkably little for men over the cohorts studied here, in stark difference to what we observe for women. While men and women in the 1940–44 cohort worked in occupations

[10] Goldin and Mitchell (2017) distinguish between an 'old' and a 'new' life-cycle pattern of employment. In the 'old' pattern, women's employment rises until women reach their 50s and then starts to decline (a hump shape), wheras the 'new' pattern means that employment starts out high, somewhat drops in the age window in which women have children, and then rises again. The cohort profiles for Germany mostly reflect the 'new' pattern, since employment tends to drop in the beginning of the life-cycle also for older cohorts, such as the 1950-54 cohort. Cohorts older than that are only observed from a later age onwards.

[11] Fitzenberger et al. (2004) use repeated cross-sections from the German Microcensus to study participation profiles. An interesting aspect of the BASiD data is that we can condition on completed fertility and use information on the *past* labor supply of these groups to construct the life-cycle profiles.

that were fairly similar in the extent to which they required interactive tasks, women now work in more interactive occupations than men. At the same time, they have not only closed the large gap in analytical tasks, but have slightly surpassed men in the more recent cohorts. For non-routine manual tasks, the gap has been reduced, but here men still dominate by a fairly large margin.

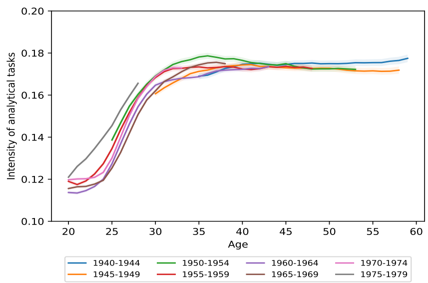
We now turn to the relationship between tasks and completed fertility. Figure 3.4 shows how the analytical task content of occupations differs by the number of children. We focus on only three cohorts here: women born in 1940–1944, 1950–1954, and 1960–1964. In each case, there is a negative relationship between the number of children and the analytical task content. From age 35 (the only age at which all three cohorts can be observed), in all cohorts women with children have lower values than childless women, and the gap increases in the number of children. However, the gaps have been reduced in the younger cohorts: while all women increasingly work in more analytical occupations irrespective of the number of children (which shows that the increase is not simply driven by changes in fertility), the increase has been more pronounced for women with children than for childless women.

The two younger cohorts also allow us to get a first glance at potential selection effects, i.e. differences in the analytical tasks contents of occupations that already appear at early ages, before most women have children. Figure 3.4 suggests that there are some differences before the prime child-bearing years (except for the case of the 1960-64 cohort), but that these differences are smaller than the gap that opens up at later ages. Interestingly, for the two younger cohorts, the relationship with the number of children is not monotonic at all ages: women with two children first work in more analytical occupations than women with one child in the beginning of their life-cycle and this pattern reverses at later ages.

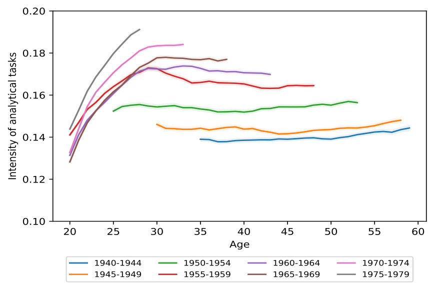
A very similar picture emerges when not the analytical task content, but the share of women working in managerial, professional, or technical occupations or the interactive task content are considered (cf. Figures C.4 and C.5 in the appendix). For interactive tasks, the gap between childless women and women with one or two children has become very small in the youngest of the three cohorts, while women with three or more children are still far behind. We also considered differences in task intensities by the total length of the (child-related) career break, which could be more directly related to labor market outcomes than the number of children per se. The total length of the break is defined as the number of years a woman spends in non-employment after the birth of the first child. The patterns are very similar to considering differences by the number of children and shown in Appendix Figure C.6. While these comparisons are suggestive of a relationship between tasks and completed fertility, we now turn more closely to the changes in task intensities around first childbirth.

Figure 3.3: Task content by cohort

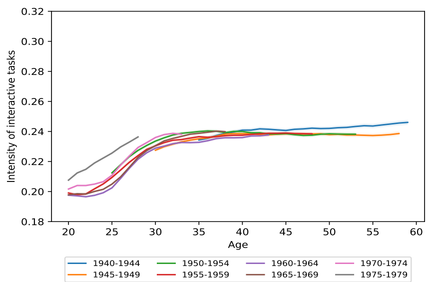
(a) Analytical (men)



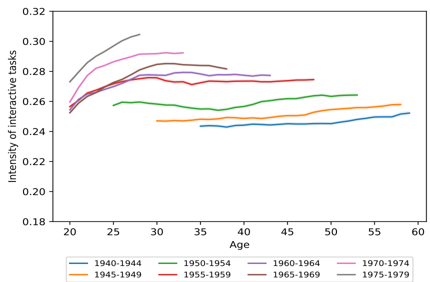
(b) Analytical (women)



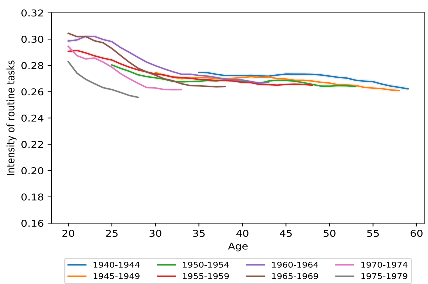
(c) Interactive (men)



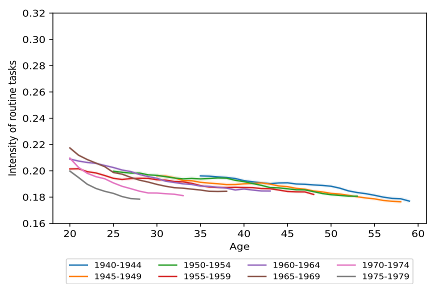
(d) Interactive (women)



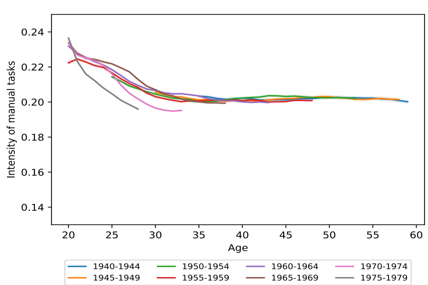
(e) Routine (men)



(f) Routine (women)



(g) Manual (men)



(h) Manual (women)

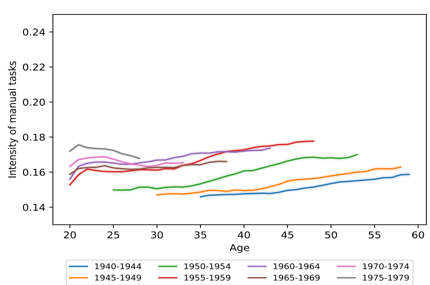
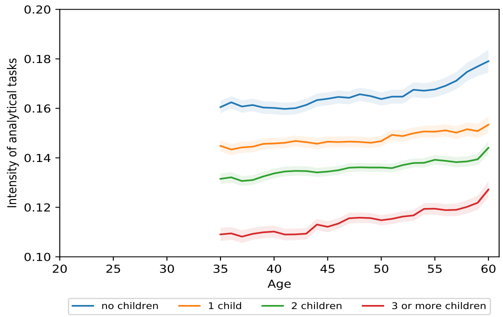
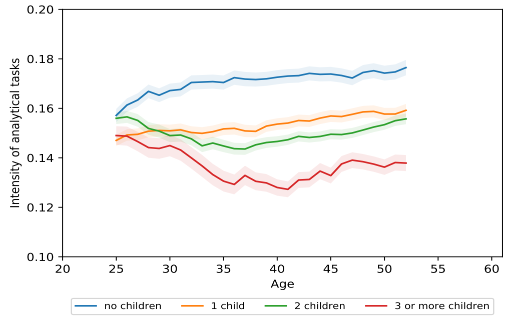


Figure 3.4: Analytical task content by cohort and completed fertility

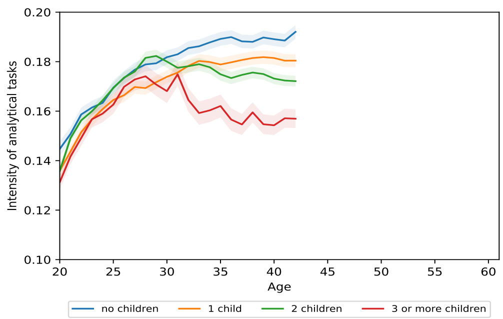
(a) Cohort 1940-1944



(b) Cohort 1950-1954



(c) Cohort 1960-1964



3.4 Dynamics around Childbirth

3.4.1 Estimation of Child-Related Effects

To investigate the role of changes around childbirth for the life-cycle patterns discussed previously, we use the event-study approach, which have often been used in the literature to investigate the impact of family-related events (see for example Kleven et al. (2018), Angelov et al. (2016), Fernández-Kranz et al. (2013), Mazzocco et al. (2013) or Attanasio et al. (2008)). Event time is denoted by t, and $t = 0$ is the quarter in which a woman has her first child. For different outcome variables Y_{ist}, we estimate the following equation:

$$Y_{ist} = \sum_{t \neq -1} \alpha_t \cdot EVENT_{it} + \sum_j \beta_j \cdot AGE^j_{is} + \sum_s \gamma_s \cdot YEAR_s + \nu_{ist}, \tag{3.1}$$

where α_t represents the deviation in the event-time t, i denotes individuals, s is calendar time, $EVENT_{it}$ is an event-time dummy, AGE^j_{is} is an age dummy for being j years old in year s, and $YEAR_s$ is a year dummy. Omitting the contribution of $\hat{\alpha}_t$ we get a t-year-age counterfactual $\hat{Y}_{t,s,j}$, and consequently a t-expected counterfactual $\hat{Y}_t = \mathbb{E}[\hat{Y}_{t,s,j}|t]$. The estimated impact of children is then defined as $\hat{P}_t = \hat{\alpha}_t / \hat{Y}_t$. The child-related effect represents the mean deviation of the outcome variable in event time t relative to the average before childbirth. In estimating this equation, we focus on the variation within the group of mothers.[12] Our sample consists of mothers who are observed between 5 years before the birth of their first child and 10 years afterwards. As a result, our analysis focuses on births that occured between 1980 and 1997. Table C.7 shows some summary statistics for the event study sample. In particular, the table shows the educational composition of the sample, which mainly consists of medium and low-skilled women.

Before turning to the task intensities, we first discuss the impact of children on employment, earnings of working hours. These are especially large in Germany, more so than in most other countries, as has recently been pointed out by Kleven et al. (2019), who estimate the effects of children on employment and earnings based on survey data from the SOEP.[13] They find an average long-run effect on earnings (including non-employment) of -61% for Germany, -31% for the United States and -21% for Denmark. In the following, we add to these findings by analyzing the role of the number of children, which shows that especially the effect on participation increases quite drastically in the number of children.

These results are shown in Figure 3.5. Unless indicated otherwise, all confidence bands refer to the 5% level. Panel (a) first shows the striking differences in the child-related effects on employment. Women who have only one child reenter the labor force much more quickly than the other groups. After 10 years, their employment rate is only reduced by 20%.[14] For women who have 2 children, the corresponding effect is 50%, which is more than twice as large. Women who have 3 or more children even reduce their labor force participation by around 70% 10 years after childbirth. Comparing the two latter groups, an important difference is also that women with two children start reentering the labor force around 6 years after the birth of the first child, so that the estimated effects begin to show an upward trend again. For women with 3 or more children, the effect is much more persistent.

Panel (b) focuses on the intensive margin and shows the effects on the probability of working full-time conditional on being employed. The long-run impact is between -42 and -58%. Interestingly, the differences by the number of children become much smaller than for participation. In other words, while women with one child take up employment earlier than the other groups, they also largely work in part-time jobs.

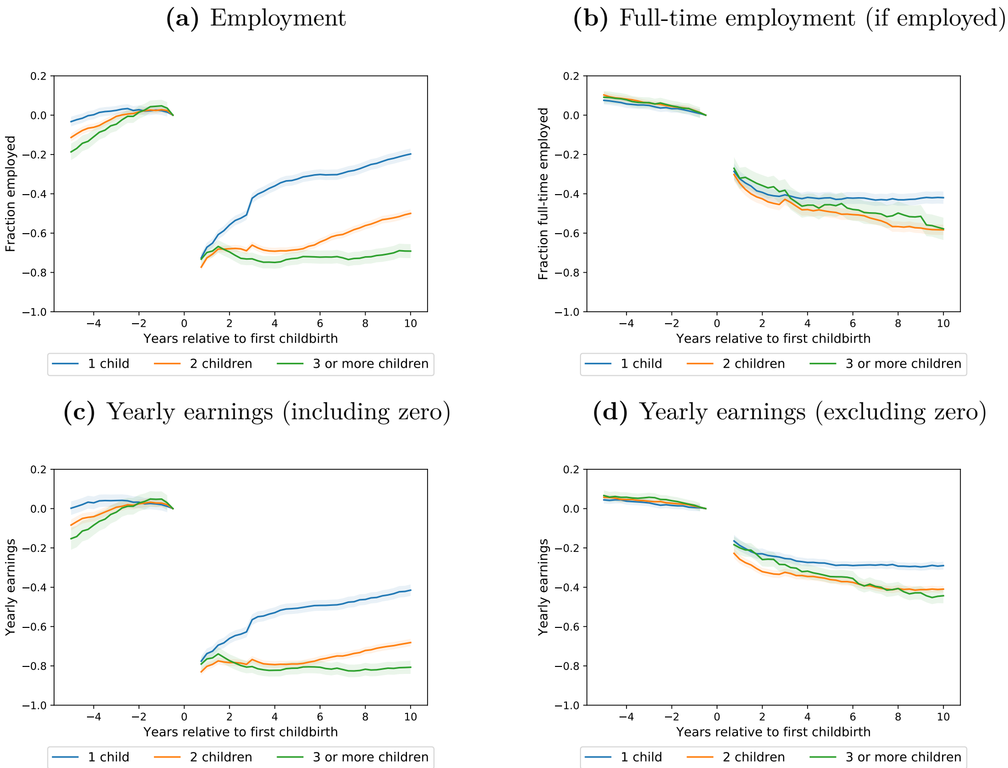
Panels (c) and (d) show the effects on yearly earnings, first for the case of non-employment being included as zero earnings and then for the case without non-employment. Comparing this to the previous results, the figures show that the effects on earnings strongly reflect the estimated effects on the extensive and intensive margin of employment. As a result, there

[12] Since the event time variable is only observed for mothers, we exclude childless women in this event study. Including childless women requires assigning 'Placebo births', as we will discuss in more detail in the next section.

[13] The SOEP is also useful for comparison because fertility information is observed for men. Kleven et al. (2019) find that the child-related effect on the earnings of men is slightly negative and close to zero.

[14] Note that we do not observe public servants and self-employed individuals in our data. This could in principle affect these estimates if the proportion of women who become civil servants after childbirth is substantial. However, supplementary analyses based on the SOEP, where these groups are observed, indicate that this does not have a substantial effect on the estimates.

Figure 3.5: Event studies of employment and earnings



are large differences by the number of children when non-employment is included (panel (c)) and smaller differences when considering working women only (panel (d)).

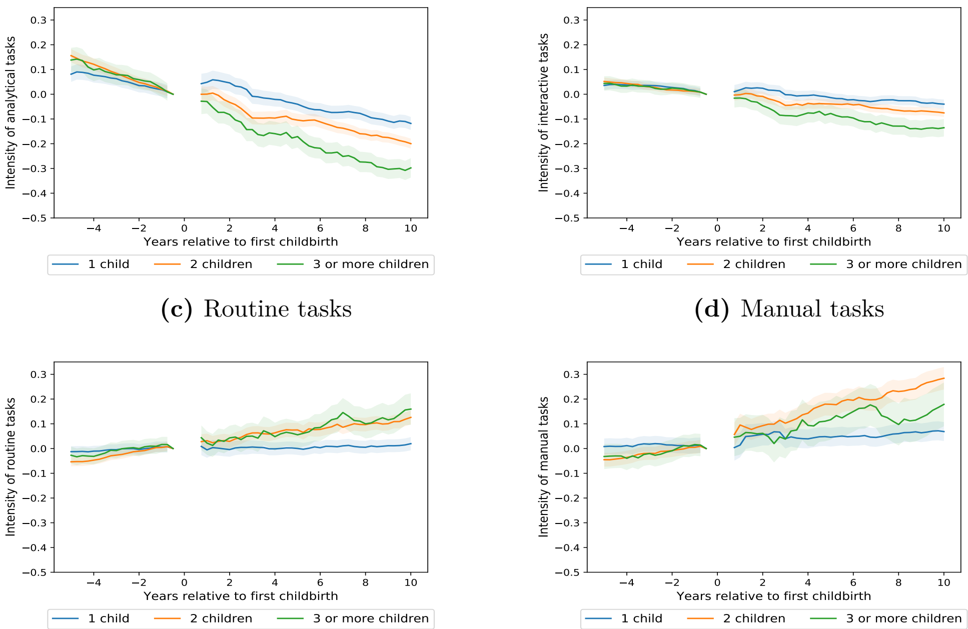
The negative impact of childbirth on earnings could potentially partly be explained by changes in task intensities. As a first step in the analysis of task intensities, we estimate equation 3.1 for each of the task variables. The results are shown in figure 3.6. For the analytical task intensity, the estimated effects are negative and grow over time. Having more children amplifies the effects. In interpreting the sign of the effect, it is important to keep in mind that equation 3.1 contains age and year effects. Thus, a negative event time coefficient implies that the task intensity is lower than would be predicted based on the age and year effects, rather than that the mean task intensity is lower than before childbirth. In other words, a negative coefficient for analytical task intensities suggests that the task intensity would be higher in the absence of children. An additional important issue for the interpretation is the role of selection, which we will discuss in more detail below. Importantly, a negative effect of having children on analytical task intensities could be reflected in wages and contribute to the gender wage gap. We return to this issue in section 3.4.2.2.

The pattern for interactive tasks is quite similar, although its magnitude is smaller. Again, the effects are increasing in the number of children and women with three or more children show the strongest long-term change in interactive task content. The routine task intensity increases for the groups of women with two and at least three children and is essentially constant for women who have only one child. For manual task intensities, women with only

Figure 3.6: Event studies of task intensities

(a) Analytical tasks (b) Interactive tasks

(c) Routine tasks (d) Manual tasks

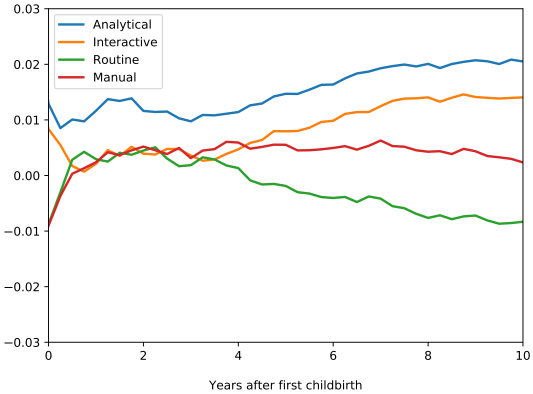


one child show a small increase. The effect is strongest for women with two children, and somewhat smaller again for women with at least three children.

As already noted, the estimated coefficients take both occupational switches and selection due to some mothers exiting the labor force into account. To get a sense for the role of the selection effects and their direction, we computed the mean task intensities that women had in their last job *before childbirth* for mothers who work or stay at home at each point in event time. Figure 3.7 shows the *difference* in the pre-birth task intensities for working and non-working mothers. For example, the figure shows that mothers who work *after* 10 years had a higher analytical task intensity (by roughly 0.02) *before* childbirth than mothers who stay at home *after* 10 years. In other words, women in more analytical jobs before childbirth are more likely to stay in the labor force after birth. Interestingly, this selection pattern goes in the other direction than the estimated child-related effect from figure 3.5, where we find a negative effect on analytical task intensities. The selection effect alone would lead to a *positive* effect, since it is mothers in more analytical occupations who return to work at a higher rate. Working mothers had higher interactive tasks and lower routine task intensities than mothers who do not work. There is little selection on manual task content. Given the presence of these selection patterns, it is important to also isolate the role of occupation switches. To do so, we now turn to analyzing the within-individual changes in task content.[15]

[15]Note that it is not straightforward to directly correct for selection in the regression equation 3.1, since this would require an instrument or a structural model, which provides the counterfactual task intensities for women for whom they are not observed.

Figure 3.7: Difference between pre-birth task intensities for working vs non-employed mothers



Notes: The figure shows the difference in the mean task intensities *before* childbirth for mothers who work and mothers who stay at home at each point in time *after* childbirth. Thus, a positive (negative) value corresponds to working mothers having a higher (lower) task intensity along this dimension before birth.

3.4.2 Occupation Switches

3.4.2.1 Placebo Event Studies

We now focus on occupation *switches* following childbirth. To what extent do women increasingly switch occupations after childbirth and how do the characteristics of the occupations change? The main purpose of this analysis, as discussed in the previous section, is to shut down the role of selection due to non-employment and only consider the within-individual variation for all women who work at a given point in time.[16]

To create a control group for the analysis, we assign a 'placebo birth' to each childless woman by drawing the age at first birth from the empirical distribution of births of the group of mothers from the corresponding education group, following the approach in Kleven et al. (2018). This allows to study the labor market outcomes of childless women in event time relative to the date of their placebo birth. The choice of childless women as a control group deserves some discussion, since having children is an endogenous outcome and childless women might differ from mothers in certain dimensions. In the following, we find that both groups are on fairly similar trajectories before the (placebo) birth (the pre-trends are similar), which suggests that childless women are a good control group. While one could also use men as a control, this would similarly require assigning placebo births, since we observe fertility only for women in our data. We restrict the sample to the cohorts born between 1955 and 1969, to come close to completed fertility.

Figure 3.8 compares the changes in outcomes for mothers and childless women and Table 3.4 summarizes the magnitude of the differences. Panel (a) of the figure first presents the fraction of employed individuals from the group of mothers and the placebo group that work

[16]More concretely, we analyze changes in task intensities relative to the task intensity before childbirth (Δt_{it}) rather than the task intensity (t_{it}) directly.

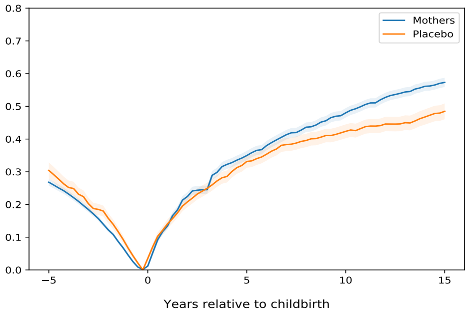
in a different occupation than before childbirth. By construction, the fraction is close to zero before childbirth for each group. After childbirth, mothers switch more often than women in the placebo group. For example, after 10 years, 48% of working mothers are in a different occupation than before childbirth, while it is 42% of non-mothers. Thus, mothers on average switch 6 percentage points more often. As will be shown in Figure 3.10, the difference in the switching rate grows with the number of children and the length of the career break - in particular, women with very long career breaks switch occupations markedly more often than childless women.

To assess the implications of this increased occupational switching rate, it is important to consider additional characteristics of the occupations, such as tasks, the earnings potential or the amenity value in terms of family-friendliness. On the one hand, switching could be a sign of a worsening labor market situation and contribute to a wage loss. This would for example be the case if mothers increasingly 'downgrade' their occupation and return to work in less skill-intensive occupations than they had worked in before childbirth. There are several theoretical mechanisms that could lead to such downgrading. For example, models of human capital depreciation (such as Adda et al. (2017)) predict that the skills in the pre-birth occupation deteriorate over time and at some point, it might become hard to reenter to the previous occupation. In addition, search frictions could prevent women from reentering in their old occupation, especially when the career break is long enough so that they cannot return to their old employer. Another potential reason for wage penalties resulting from switching is that human capital is often argued to be task-specific (see for example in Gathmann and Schönberg (2010) or Adda et al. (2017)). In such models, experience in working with specific tasks increases the future productivity of these tasks. If the new and old occupations differ strongly in the tasks that are required, women can only partially transfer the skills they acquired in their previous working life. This would generate wage losses especially for women who move to occupations with a very different task content. On the other hand, Groes et al. (2014) highlight that occupational switching per se does not necessarily lead to wage losses, since individuals can also progress to a 'better' ocupation with a higher earnings potential. In addition, it could be that the presence of children amplifies the need to work in a family-friendly environment, so that women could switch occupations to adapt to this. In the following, we use our data to investigate each of these points in more detail.

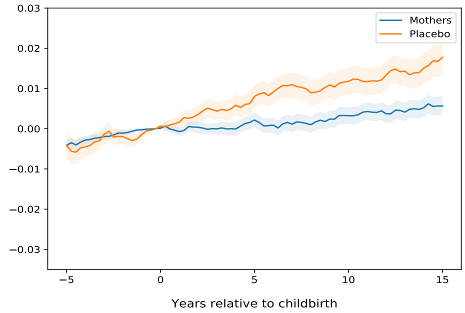
In panels (b) - (e), we investigate the issue of occupational up- and downgrading in terms of tasks. Skill-intensive and well-paying occupations tend to be more analytical and interactive and occupational downgrading would be reflected in switches towards occupations with lower task intensities along these dimensions. As discussed previously, this is closely related to the event studies which were shown in figure 3.5. However, we now use different variation by considering the difference between current and pre-birth tasks in order to isolate the role of of occupational switching. Each occupation o is characterized by a task vector $(t_o^a, t_o^i, t_o^r, t_o^m)$.

Figure 3.8: Occupational switches for mothers and childless women

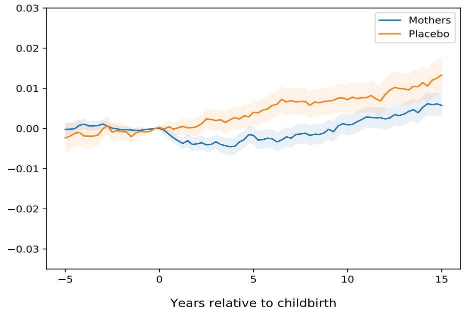
(a) Fraction of occupation switches



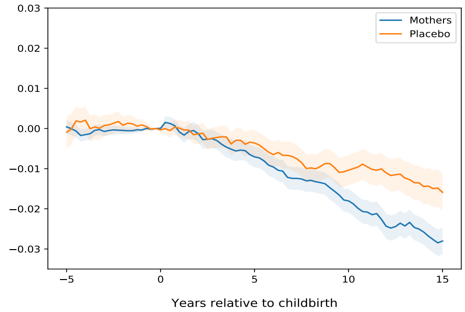
(b) Change in analytical tasks ($\Delta t^a_{i,t}$)



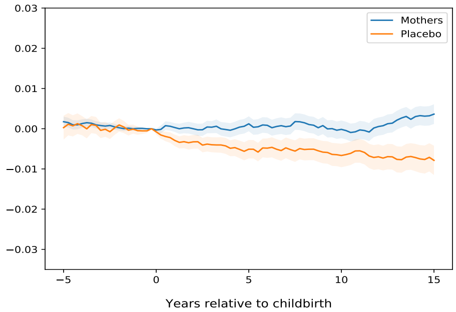
(c) Change in interactive tasks ($\Delta t^i_{i,t}$)



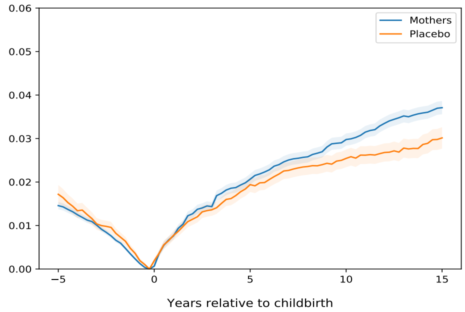
(d) Change in routine tasks ($\Delta t^r_{i,t}$)



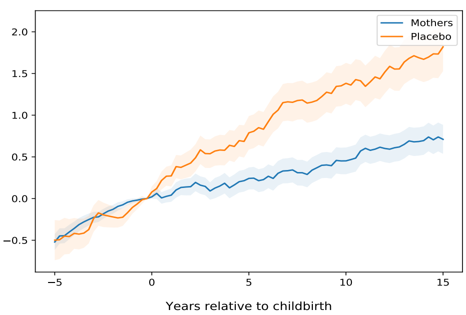
(e) Change in manual tasks ($\Delta t^m_{i,t}$)



(f) Distance from previous tasks ($\Delta T_{i,t}$)



(g) Change in yearly occ. earnings (in 1000€)



(h) Change in occupational part-time share

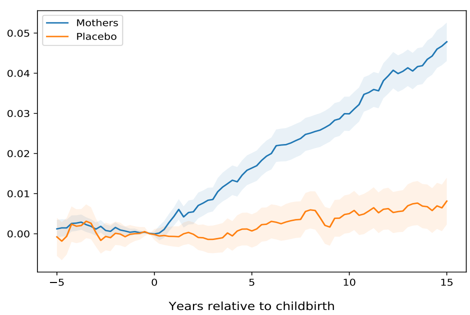


Table 3.4: Differences between mothers and childless women

	After 5 years	After 10 years	After 15 years
(1) Switched occupation	0.018	0.057	0.088
(2) Analytical	-0.006	-0.009	-0.012
(3) Interactive	-0.006	-0.006	-0.008
(4) Routine	-0.003	-0.008	-0.012
(5) Manual	0.006	0.006	0.011
(6) Task deviation	0.001	0.004	0.007
(7) Median earnings in occupation	-0.547	-0.930	-1.109
(8) Share of part-time work in occupation	0.016	0.025	0.040

Notes: The table shows the difference between mothers and childless women (i.e. $y^M - y^C$, with M and C corresponding to the two groups) for each of the outcomes from Figure 3.8. The units of the outcomes are percentage points for row (1), task intensity units for rows (2), (3), (4) and (5), the squared task distance for row (6), yearly earnings (in thousand) for row (7) and percentage points for row (8).

Denoting o_{it} as the current occupation and $\bar{o}_i$ as the last occupation before childbirth, we compute the following outcome variables:

$$\Delta t^j_{i,t} = t^j_{o_{i,t}} - t^j_{\bar{o}_i}$$

Variation in $\Delta t^j_{i,t}$ purely reflects occupational changes, since the variable would always be zero if all women stay in their previous occupation. As a result, focusing on changes relative to the pre-birth occupation circumvents the issue of selection.[17] Figure 3.8 then shows the mean of $\Delta t^j_{i,t}$ for mothers and the placebo group at each point in event time. Panel (b) shows the results for analytical task content.[18] Interestingly, mothers do not directly downgrade their occupation, in the sense that they would start to work in occupations with lower analytical task intensities after childbirth. Instead, they slightly *increase* the task intensity relative to their pre-birth occupation. However, around childbirth, a gap emerges between mothers and childless women, since the latter group still increases the analytical task intensity on average by around 0.02, whereas the increase for mothers is marginal. To put the magnitude of this change into perspective, it is useful to compare it to the mean task intensity before childbirth. For childless women, this is 0.16. Thus, the average growth after 15 years for the placebo group is 12.5% as large as the initial level. Assuming that mothers

[17] Thus, there is no selection at each point in event time, since we only look at women who work in this event quarter and before childbirth. Note that the composition of the labor force still changes between different points in event time, since more and more women reenter over time.

[18] As a robustness check regarding the definitions of the task intensities, we also replicated these results for alternative task measures, where task intensities are rescaled so that they sum to 1 for each occupation. These results are shown in figure C.9 and broady similar. Mothers switch less to analytical and to less routine-intensive occupations and to more manual-intensive ones. There is no noticeable difference in interactive tasks between the groups for the rescaled task measures.

would have faced the same career trajectory as childless women, these results suggest that the 'penalty' of motherhood with respect to analytical tasks mainly operates through the effect that mothers would have been able to further increase their task intensity if they had decided not to have children.

For interactive tasks, the point estimates suggest a small negative effect of motherhood on task intensities. However, the confidence bounds, which are computed for the 5% level, are relatively wide. Panel (d) and (e) finally show the event studies for routine and manual tasks. Mothers reduce their routine task content more strongly than the placebo group. The manual task intensity stays roughly constant, while the placebo group reduces its manual task content. Note that the magnitude of these changes is relatively modest, compared to the differences in task intensities across occupations. In general, with the exception of interactive tasks, the difference between the two groups becomes more pronounced over time.

Panel (f) of Figure 3.8 shows the results for the 'distance' between the current and the old occupation. We compute the task distance between occupations as the sum of the squared deviations between current and pre-birth tasks.[19]

$$\Delta T_{i,t} = \sum_j (t^j_{o_{i,t}} - t^j_{\bar{o}_i})^2$$

The results show that mothers overall work in occupations that are more distant from their pre-birth occupations, particularly in later periods (such as 10 or 15 years after childbirth). Interestingly, the effects closely mirror the increased rate of occupational switching, which was shown in panel (a). This suggests that the increased task distance relative to the pre-birth occupation is largely due to a higher number of occupational switches, rather than moving towards more different occupations conditional on switching. Note the this task *distance* is conceptually different from the task *changes* analyzed in panels (b) - (e). For example, if equally many women increase and decrease their analytical task intensity, it could be that the mean change is zero, but the mean distance is non-zero. From the perspective of the model in Gathmann and Schönberg (2010), the task distance reflects the matching between the previously accumulated human capital and the current tasks. This implies that having children could not only affect occupational choice in terms of moving between 'low-wage' and 'high-wage' (or 'skilled' and 'unskilled') occupations, but also the allocation of individual skills to tasks, which is a unique implication of the task framework. In other words, also a move between two 'unskilled' occupations could be wage-reducing if the tasks in the new occupation differ from the skills of the individual.

To complement our analysis of task intensities, the remaining two panels show additional characteristics of occupations. First, we investigate whether women switch to occupation

[19] As a robustness check, we also compute the distance using the absolute rather than the squared value and the results are very similar. The squared difference attaches a higher weight to large deviations from the pre-birth tasks.

with a higher or lower average earnings potential. As a proxy for the earnings potential in an occupation, we compute the median (annual) earnings of all women working in this occupation, focusing on childless women to reduce the impact of hours choices. Panel (g) shows how the occupational earnings rank changes for mothers and childless women. As in the previous analysis, the outcome is expressed relative to the earnings in the last occupation before childbirth, so that changes only reflect occupation switches. The graph shows that both groups switch to better-paying occupations over time, but this increase is more pronounced for the placebo group. Thus, interestingly, there is again little direct occupational downgrading, but rather a stagnation relative to childless women. Childless women on average still advance to occupations where the median (annual) earnings are around 2000€ higher after 15 years, while mothers only experience a lower increase.

Finally, we consider a proxy for the family-friendliness of an occupation, which we define as the fraction of women who work part-time in an occupation. Occupations in which this fraction is low could be harder to combine with part-time work and might thus be seen as less family-friendly. Panel (h) shows the results for the measure. This stays essentially constant for the placebo group, while mothers switch to occupations with a higher fraction of part-time work. Thus, taken together, these findings suggest that mothers switch less to higher paying occupations (panel (g)) and more to occupations in which part-time work is prevalent (panel (h)). Thus, our results are consistent with a trade-off between the earnings potential of an occupation and its family-friendliness, which has recently been suggested by Goldin (2014). While Goldin (2014) discusses the cross-sectional differences in occupations in terms of their family-friendliness, there is not much direct evidence so far on the role of *switches* towards family-friendly work environment following childbirth. Kleven et al. (2018) find that mothers become more likely to work in the public sector and in family-friendly *firms*, proxied by the presence of women with young children in the management, whereas we add evidence on *occupational* switching towards family-friendly occupations.

In light of this discussion, an important question is whether there is a systematic relationship between tasks and the family-friendliness of an occupation. If, for example, certain task combinations often occur in less family-friendly occupations and mothers increasingly switch to family-friendly occupations, this would explain part of the observed shifts in task intensity. As a simple way of investigating such a relationship, we calculated the correlation coefficients between the part-time share in an occupation and each of the task intensities. Interestingly, analytical task content is slightly *positively* correlated with part-time work (0.04). This suggests that is not necessarily more difficult to work part-time in more analytical occupations. The correlation is also positve and somewhat larger for interactive tasks (0.23) and close to zero for manual tasks (0.01). The association is strongest for routine work, which negatively correlates with part-time work (-0.53).

3.4.2.2 Implications for Wages

Given the previous results about changes in task intensities, an important question is to what extent they could be reflected in hourly wages. We perform a simple back-of-the-envelope calculation to assess how important the changes in tasks could be quantitatively, focusing on the increases or decreases in task intensities of mothers relative to childless women, which were shown in figure 3.8 (panels (b) - (e)).

To calculate the hourly wage rate in our data, we assume that individuals work 20 hours in part-time and 40 hours in full-time.[20] We then estimate the following equation:

$$\log(w_{it}) = \beta_0 + \sum_k \beta_k t_{it}^k + u_{it}$$

Based on the estimated coefficients for each of the task intensities, the effect of a task change Δt^k can simply be computed as follows:

$$\Delta w^k = (\exp(\beta_k) - 1)\Delta t^k \cdot 100$$

Here, Δw^k is the percentage change in the wage which is associated with the task change Δt^k.

Table 3.5: Implications for wages

	β_k	Δt^k	Δw^k
Analytical	1.970	-0.012	-7.466
Interactive	-0.230	-0.008	0.156
Routine	-0.010	-0.012	0.012
Manual	-0.307	0.011	-0.304

Notes: The table shows the results of the back-of-the-envelope calculation for wages. It reports the regression coefficient (column 1), the average change in the task variable after 15 years (column 2) and the predicted wage change (column 3). The wage change is expressed in percentage points.

Thus, this simple calculation suggests that analytical tasks have the most potential to contribute to effect of motherhood on hourly wages. The calculation suggests that the hourly wage rate of mothers would be 7.5% higher if they also increased their analytical task intensity like childless women. For the other task dimensions, the quantitative implication on wages is quite small, at least in the context of this simple calculation. This is mostly the

[20] Recall that the data only contains a binary hours variable, so that such an assumption is necessary. It implies that more fine-grained differences in hours are captured as a difference in the hourly wage.

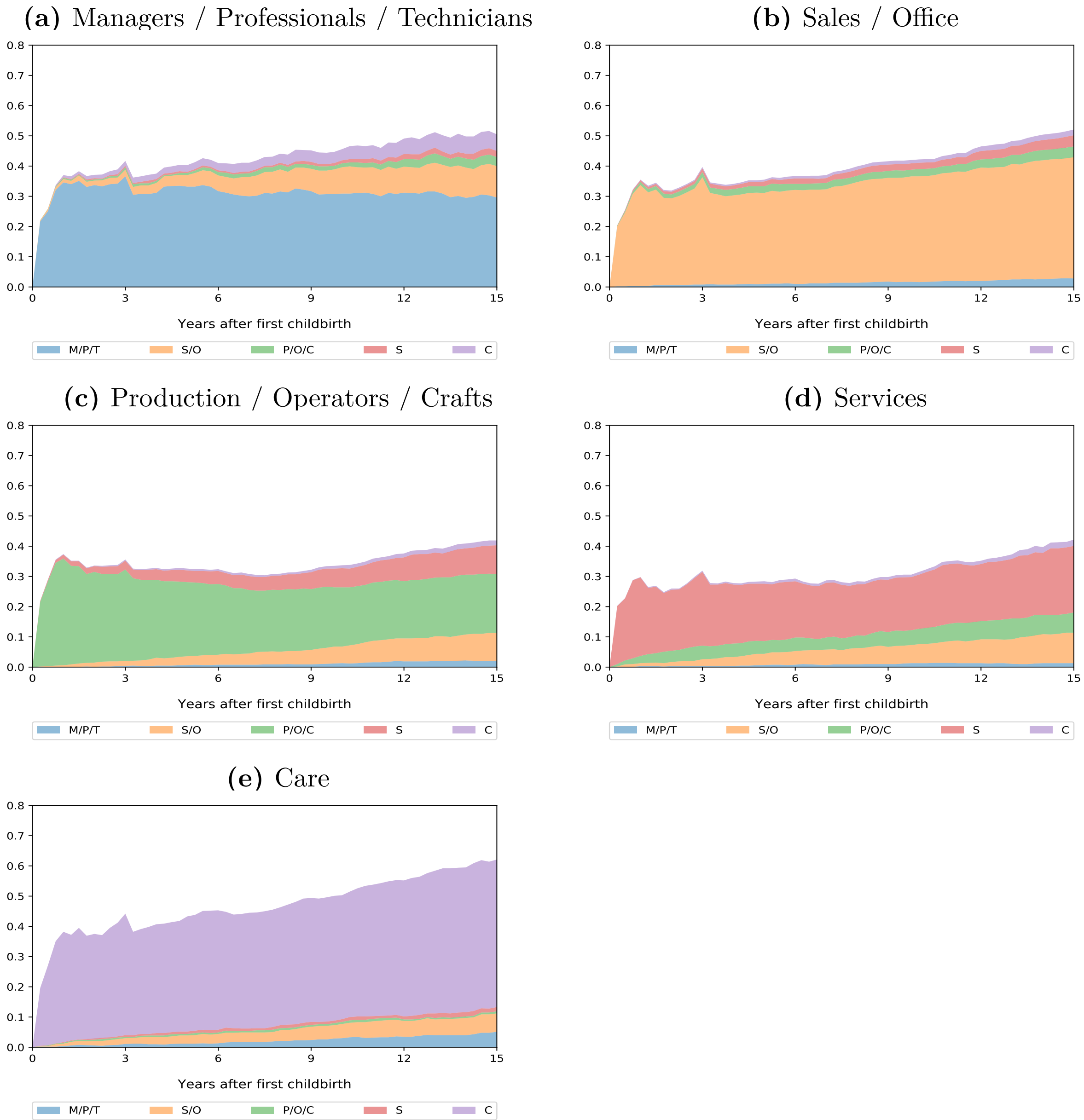
case because the estimated regression coefficients suggest that analytical tasks have a much larger impact on wages than the other tasks.[21]

3.4.2.3 Occupational Switching between Sectors

To better understand the occupational switches following childbirth and build more intuition about the types of switches that occur, we also grouped the 120 occupations into broader categories (manager / professional / technicians (MPT), sales / office, production / operators / crafts (POC), services and care) and studied switches between these five broad sectors directly. Table 3.6 summarizes the occupational switches relative to the pre-birth occupation ten years after childbirth and Figure 3.9 illustrates the transitions graphically and shows the dynamics in more detail. For each pre-birth occupation group, the columns of Table 3.6 report the fraction of women who work in each of the occupation groups (conditional on working). The diagonal reveals some heterogeneity in switching behavior across sectors. Women who work in 'Sales/Office' and 'Care' are least likely to switch (around 80% stay within their occupation group), whereas women working in the service sector and POC are most likely to switch (only around 60% stay in their group). Managers, professionals and technicians are somewhat in between these two cases, since 68% of women stay in the group. The table further shows which type of broad occupational moves are most common. In particular, there are noteworthy fractions of women switching from MPT to sales/office (19%), from production, operators and crafts to sales/office (16%) and services (19%), and from services to sales/office (20%) and POC (16%). Table C.4 shows the change in tasks for different occupational moves between these sectors. Given that the calculation from the previous section suggests that analytical task intensities have the largest implications for wages, it is perhaps most interesting to look at changes in analytical tasks - for example, a worker who moves from MPT to sales/office reduce the analytical task intensity by 0.11. Moving from POC to sales/office increases it by 0.13 and moving from services to sales/office increases it by 0.11.

[21] There are several important dimensions in which this simple calculation could be extended. In particular, it would be interesting to allow for skill accumulation and depreciation, heterogeneity in ability and sorting across occupations.

Figure 3.9: Fraction working in different occupations conditional on occupation before birth



Notes: The figure shows the transitions between occupational sectors for women who worked in a specific occupation group before childbirth. The figure includes transitions into non-employment.

Table 3.6: Occupational Switching - 10 years after childbirth

Post-birth Pre-birth	Man./Prof./Tech.	Sales/Office	Prod./Op./Crafts	Services	Care
Man./Prof./Tech.	0.68	0.19	0.03	0.02	0.08
Sales/Office	0.04	0.83	0.06	0.05	0.03
Prod./Op./Crafts	0.04	0.16	0.58	0.19	0.03
Services	0.04	0.20	0.16	0.57	0.03
Care	0.06	0.09	0.01	0.02	0.82

Notes: The table shows the transition rates between the 5 broad occupational groups. Each row corresponds to a pre-birth occupation. The columns contain the probability of working in each of the occupations 10 years after childbirth, excluding transitions into non-employment (i.e. it focuses on women who work both before childbirth and 10 years afterwards).

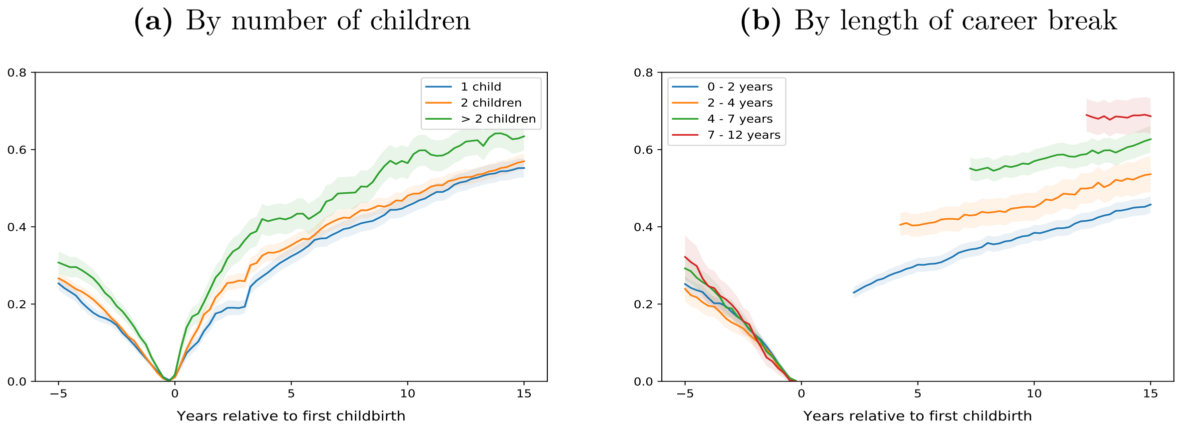
3.4.2.4 Heterogeneity

While the analysis of occupational switches so far has focused on the pooled sample of all women, there are several interesting dimensions of heterogeneity along which the effects could potentially differ. We also conducted the analysis separately by education group to see to what extent the average effect is driven by one of these groups. For example, it could be the case that highly educated women start their career in more analytical occupations, but downgrade the occupation after childbirth. Instead, we find that highly educated mothers, in terms of having a university degree, exhibit very small differences in their switching behavior from the corresponding placebo group (Figure C.10). The effects are largest for the medium education group, which consists of women with a high school degree (including A-levels/*Abitur*) or vocational training. The average results from Figure 3.8 mostly reflect the changes for this group of women. For the lowest education category, the effects are small again.

We further investigate the role of the number of children and the length of the career interruption. The theoretical arguments for why mothers might be more inclined to switch occupations - such as a higher need to work in a family-friendly occupation or human capital depreciation - are likely to be amplified by these variables. Note that the variables are interrelated, since a higher number of children typically generates longer career breaks, as was shown by Figure 3.5. In Figure 3.10, we show how the fraction of occupation switches is related to these variables.

Panel (a) shows the fraction of women who has changed their occupation relative to the pre-birth occupation by the number of children. Woman who have two or more children experience a larger switching rate than the other two groups, for which the switching probabilities quite

Figure 3.10: Occupational switches - by children and length of break



similar. The differences between the groups become larger when comparing mothers with different lengths of their career break, defined as the number of years that pass until the individual is observed to go back to work (panel (b)). Women with a very long break exhibit a markedly higher switching rate than women with short breaks.[22] This is in line with theoretical mechanisms such as increasing human capital depreciation or search frictions due to an increasing difficulty to return to one's old employer. Another interpretation would be selection, in the sense that taking a long career break could be related to having worse labor market prospects in the first place, which could also increase the likelihood to switch.

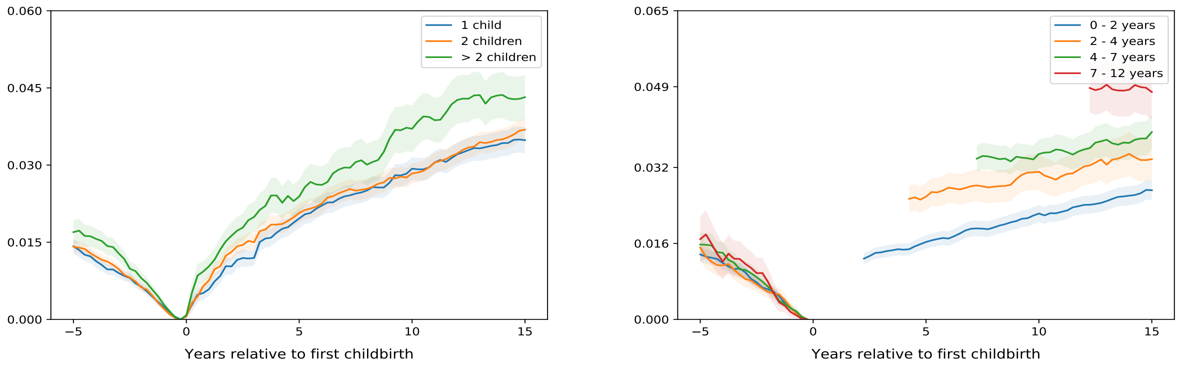
The higher rates of occupational switching, both for women with a higher number of kids or a longer career break, are reflected in a greater distance between current and pre-birth tasks, as is shown in Figure 3.11. Given the results from Gathmann and Schönberg (2010), such a task distance could lead to a reduction of the earnings potential of mothers, as they perform tasks they have less prior experience with. In Appendix Figure C.12, we also computed the task distance separately for each of the task intensities, to see whether the overall impact shown in Figure 3.11 is driven by a particular task dimension. The qualitative pattern that longer breaks tend to increase the task distance is present for each of the task intensities. The magnitude is strongest for routine and interactive tasks, although it is comparable for the other dimensions.

In Table C.11, we also look at up- and downgrading in the task content of the job. A first issue to note is that breaking down the sample into subgroups makes the confidence bounds for the task variables quite large. As a result, the figure shows the confidence intervals at the 10% level, although they are still relatively wide for both the breakdown by the number of children and by the length of the career break. For the number of children, the point estimates suggest that more children lead to slightly less analytical jobs, whereas there is less of a relationship between analytical task content and the length of the career break. Thus,

[22]Interestingly, mothers with very short career breaks (0-2 years) switch at a (slightly) lower rate than the placebo group.

Figure 3.11: Task distance - by children and length of break

(a) By number of children **(b)** By length of career break



interestingly, these results suggest that women with long career breaks do *not* downgrade their occupation in terms of the analytical task content.

Finally, in Figure C.13 we consider how the other occupation characteristics (the median annual earnings and the fraction of part-time workers) vary with the length of the career break. Another potential sign of occupational downgrading for women with a long career break would be if they switched to occupations with lower median annual earnings than the occupation they worked in before childbirth. Panel (a) shows the change of the occupational median earnings (as was also defined previously). There are no statistically significant differences (at the 5% level) by the length of the break. This is another indication that there is little direct occupational downgrading, since even women who take very long career breaks do not (on average) return to work in less well-paying occupations, but only experience less of an increase than childless women. For the part-time intensity of the occupation, the differences between the groups are larger. Especially women who who take the longest breaks (between 7 and 12 years) switch more to occupations in which a higher fraction of individuals works part-time.

3.5 Conclusion

This chapter has analyzed the relationship between fertility and women's careers, with a particular emphasis on the role of occupations and the tasks that women perform on the job. Our analysis was based on administrative labor market data from German social insurance records and detailed survey data on tasks that individuals typically perform in different occupations.

Our analysis started by comparing cohort profiles of participation, income and task intensities. In particular, we document that women have experienced a strong rise in analytical and interactive task intensities across cohorts, whereas the task intensities have changed little for men. However, there still is some divergence in tasks over the life-cycle, particularly for women with a higher number of children. We then consider the role of childbirth more explicitly by analyzing its impact in an event-study framework. The direct impact of childbirth on women's labor force participation is very large in Germany. Among the task intensities, analytical and manual tasks react strongest to childbirth. In particular, analytical task intensities increase less than would be predicted in the absence of children.

We finally consider the role of occupational switching in more detail and directly compare the occupational switches of mothers and childlesss women. While mothers switch occupations more often than childless women, their analytical task intensity increases less than for the comparison group. This can be interpreted as mothers missing out on some career progression relative to childless women. A simple back-of-the-envelope calculation suggests that the hourly wage rate of mothers would be 7% higher if they experienced the same increase in the analytical task intensity as childless women. Interestingly, there is little evidence for direct occupational 'downgrading' after childbirth, even for women with long career breaks, which would mean that women return to the labor force in occupations with a lower earnings potential or lower skill requirements. Another interesting aspect is that mothers switch to occupations which differ more strongly from the task requirements of the occupation they had before childbirth, particularly those who take a very long career break. In the light of theories of task-specific human capital accumulation, this could imply that mothers can transfer less of their previously accumulated human capital to the new occupation. In addition, mothers tend to switch to occupations in which a higher fraction of women works in part-time, which is consistent with an increased need to work in a family-friendly environment.

There are some interesting directions in which our analysis could be extended. In particular, it would be interesting to consider a more fine-grained classification of tasks, which would allow to study more closely whether there are certain tasks, which are per se less suitable for part-time work. For example, our definition of interactive tasks includes both 'selling', which is often done in part-time work, and 'supervising others', which one would expect to be more common in full-time work. This more fine-grained classification would allow to shed more light on the relationship between occupational tasks and family-friendly jobs.

In addition, the simple back-of-the-envelope calculation about the potential implication for wages could be extended to a more realistic empirical model which captures, for example, the accumulation of experience. This would help to further quantify the wage impact of changes in the task composition. Finally, it would be interesting to analyze the patterns documented in this paper through the lens of a structural model, which would allow to assess the welfare implications of occupational switching.

Appendix A

Appendix to chapter 1

A.1 Numerical Solution of Model

In this section, we outline the algorithm used to solve for the equilibrium of the model.

General approach. We start by guessing a matrix of hiring rates $g_j(t)$. Given these values and the functional forms described in Section 1.4, we can solve the agent problem backwards. In each period, the optimal level of search intensity has a closed-form solution:

$$s_{j,t} = A\Big(\beta g_j(t)(V_j^e(t) - V_j^u(t))\Big)^\lambda$$

To obtain policy functions for savings, we use the method of endogenous grid points (Carroll (2006)). In period T, agents will consume their remaining assets. For each previous period, we can rearrange the Euler equations so that k_t is expressed as a function of k_{t+1} and k_{t+2}. Since we know the policy function for period $t+1$ and can replace k_{t+2} by a function of k_{t+1}, this results in an equation that just contains k_t and k_{t+1}. We use a grid of 50 points for k_{t+1} and can compute the corresponding k_t. To obtain the full policy function, we interpolate linearly between the grid points (Judd (1998)).

Given the solution to the agent problem, the update of the firm problem consists of two steps. First, we have to update the hiring probabilities $g_j(t)$ via the equation described in the model section (and, in more detail, below). Second, we need to update v using the free-entry condition. The equilibrium is computed by iterating these steps until convergence.

Computing the hiring rates. Recall the following two expressions needed for the hiring rates:

$$\begin{aligned} p(t,\phi) &= \sum_{k=1}^{J} \frac{a_k}{a} \cdot \pi_k \cdot P\Big(\Pi(\tilde{\phi},\tilde{t}) \geq \Pi(\phi,t)|k\Big) \\ g_j(t) &= \pi_j \int_\phi \exp\Big(-p(\phi,t)\cdot\mu\Big) dF_j(\phi) \end{aligned}$$

We compute these expressions as follows:

- $P(\cdot|k)$ is the probability that a random draw of type j from the pool is better than a given applicant. This is the following probability:

$$\int_{\tilde{\phi}} \Big(\sum_{\tilde{t}=1}^{T} \mathbb{1}(\Pi(\tilde{\phi}, \tilde{t}) \geq \Pi(\phi, t)) \frac{S_{j,\tilde{t}} s_{j,\tilde{t}} \alpha_j}{\sum_t S_{j,t} s_{j,t} \alpha_j} f_j(\tilde{\phi}) \Big) d\tilde{\phi}$$

  We evaluate the integral using Gauss-Legendre quadrature.

- Given these probabilities, we calculate $g_j(t)$ using Gauss-Hermite quadrature with 5 nodes.

A.2 Institutional Details

We create two samples of unemployment spells. One from 2000 until 2011 as specified in Section 1.2.1 and a sample from 1983 until 2010. The second is necessary to create a sample of unemployed individuals that receive two or more unemployment spells in their work history, because in the estimation part we use some moments from a multiple spell sample to identify the heterogeneity parameters. In the following we describe the sample creation for the 1983 sample, because the 2010 sample is just a simple subsample of the former. To account for changing rules and laws over the sample period that determine UI eligibility we use an eligibility simulator and drop all individuals that are not eligible for 12 months of UI. The simulator includes age cutoffs (older individuals receive benefits for longer), employment history regulations and drops individuals that might be subject to carry-forward rules that come into play for individuals with multiple unemployment spells. Shorter durations are applied to individuals with unstable working histories; longer durations to older workers. In order to obtain a proper sample of unemployment spells it is necessary to implement the main features of the German unemployment insurance system. To do so, we restrict ourselves to unemployment spells starting from January 1^{st}, 1983 until the end of the last day of 2011. Since our data ends in 2014 we only consider unemployment spells that we observe for at least three years. We choose 1983 as the beginning, since we need to observe the employment history of individuals four years prior to their unemployment spell in order to determine UI eligibility. In Germany, the duration of UI recipiency depends on the employment history in the last four years from January 1^{st} 1983 until June 30^{th} 1987, the last three years from July 1^{st} 1987 until January 31^{st} 2006 and the last two years from from February 1^{st} 2006 until December 31^{st} 2011. The number of years that are considered for the employment history is legally called base period (*Rahmenfristen*). In our analysis, we will only consider individuals that are eligible for 12 months of unemployment benefits when they lose their job. The general rule is determined by an abeyance ratio (*Anwartschaftsverhältnis*). The abeyance rule says that the months worked in the base period divided by 3 (from 1.1.1983 until 30.6.1987) or 2 (from 1.7.1987 until 31.12.2011) determines the maximal UI eligibility (abstracting from age cutoffs). Table A.1 summarizes the mapping from the months worked in the base period into the months of UI eligibility for the period from 1983 until 2011. (See Hunt (1995); Schmieder et al. (2010) for similar tables.) For individuals with a certain age, special rules apply that extend the potential UI duration to more than 12 months. For these individuals the base period is seven years. These individuals are not in our sample and the table does not show the potential durations for these individuals[1]. The table entries with ages in brackets show when individuals become eligible for longer durations due to their age. All individuals that are below the age cutoff receive 12 months of benefits. We drop all unemployment spells from our sample to which certain age restrictions apply.

For the estimation, we use some moments that use information from the second unemployment spell of individuals. However, for individuals that experience their second unemployment spell complex carry-forward rules apply if the second spell is not more than four years after the beginning of the first spell. To avoid modelling these rules we restrict second spells to be at least four years after the beginning of the first spell. Second, we restrict unemployment spells to individuals aged between 20 and 55. For individuals older than 55 the German social security system offers several early retirement schemes. For individuals below the age of 20, there is often the opportunity to go back to some form of school. We then drop third and fourth unemployment spells from the data, even though only a handful individuals are eligible for UI three or more times. Further, we exclude any ambigous spells from the sample. These are in particular the following cases that can arise: (a) individuals that receive UI and UA at the same time for more than 30 days and (b) individuals that are employed and receive UI at the same time for more than 14 days.[2] If we observe two consecutive unemployment spells within 14 days we pool them together and count them as one spell. With all these restrictions we arrive at a final estimation sample of $179,696$ individuals, where $18,432$ individuals experience an additional second spells. In our sample from 2000 onwards we have $59,793$ first unemployment spells.

An unemployment spell is defined as the transition from employment to UI within 30 days. Individuals that register more than 30 days after their last job has ended are dropped to avoid voluntary quitters that have a waiting period of 3 months and to avoid to wrongly measure unemployment spells due to individuals that do not take-up UI within a month. Employment consists of either socially insured employment, apprenticeships, minor employment, or other forms of registered employment. We define unemployment duration as the time between the start of UI recipiency until next employment starts (similar as in Card et al. (2007) and Schmieder et al. (2012)), though we also count moves to apprenticeship, or minor employment relationships as re-employment. We also cap unemployment durations at 36 months. This is necessary, because in the data there are many spells with long tails and some individuals

[1]I.e. the table ignores working histories of more than 48 months.
[2]It is not entirely clear where these cases come from, however there are only a few of them.

that never return to work or have an additional entry. The re-employment wage is defined as the wage the individual earns at the first employed position after unemployment.

A.3 Additional Figures & Tables

Table A.1: Potential unemplyoment benefit durations

Months worked in base period	1.1.83 - 31.12.84 (4 years)	1.1.85 - 31.12.85 (4 years)	1.1.86 - 30.6.87 (4 years)	1.7.87 - 31.3.97 (3 years)	1.4.97 - 31.12.04 (3 years)	1.1.05 - 31.1.06 (3 years)	1.2.06 - 31.7.08 (2 years)	1.8.08 - 31.12.11 (2 years)
12	4	4	4	6	6	6	6	6
16	4	4	4	8	8	8	8	8
18	6	6	6	8	8	8	8	8
20	6	6	6	10	10	10	10	10
24	8	8	8	12	12	12	12	12
28	8	8	8	14(≥42)	14(≥45)	12	12	12
30	10	10	10	14(≥42)	14(≥45)	15(≥55)	15(≥55)	15(≥50)
32	10	10	10	16(≥42)	16(≥45)	15(≥55)	15(≥55)	15(≥50)
36	12	12	12	18(≥42)	18(≥45)	18(≥55)	18(≥55)	18(≥50)
40	12	12	12	20(≥44)	20(≥47)	18(≥55)	18(≥55)	18(≥50)
42	12	14(≥49)	14(≥44)	20(≥44)	20(≥47)	18(≥55)	18(≥55)	18(≥50)
44	12	14(≥49)	14(≥44)	22(≥44)	22(≥47)	18(≥55)	18(≥55)	18(≥50)
48	12	16(≥49)	16(≥44)	24(≥49)	24(≥52)	18(≥55)	18(≥55)	24(≥50)

Notes: This table is based on Hunt (1995); Schmieder et al. (2010) and own calculations. For individuals with a certain age, special rules apply that extend the potential UI duration to more than 12 months. For these individuals the base period is seven years. These individuals are not in our sample and the table does not show the potential durations for these individuals. The table entries with ages in brackets show, if individuals become eligible for longer durations due to their age (for working histories of less than 48 months). All individuals that are below the age cutoff receive 12 months of benefits.

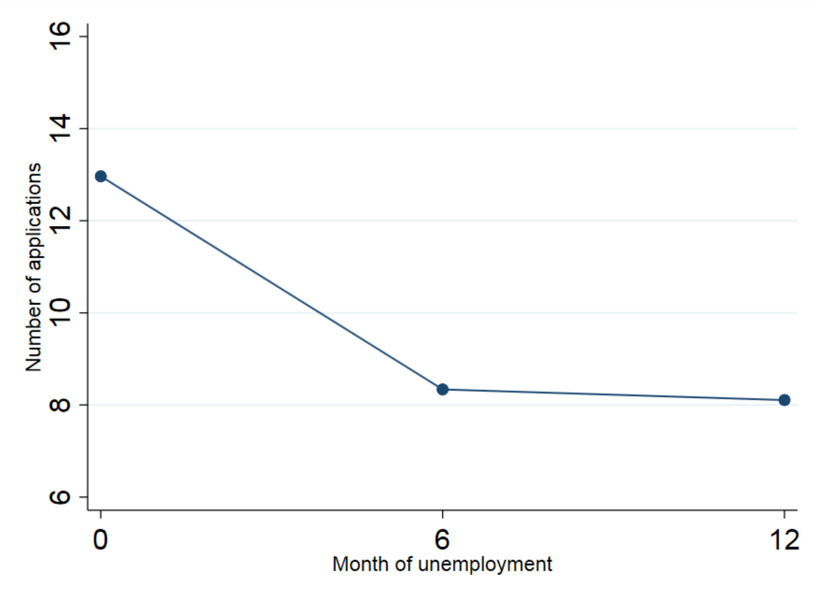


Figure A.1: Conditional Search Effort

Notes: The figure shows the average search effort conditional on staying unemployed for at least one year. Search effort is measured on the y-axis in terms of the number of applications. Source: IZA ED.

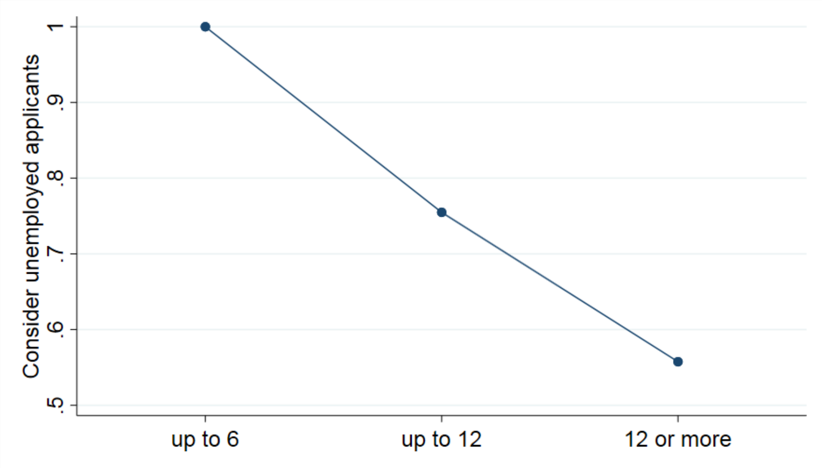


Figure A.2: Consider unemployed applicants

Notes: This graph shows the response to whether vacancies consider unemployed applicants as a function of the unemployment duration in months. The answers in the figure are conditional on reviewing unemployed applicants at all. The x-axis shows the categories in the survey question (consider applicants with up to 6 months of UI duration, up to 12 months of UI duration or longer than 12 months of UI). The y-axis plots the fraction of firms that still consider certain applicants. Source: JVS.

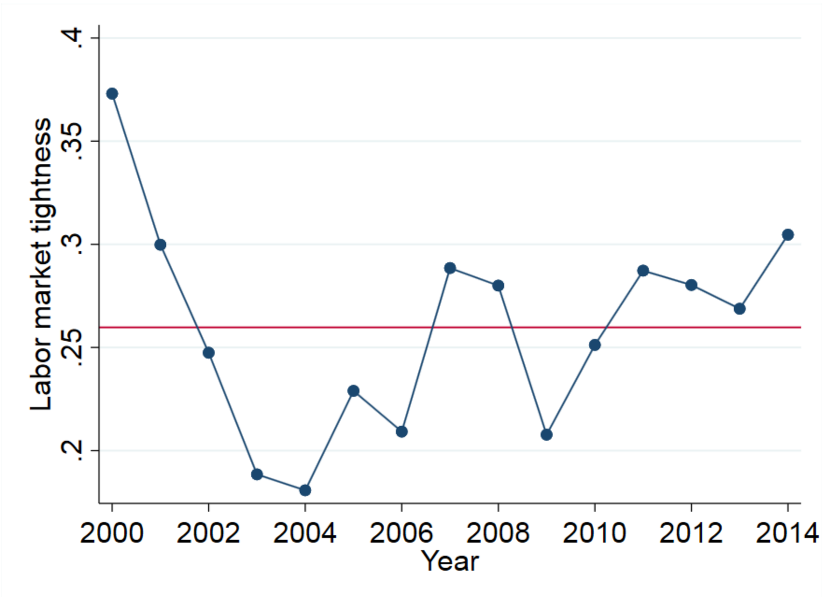


Figure A.3: Labor market tightness

Notes: This figure plots the labor market tightness for Germany from 2000 until 2014. Labor market tightness is defined as the ratio of open vacancies over the number of registered unemployed. The horizontal line denotes the average labor market tightness over the period. This figure shows that there are fewer vacancies than unemployed and that even when each vacancy is filled there remain some job seekers, which provides additional evidence that crowding-out factors and multiple applications among job seekers might be of importance. Source: Insitute for Employment Research (IAB).

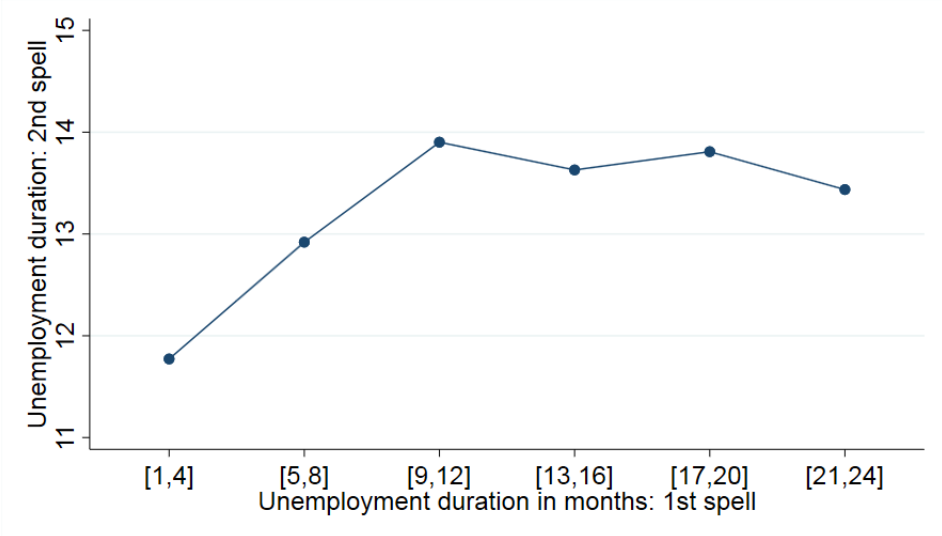


Figure A.4: Mean duration in second unemployment spell

Notes: The x-axis of this figure puts the unemployment duration of the first UI spell into 4-month bins and shows the mean duration in the second spell on the y-axis. The sample of spells is extended to the period from 1983 until 2011. Source: SIAB.

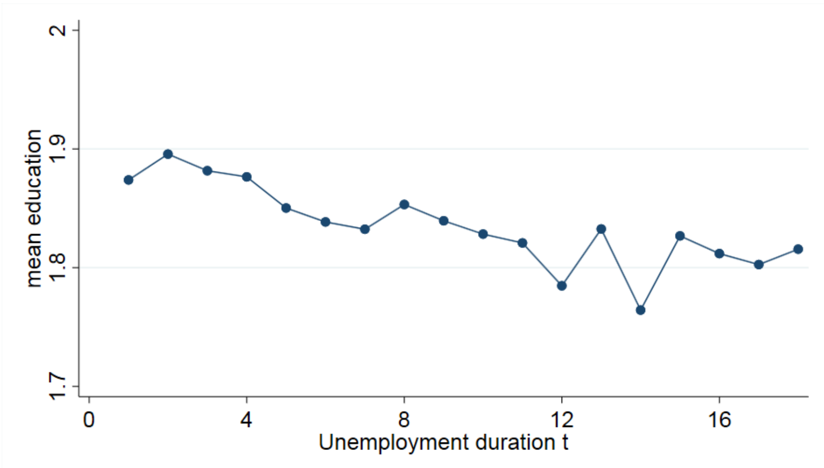


Figure A.5: Mean education over UI spell

Notes: In this graph we plot the mean education of unemployed as a function of the UI duration. The education variable is defined as follows: 0 no school degree. 1 school degree. 2 apprenticeship. 3 college. Source: SIAB.

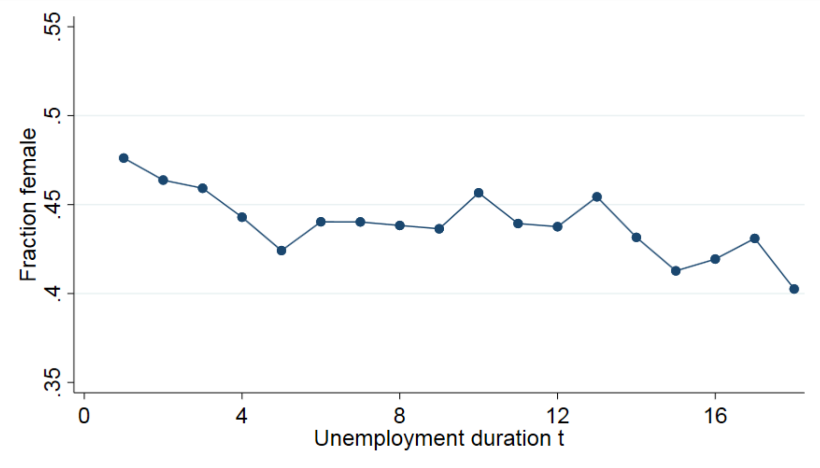
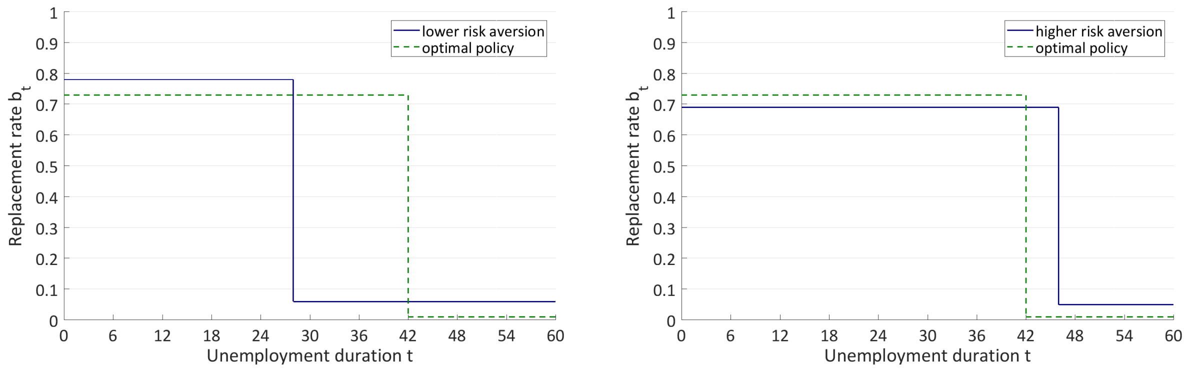


Figure A.6: Fraction female over UI spell

Notes: In this graph we plot the fraction of female unemployed as a function of the UI duration. Source: SIAB.

A.4 Alternative Parametrizations



(a) Lower risk aversion　　**(b)** Higher risk aversion

Figure A.7: Different risk aversion

Notes: This figure compares the optimal policy of our baseline model (dashed line in both panels) with a setting where agents are either less risk averse with $\gamma = 1.8$ (solid line, panel (a)) or more risk averse with $\gamma = 2.2$ (solid line, panel (b)).

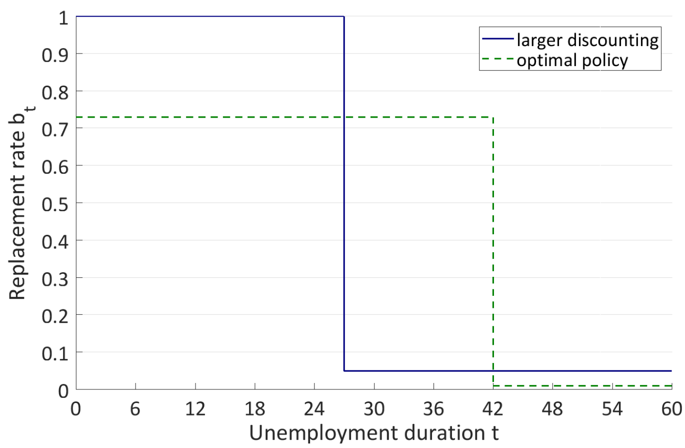


Figure A.8: Higher discounting

Notes: This figure compares the optimal policy of our baseline model (dashed line) with a setting where agents have a larger monthly discounting factor, i.e. $\beta = 0.95$. The optimal policy under this assumption is illustrated with the solid line.

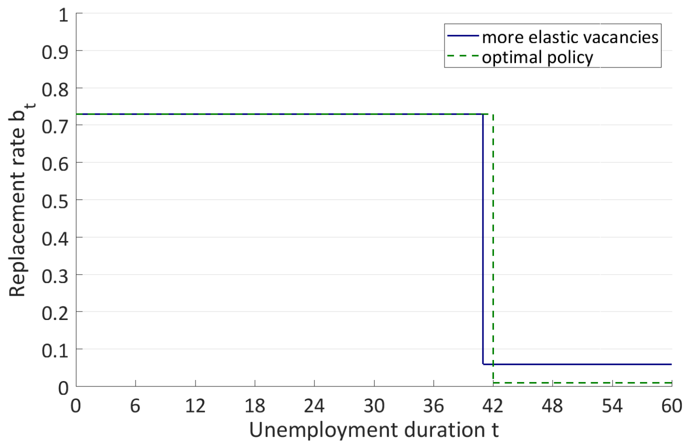


Figure A.9: Higher elasticity of vacancy creation

Notes: This figure compares the optimal policy of our baseline model (dashed line) with a setting where $\rho = 0.5$, i.e. the vacancy creation is more elastic and vacancy costs are closer to linear. The optimal policy can be seen in the solid line.

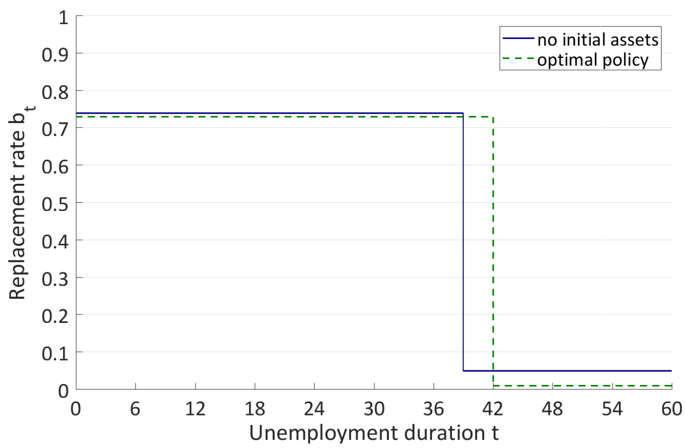


Figure A.10: No initial assets

Notes: This figure compares the optimal policy of our baseline model (dashed line) with a setting where no agent has initial assets (solid line).

Appendix B

Appendix to chapter 2

B.1 Computational Details

Solving the model requires finding a fixed point between the *expected* distributions of singles, which individuals use to compute the expected valules over the future, and the *actual* distributions of singles, which depend on the marriage and divorce choices that individuals make. The equilibrium essentially requires rational expectations, so that the beliefs about the future coincide with the actual distributions of singles. The basic algorithm is to

- start with a guess of the expected distributions of the characteristics of singles at each age $\Lambda_{t,g}(\epsilon, H, \alpha, A)$
- solve the life-cycle problem for the value functions given these expected distributions
- simulate the model, where individuals make choices about consumption and time use, marriage and divorce given the expected value functions from the previous step
- compute the implied distributions based on the simulations
- update the expected distributions and repeat until convergence

Note that little can be said about the theoretical equilibrium properties, regarding existence and uniqueness, as is standard in search models of marriage (also see Guvenen and Rendall (2015) or Greenwood et al. (2016)). In numerical checks, this does not pose an issue given the calibrated parameter values.

The life-cycle problem

Given a guess for the expected distribution of singles at each age, the life-cycle problem can be solved via backwards-induction starting in the terminal period T. The resulting value function has the interpretation as the *expected* value function given expectations about the future availability of singles. Both the assets and the Pareto weights are treated as continuous state variables. Values outside of the corresponding grids are interpolated linearly. The distributions of singles at each age are discretized on a fine grid in the asset dimension. Note that temporary wage shocks do not introduce separate state variables, which facilicates including them without greatly increasing the size of the state space. Since the temporary shocks are i.i.d. and do not influence continuation values, for example a positive wage shock is equivalent to an increase in assets (conditional on labor supply), so that the temporary shocks can be evaluated on the same grid that is also used to evaluate different asset levels.

Simulation and updating the distributions

Having computed the value function, one needs to update the guess of the distributions of singles, which is done via simulation. In the simulation, potential partners are always drawn from the *actual* distribution of singles, which can be computed from the currently available singles in the simulation (given the assumption of meetings only occuring within cohorts). The guess for the *expected* distributions of singles enters (only) through the value functions, which individuals use for example to compare the value from getting married to the value of staying single. In the end, the guess for the distributions is updated (with a weight κ on the old distributions).

Implementation

The model is implemented with Python and Numba and solved on a computing cluster. The parallelization is based on MPI. To get good performance for this type of model, it is helpful to use MPI only for inter-node communication and use shared-memory parallelism (such as Numba's *prange*) within a node. This limits communication time and allows to have a full node as a master MPI rank with sufficient memory for the simulations.

B.2 Alternative Inequality Measures

For robustness, I also conducted the decompositions for two alternative inequality measures that are commonly used. For the purpose of this chapter, the Theil index and the Mean Logarithmic Deviation (TI and MLD in the following) are most suitable, because they can be decomposed into within and between components. As a result, the same decomposition formulas that were used in the main text can be applied. The two indices are defined as:

$$\begin{aligned}\text{TI} &= \frac{1}{N}\sum_{i=1}^{N}\frac{X_i}{\bar{X}}\log\left(\frac{X_i}{\bar{X}}\right)\\ \text{MLD} &= \frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{\bar{X}}{X_i}\right)\end{aligned}$$

The tables replicates the main results for these two inequality measures. Since they require the variable to be positive, I report the results for the equivalence scale (instead of the utility measure) and private consumption. The results are in tables B.1 and B.2. The first row in each section shows the decomposition of the level of the inequality measure and the second row shows the decomposition of the change. The main conclusions from the tables are similar as when using the variance. In addition, the differences between the Theil index and the MLD are small.

Table B.1: Decomposition of Mean Logarithmic Deviation

	Total	Singles (%)	Within couples (%)	Between couples (%)	Betw. sin. mar. (%)
(a) $\mathbf{c_i}$					
MLD ($\psi_2 = 0.15$)	4.7	21.7	23.6	54.3	0.4
Change ($\psi_2 = 0.21$)	-11.2	0.15	0.23	0.62	-0
(b) Log eq. scale					
MLD ($\psi_2 = 0.15$)	9.8	13.5	6.5	69.7	10.2
Change ($\psi_2 = 0.21$)	-4.9	0.27	0.1	0.61	0.02
(c) Log eq. scale (LT)					
MLD ($\psi_2 = 0.15$)	6.9	16.9	8.5	74.5	0.1
Change ($\psi_2 = 0.21$)	-8	0.22	0.09	0.69	-0

Notes: This table replicates the decomposition for the MLD.

Table B.2: Decomposition of Theil index

	Total	Singles (%)	Within couples (%)	Between couples (%)	Betw. sin. mar. (%)
(a) $\mathbf{c_i}$					
Theil index ($\psi_2 = 0.15$)	4.7	21.1	23.3	55.1	0.4
Change ($\psi_2 = 0.21$)	-11.4	0.16	0.24	0.61	-0
(b) Log eq. scale					
Theil index ($\psi_2 = 0.15$)	9.8	9.7	7.5	73.3	9.5
Change ($\psi_2 = 0.21$)	-5.4	0.19	0.13	0.67	0.02
(c) Log eq. scale (LT)					
Theil index ($\psi_2 = 0.15$)	7.2	16.8	8.3	74.8	0.1
Change ($\psi_2 = 0.21$)	-8.4	0.23	0.1	0.68	-0

Notes: This table replicates the decomposition for the Theil index.

B.3 Details of the Decomposition Formula

The expression for the variance was given by:

$$V(X) \quad = \quad p_s V^S + p_m V^{M,W} + p_m V^{M,B} + V^{BMS}$$

The variance between couples and singles can be split into two parts:

$$\begin{aligned} V^{BMS} &= p_s V^{BMS,S} + p_m V^{BMS,M} \\ V^{BMS,S} &= (E(X|\text{ Single}) - E(X))^2 \\ V^{BMS,M} &= (E(X|\text{ Married}) - E(X))^2 \end{aligned}$$

Overall, this leads to the following decomposition of the growth rate:

$$\hat{V}(X) = \omega_0 \hat{p}_m + \omega_1 \hat{V}^S + \omega_2 \hat{V}^{M,B} + \omega_3 \hat{V}^{M,W} + \omega_4 \hat{V}^{BMS,S} + \omega_5 \hat{V}^{BMS,M} + \hat{R}$$

The weights ω_i and $\hat{R}$ and residual are the following:

$$\begin{aligned} \omega_0 &= \frac{-p_m V^s + p_m V^{M,B} + p_m V^{M,W} - p_m V^{BMS,S} + p_m V^{BMS,M}}{V} \\ \omega_1 &= \frac{p_s V^S}{V} \\ \omega_2 &= \frac{p_m V^{M,B}}{V} \\ \omega_3 &= \frac{p_m V^{M,W}}{V} \\ \omega_4 &= \frac{p_s V^{MBS,1}}{V} \\ \omega_5 &= \frac{p_m V^{MBS,2}}{V} \\ \hat{R} &= \hat{p}_m \hat{V}^S \cdot \frac{-p_m V^s}{V} + \hat{p}_m \hat{V}^{M,B} \cdot \frac{p_m V^{M,B}}{V} + \hat{p}_m \hat{V}^{M,W} \cdot \frac{p_m V^{M,W}}{V} \\ &+ \hat{p}_m \hat{V}^{BMS,S} \cdot \frac{p_m V^{BMS,S}}{V} + \hat{p}_m \hat{V}^{BMS,M} \cdot \frac{p_m V^{BMS,M}}{V} \end{aligned}$$

B.4 Calibration of the Wage Process

Since the stochastic process for wages in exogenous, it can be estimated separately from the rest of the model. I estimate the parameters of the process by matching moments of the process to data moments from the Dutch Socio-Economic Panel. Since the process is assumed to be the same for men and women, a convenient way to address potential selection problems is to estimate the process on data for men only. The parameters that need to be calibrated are the persistent and variance of the wage process (which is discretized according to the Rouwenhorst method), the variance of the temporary earnings shock, and the mean and variance of the ability distribution. The ability distribution is a discretized log-normal distribution. To separately pin down variability over time (through shocks) and persistent heterogeneity, I include moments on the variance of wage changes as well as the cross-sectional variance of wages. In addition, I include two autocorrelation moments to pin down the persistence parameter of the process and the variance of the temporary shock. The fit of the wage process and the parameters are summarized in tables B.3 and B.4.

Table B.3: Wage moments - fit

	Model	Data
$Var(\log(w_{t+1,i}) - \log(w_{t,i}))$	0.03	0.02
$Corr(\log(w_{t+1,i}), \log(w_{t,i})$	0.93	0.86
$Corr(\log(w_{t+2,i}), \log(w_{t,i})$	0.89	0.75
$E(\log(w_{t,i}))$	2.78	3.03
$Var(\log(w_{t,i}))$	0.21	0.15
$q^{10}_{\log(w_{t,i})}$	2.19	2.60
$q^{90}_{\log(w_{t,i})}$	3.43	3.49

Notes: This table compares the moments of the wage process to the data.

Table B.4: Parameters of the wage process

	Value
σ_w	0.41
ρ_w	0.94
Variance of temp. shock	0.05
Mean of persistent heterogeneity	1.10
Variance of persistent heterogeneity	0.09

Notes: This table shows the calibrated wage parameters.

The process for women is identical to the process for men except for the parameter for the gender wage gap. This parameter is set to 0.83, which reproduces the difference in mean earnings between men and women in the data. Note that this abstract from potential selection issues as well as from endogenous human capital accumulation, which could be included in a richer specification.[1]

[1]With human capital accumulation, the gender wage gap would be endogenous.

B.5 Additional Tables

Table B.5: Composition of inequality under the two policy regimes

	$\psi_2 = 0.15$	$\psi_2 = 0.21$
(b) $\log(c_i)$		
Within couples (% of total)	23.9	23.9
Between couples (% of total)	54.3	53.3
Singles (% of total)	21.4	22.2
Betw. sin. and mar (% of total)	0.4	0.6
(a) $u(c_{it}, C_{it}, l_{it}, D_{it})$		
Within couples (% of total)	6.2	6.1
Between couples (% of total)	67.9	68.6
Singles (% of total)	17.3	16.4
Betw. sin. and mar (% of total)	8.6	8.8

Table B.6: Variance decomposition - Reduction of progressivity to $\psi_2 = 0$

	$\log(\mathbf{c_i})$	$\mathbf{u}(\mathbf{c_i}, \mathbf{C_i}, \mathbf{l_i}, \mathbf{D_i})$	$\mathbf{u_{it}^{LT}}$
Change (%)	27.8	12	17.5
Decomposition of change			
Within couples (%)	25.54	8.33	9.71
Between couples (%)	61.87	53.33	70.86
Singles (%)	13.67	26.67	21.14
Betw. sin. and mar. (%)	-1.08	10.83	-0.57
Single probability (%)	0	0	-0.57

Notes: This table shows the effect of the policy change on the variance of private consumption, the per-period utility and life-time utility. The first row shows the total change. The other rows decompose this total change in inequality and add up to 100%. Compared to the results from table 2.5, the table illustrates that increasing or reducing progressivity has relatively symmetric effects.

B.6 More Detailed Description of Data

In this appendix, I discuss the data in more detail. For data on consumption and time use, the LISS panel is used. I pool the surveys of 2009, 2010 and 2012 and keep households in which there are two adult members between ages 20 and 55, which results in a sample with 2608 person-year observations. Note that the survey design has changed over time, so that the questions about consumption from later surveys are only partially comparable to the initial waves. The measure of assignable private consumption contains expenditure on:

- Food and drinks outside the house
- Cigarettes and other tobacco products
- Clothing
- Personal care products and clothing
- Health costs and medical care
- Leisure time expenditure
- (Further) schooling
- Donations and gifts
- Other personal expenditure

For the computation of the wage moments, I use data from the Dutch Socio-Economic Panel (*Sociaal-economisch panel onderzoek*). This provides a significantly larger sample size and is less subject to panel attrition than the LISS panel. Pooling the period of time between 1984 and 2002 yields 17973 person-year observations for men. Note that the time window of the Socio-Economic Panel precedes the LISS panel.[2]

[2] As an alternative data source on wages, there also is the *Arbeidsaanbodpanel*, which is available between 1985 and 2014 and also provides a larger sample size than the LISS panel. However, this dataset only provides information on *net* (rather than gross) wages, so that a tax simulator would be required to recover the underlying gross wages.

Appendix C

Appendix to chapter 3

C.1 More Details on Task Variables

In this appendix, we provide some additional information on our grouping of occupations and the computed task intensities. Figure C.1 first shows the distribution of each of the task intensity measure across the 120 occupations. The task intensities typically range between values close to 0 and 0.5, with substantial variation across occupations. Interestingly, the distribution is most skewed for analytical tasks, since a large number of occupations involves low analytical task intensities, whereas the distribution is more centered for the other task variables.

Tables C.1, C.2 and C.3 list all the occupations which we use in the analysis. The classification is based on a broader classification with 343 occupations and groups related occupations according to the digits of their code. The original classification is the KldB 1988 (*Klassizierung der Berufe*), which was developed by the German Employment Agency. To keep the exposition simple, the table lists the *first* occupation of the original classification which is in that group. For example, the group 'gardeners' includes the sub-occupations 'gardeners, garden workers', 'garden architects, garden managers', 'florists', 'forestry managers, foresters, hunters', and 'forest workers, forest cultivators'.[1]

[1]A complete list of occupations, including both German and English occupational titles, is provided by the IAB (`http://doku.iab.de/fdz/Klassifikationen_de_en.xlsx`).

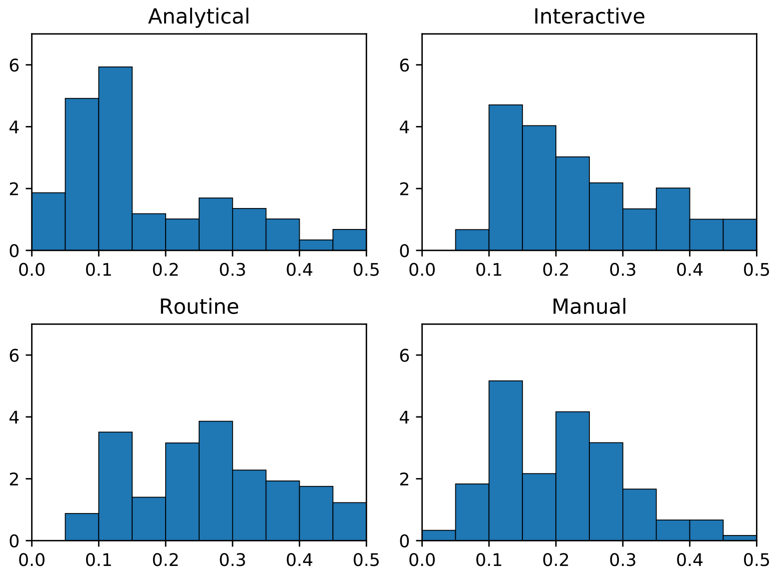


Figure C.1: Distribution of task intensities across occupations

Table C.1: Occupations and tasks

Occupation	Analytical	Interactive	Routine	Manual
Farmers	0.10	0.19	0.43	0.24
Gardeners	0.11	0.26	0.33	0.23
Miners	0.06	0.12	0.38	0.22
Ceramic workers	0.08	0.12	0.42	0.11
Chemical plant operatives	0.14	0.15	0.54	0.22
Chemical laboratory workers	0.15	0.12	0.40	0.13
Plastics processors	0.12	0.17	0.54	0.21
Paper makers	0.10	0.17	0.53	0.20
Typesetters	0.13	0.16	0.38	0.11
Special printers	0.05	0.11	0.44	0.13
Wood prepaper	0.05	0.11	0.41	0.18
Iron, metal producers, melters	0.08	0.13	0.48	0.17
Sheet metal pressers	0.04	0.10	0.41	0.11
Turners	0.08	0.13	0.49	0.15
Drillers	0.09	0.13	0.51	0.17
Metal grinders	0.07	0.12	0.50	0.20
Metal polishers	0.10	0.15	0.41	0.16
Welders, oxy-acetylene cutters	0.07	0.11	0.34	0.22
Steel smiths	0.11	0.17	0.29	0.31
Sheet metal workers	0.12	0.19	0.31	0.33
Plumbers	0.10	0.23	0.31	0.36
Locksmiths	0.05	0.09	0.25	0.25
Engine fitters	0.18	0.22	0.52	0.37
Plant fitters, maintenance fitters	0.13	0.22	0.39	0.40
Motor vehicle repairers	0.10	0.18	0.22	0.36
Agricultural machinery repairers	0.15	0.18	0.36	0.25
Other mechanics	0.11	0.17	0.30	0.27
Toolmakers	0.12	0.14	0.42	0.20
Dental technicians	0.16	0.27	0.43	0.24
Electrical fitters, mechanics	0.13	0.19	0.28	0.32
Telecommunications mechanics, craftsmen	0.11	0.18	0.23	0.27
Electrical appliance fitters	0.21	0.23	0.37	0.29
Electrical appliance, electrical parts assemblers	0.12	0.14	0.38	0.16
Other assemblers	0.09	0.14	0.37	0.17
Other metal workers	0.06	0.10	0.39	0.16
Spinners, fibre preparers	0.06	0.14	0.36	0.14
Cutters	0.04	0.11	0.30	0.10
Bakery goods makers	0.06	0.20	0.49	0.14
Butchers	0.05	0.20	0.45	0.14
Cooks	0.10	0.24	0.36	0.23
Wine coopers	0.12	0.18	0.55	0.23

Table C.2: Occupations and tasks (continued)

Occupation	Analytical	Interactive	Routine	Manual
Wine coopers	0.12	0.18	0.55	0.23
Bricklayers	0.06	0.15	0.21	0.29
Carpenters	0.09	0.18	0.27	0.31
Roofers	0.06	0.20	0.20	0.40
Paviors	0.06	0.16	0.26	0.32
Tracklayers	0.09	0.16	0.27	0.29
Building labourer, general	0.02	0.09	0.17	0.28
Stucco workers, plasterers	0.08	0.18	0.26	0.31
Tile setters	0.09	0.20	0.24	0.30
Room equippers	0.12	0.23	0.29	0.24
Furniture makers	0.12	0.20	0.42	0.28
Painters, lacquerers	0.07	0.19	0.15	0.27
Goods painters	0.10	0.19	0.34	0.27
Goods examiners, sorters	0.24	0.20	0.31	0.12
Packagers, goods receivers	0.04	0.11	0.31	0.13
Assistants (no further specification)	0.02	0.07	0.29	0.11
Generator machinists	0.07	0.14	0.45	0.27
Machine attendants	0.13	0.17	0.50	0.26
Mechanical, motor engineers	0.54	0.36	0.20	0.08
Electrical engineers	0.51	0.38	0.25	0.12
Architects	0.49	0.38	0.18	0.14
Survey engineers	0.44	0.40	0.22	0.13
Chemists	0.46	0.36	0.25	0.09
Mechanical engineering technicians	0.45	0.33	0.31	0.17
Electrical engineering technicians	0.31	0.28	0.29	0.24
Measurement technicians	0.36	0.25	0.31	0.13
Other technicians	0.29	0.29	0.25	0.16
Foremen, master mechanics	0.26	0.37	0.35	0.23
Biological specialists	0.35	0.21	0.34	0.17
Chemical laboratory assistants	0.38	0.19	0.38	0.13
Technical draughtspersons	0.36	0.14	0.09	0.03
Wholesale and retail trade buyers, buyers	0.15	0.40	0.17	0.13
Salespersons	0.07	0.34	0.20	0.15
Publishing house dealers, booksellers	0.14	0.35	0.21	0.15
Commercial agents, travellers	0.23	0.46	0.12	0.14
Bank specialists	0.21	0.35	0.12	0.07
Insurance specialists	0.23	0.39	0.09	0.08
Forwarding business dealers	0.20	0.36	0.21	0.13
Tourism specialist	0.28	0.45	0.13	0.12
Railway engine drivers	0.11	0.19	0.23	0.25
Motor vehicle drivers	0.04	0.12	0.20	0.26

Table C.3: Occupations and tasks (continued)

Occupation	Analytical	Interactive	Routine	Manual
Motor vehicle drivers	0.04	0.12	0.20	0.26
Navigating ships officers	0.16	0.25	0.28	0.24
Post masters	0.08	0.18	0.19	0.11
Warehouse managers, warehousemen	0.11	0.21	0.32	0.19
Transportation equipment drivers	0.04	0.11	0.28	0.21
Stowers, furniture packers	0.04	0.12	0.27	0.20
Entrepreneurs, managing directors	0.33	0.52	0.19	0.15
Management consultants	0.41	0.43	0.13	0.08
Members of Parliament	0.36	0.36	0.10	0.07
Cost accountants, valuers	0.17	0.22	0.10	0.04
Cashiers	0.03	0.27	0.10	0.13
Data processing specialists	0.47	0.36	0.26	0.13
Office specialists	0.17	0.25	0.13	0.07
Stenographers, shorthand-typists, typists	0.10	0.19	0.13	0.05
Office auxiliary workers	0.11	0.14	0.12	0.09
Factory guards, detectives	0.15	0.25	0.15	0.29
Doormen, caretakers	0.11	0.21	0.24	0.45
Soldiers, border guards, police officers	0.36	0.33	0.15	0.24
Journalists	0.33	0.38	0.14	0.08
Musicians	0.35	0.38	0.21	0.10
Artistic and assisting occupations	0.27	0.31	0.25	0.15
Physicians	0.34	0.42	0.24	0.24
Non-medical practitioners	0.26	0.36	0.12	0.24
Nurses, midwives	0.25	0.34	0.28	0.41
Nursing assistants	0.16	0.22	0.21	0.40
Nursing assistants	0.25	0.26	0.35	0.18
Medical receptionists	0.17	0.30	0.24	0.24
Social workers, care workers	0.30	0.42	0.15	0.33
Home wardens, social work teachers	0.31	0.50	0.13	0.32
Nursery teachers, child nurses	0.27	0.49	0.08	0.33
University teachers, lecturers	0.31	0.45	0.12	0.11
Music teachersr	0.27	0.48	0.14	0.17
Economic and social scientists, statisticians	0.37	0.40	0.11	0.08
Hairdressers	0.08	0.29	0.09	0.21
Restaurant, inn, bar keepers	0.09	0.31	0.16	0.30
Others guest attendants	0.06	0.26	0.17	0.27
Housekeeping managers	0.10	0.23	0.17	0.31
Laundry workers, pressers	0.03	0.12	0.26	0.20
Household cleaners	0.02	0.09	0.05	0.23
Street cleaners, refuse disposers	0.10	0.16	0.24	0.36

Table C.4: Task changes for switches between selected occupation groups

From	To	Analytical	Interactive	Routine	Manual
MPT	Sal./Off.	-0.11	0.06	0.05	0.01
POC	Sal./Off.	0.13	0.23	-0.20	-0.15
POC	Serv.	0.02	0.12	-0.21	0.08
Serv.	Sal./Off.	0.11	0.11	0.01	-0.23
Serv.	POC	-0.02	-0.12	0.21	-0.08
Care	Sal./Off.	0.03	0.11	0.04	-0.17

C.2 Additional Material on Cohort Profiles

Table C.5: Number of observations by cohort

	Men	Women
Cohort		
1940-44	41247	63717
1945-49	46087	56718
1950-54	55063	56209
1955-59	72838	66582
1960-64	108488	88542
1965-69	121447	101265
1970-74	122364	121732
1975-79	100515	103458

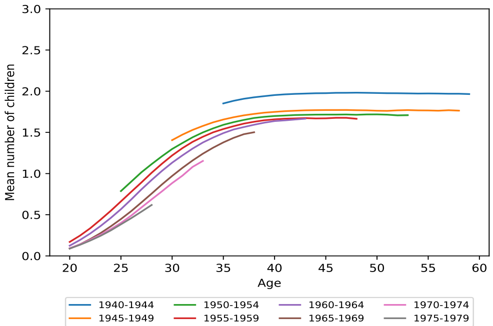
Notes: The table shows the number of men and women, which are used for constructing the cohort profiles. Since the panel data contains gaps, the number of observations fluctuates with age and we report the number of observations at the age at which the cohort enters the sample.

Table C.6: Number of observations by cohort and completed fertility

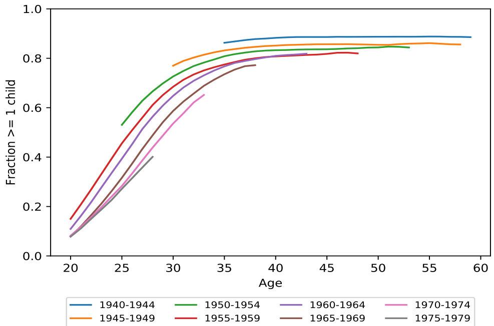
	No kids	1 kid	2 kids	3 or more kids
1940-44	7459	13100	22382	11042
1945-49	7077	11774	18614	7706
1950-54	7218	9804	15405	6357
1955-59	8400	8550	14740	5720
1960-64	9306	9094	15487	5706
1965-69	5178	4513	7305	2321

Notes: The table shows the number of women separately by completed fertility, which are used for constructing the cohort profiles.

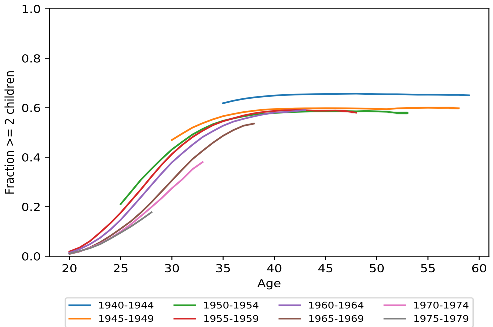
Figure C.2: Fertility by cohort



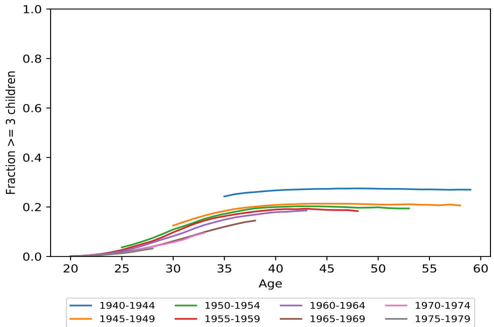
(a) Number of children



(b) Fraction with at least 1 child

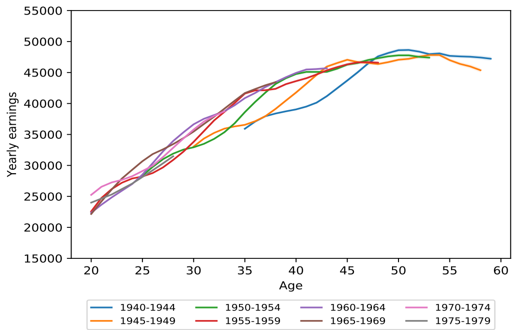


(c) Fraction with at least 2 children

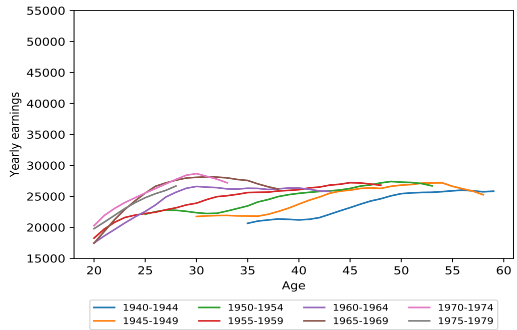


(d) Fraction with at least 3 children

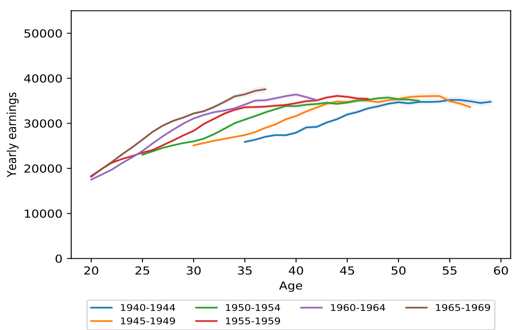
Figure C.3: Annual earnings by cohort



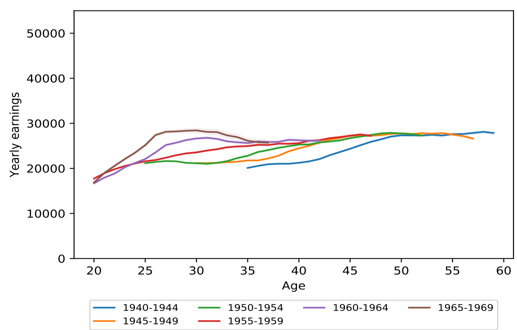
(a) Men



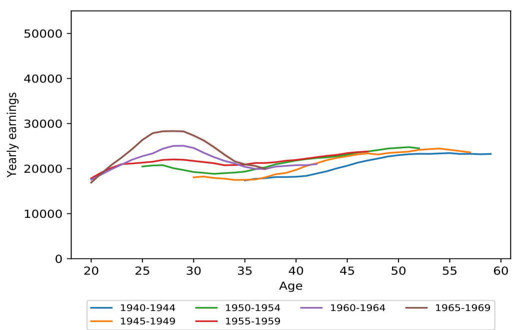
(b) Women



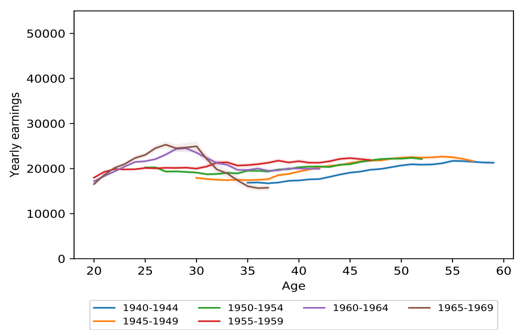
(c) Women (no children)



(d) Women (one child)

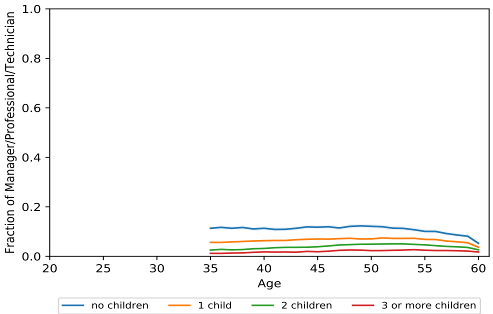


(e) Women (two children)

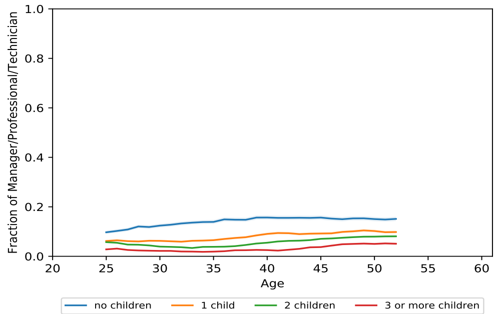


(f) Women (three or more children)

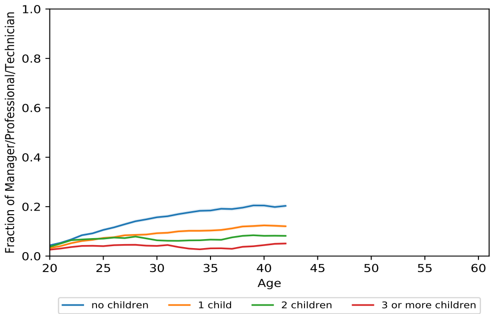
Figure C.4: Fraction of Managers/Professionals/Technicians by cohort and completed fertility



(a) Women (cohort 1940 - 1944)

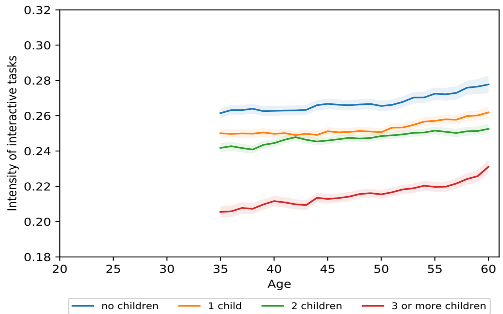


(b) Women (cohort 1950 - 1954)

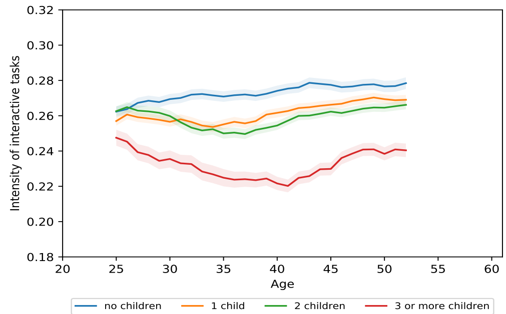


(c) Women (cohort 1960 - 1964)

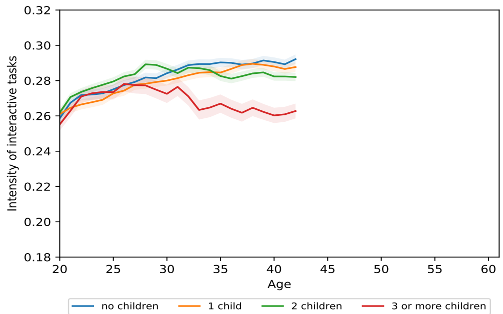
Figure C.5: Interactive task content by cohort and completed fertility



(a) Women (cohort 1940 - 1944)



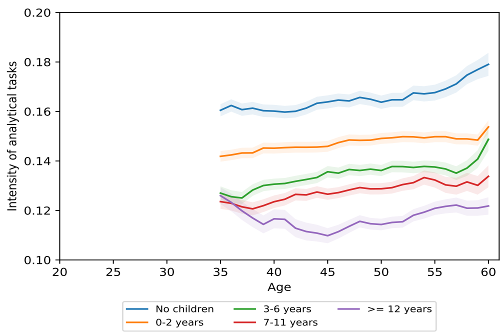
(b) Women (cohort 1950 - 1954)



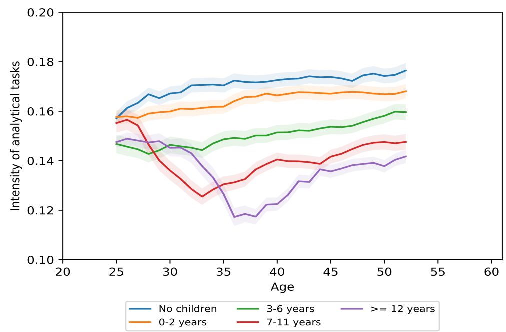
(c) Women (cohort 1960 - 1964)

Figure C.6: Analytical task content by cohort and total length of breaks

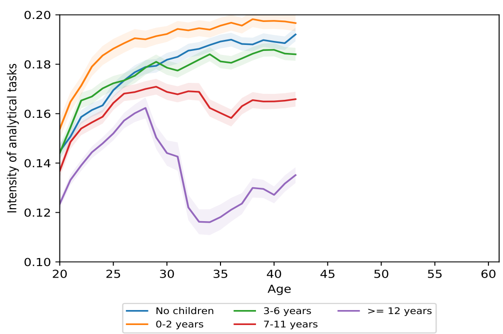
(a) Cohort 1940-1944



(b) Cohort 1950-1954



(c) Cohort 1960-1964



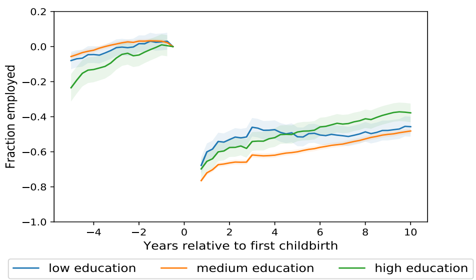
C.3 Additional Material on Event Studies

Table C.7: Summary statistics - Event study sample

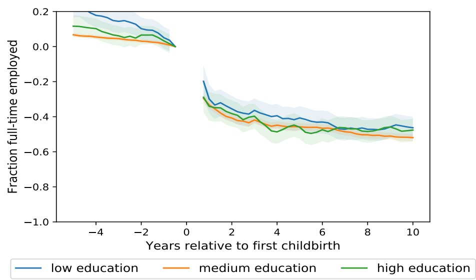
Number of observations	22935
Fraction with low education	0.15
Fraction with medium education	0.76
Fraction with high education	0.09
Fraction with 1 child	0.39
Fraction with 2 children	0.45
Fraction with 3 or more children	0.16

Figure C.7: Event studies by education

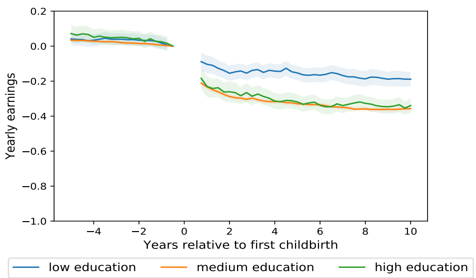
(a) Employment



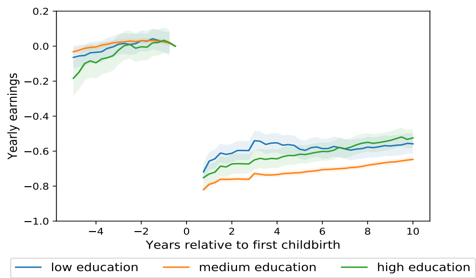
(b) Full-time employment (if employed)



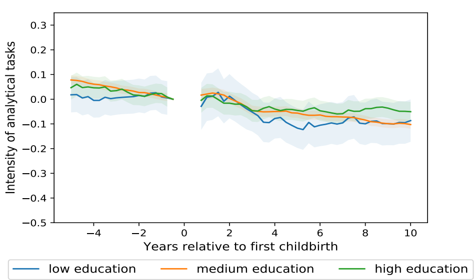
(c) Yearly earnings



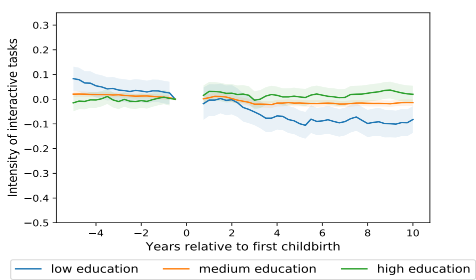
(d) Yearly earnings (missing = 0)



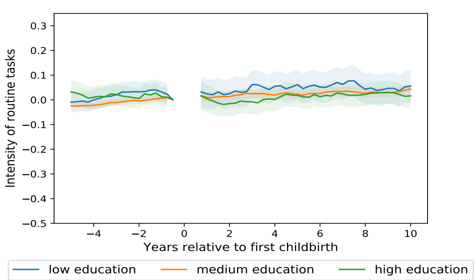
(e) Analytical tasks



(f) Interactive tasks



(g) Routine tasks



(h) Manual tasks

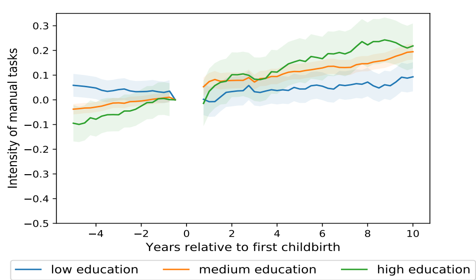
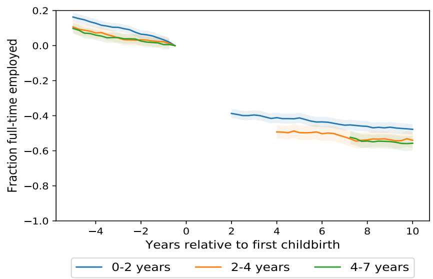
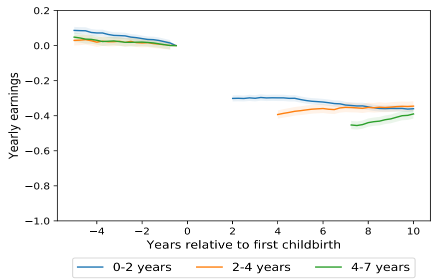


Figure C.8: Event studies by length of career break

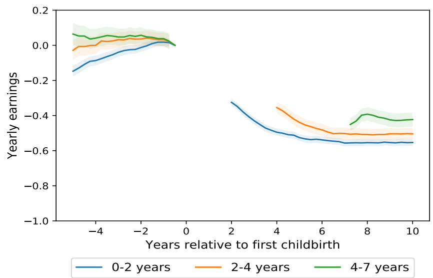
(a) Full-time employment (if employed)



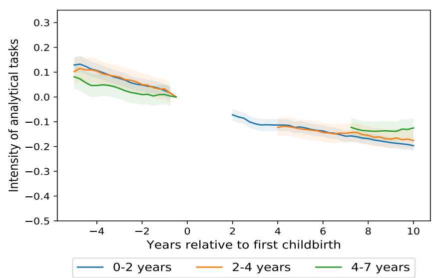
(b) Yearly earnings



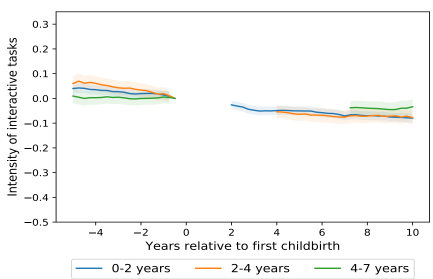
(c) Yearly earnings (missing = 0)



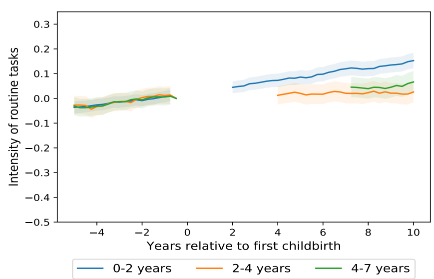
(d) Analytical tasks



(e) Interactive tasks



(f) Routine tasks



(g) Manual tasks

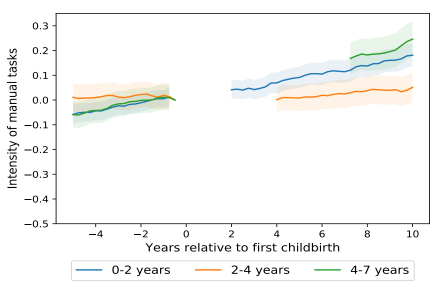
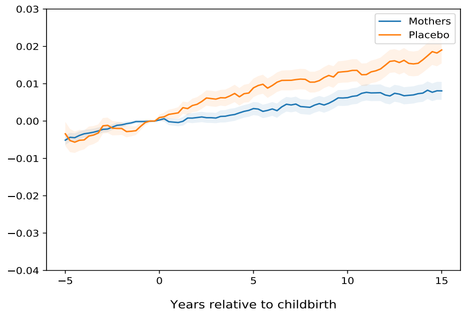
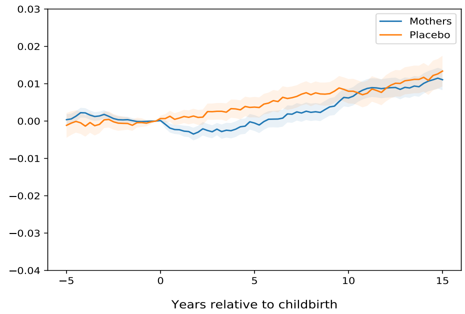


Figure C.9: Robustness: Placebo event studies for rescaled task intensities

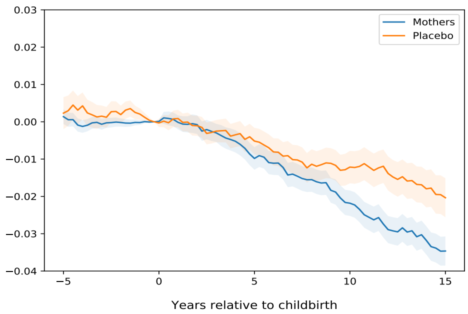
(a) Analytical



(b) Interactive



(c) Routine



(d) Manual

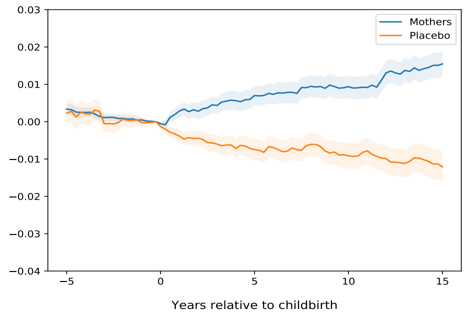
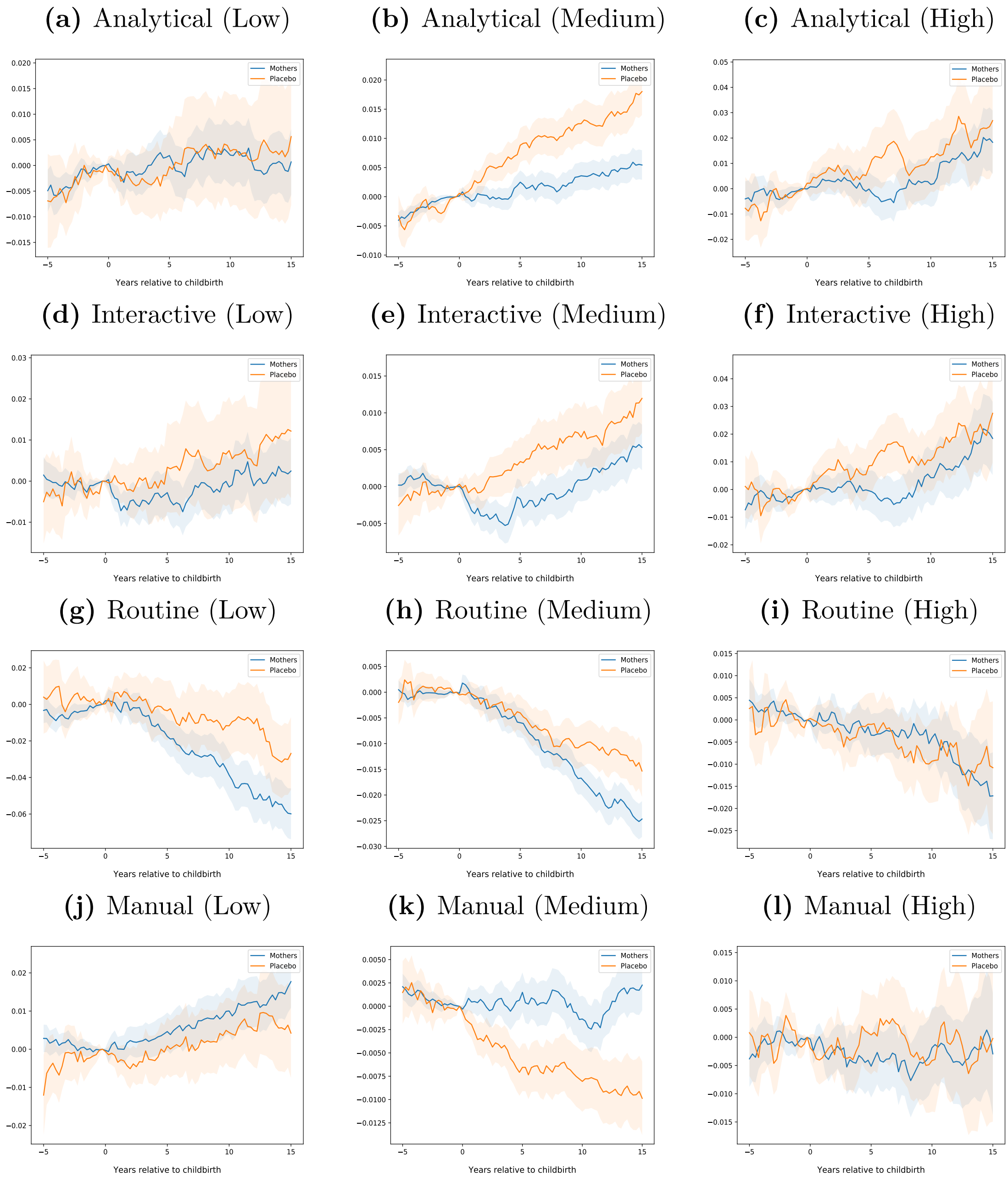


Figure C.10: Occupation switches - by education



(a) Analytical (Low) **(b)** Analytical (Medium) **(c)** Analytical (High)

(d) Interactive (Low) **(e)** Interactive (Medium) **(f)** Interactive (High)

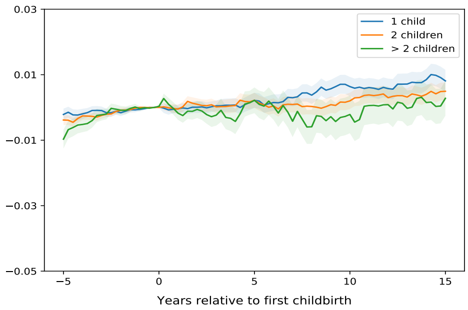
(g) Routine (Low) **(h)** Routine (Medium) **(i)** Routine (High)

(j) Manual (Low) **(k)** Manual (Medium) **(l)** Manual (High)

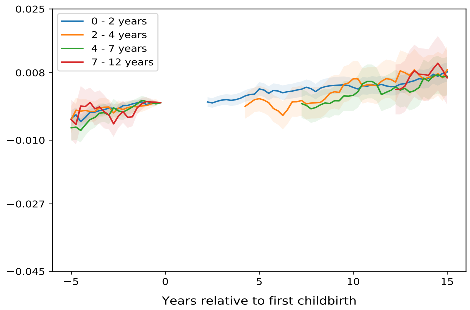
Notes: The figure conducts the Placebo event study analysis separately by education group. ”Low” corresponds to women with no degree, ”Medium” to women with a high school or vocational degree, and ”High” to those with a university degree.

Figure C.11: Occupation switches - by children (CH) and length of break (B)

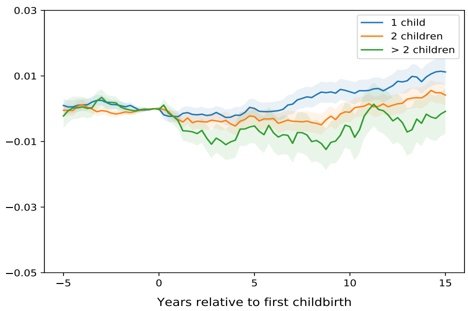
(a) Analytical (CH)



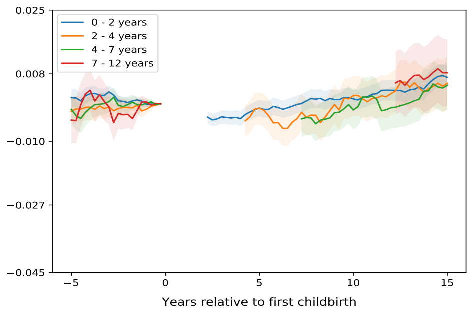
(b) Analytical (B)



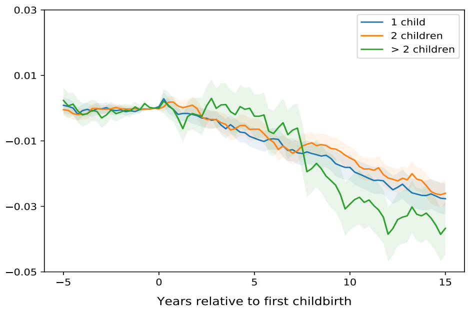
(c) Interactive (CH)



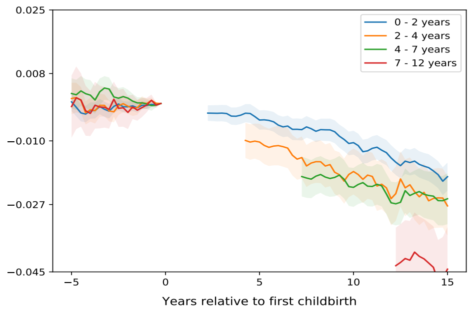
(d) Interactive (B)



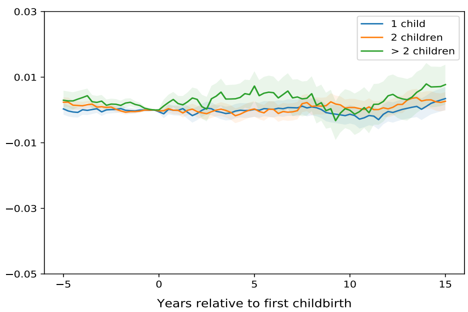
(e) Routine tasks (CH)



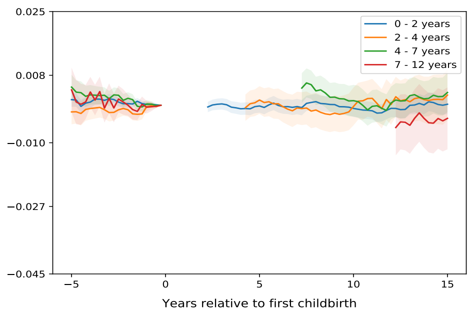
(f) Routine tasks (B)



(g) Manual tasks (CH)



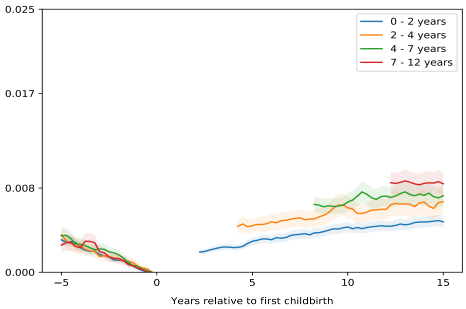
(h) Manual tasks (B)



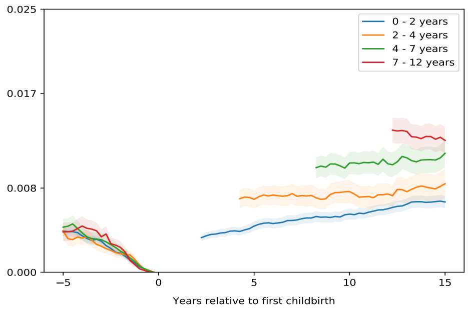
Notes: The confidence bands refer to the 10% level.

Figure C.12: Task distance for each task intensity, by length of career break

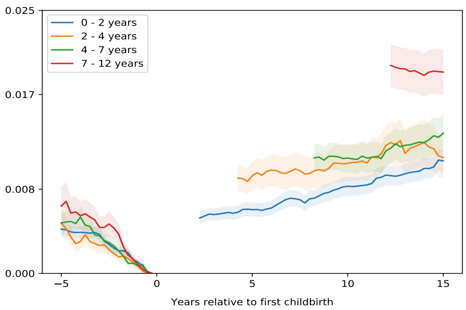
(a) Analytical



(b) Interactive



(c) Routine



(d) Manual

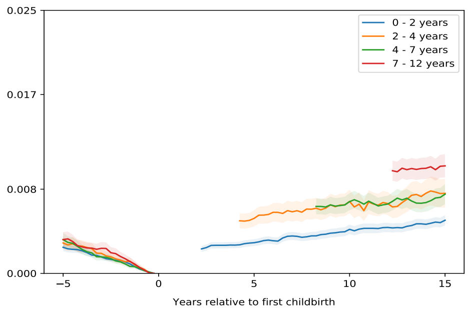
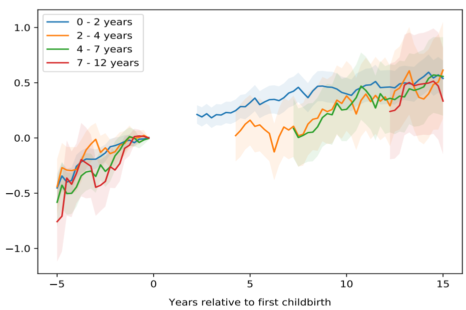
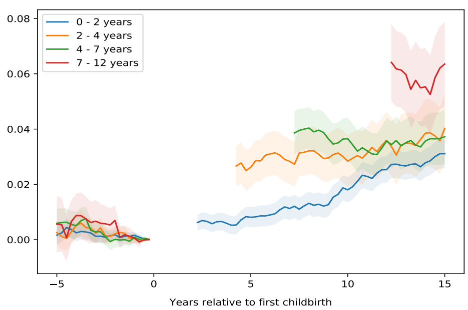


Figure C.13: Other occupation characteristics, by length of career break

(a) Change in occupational earnings



(b) Change in part-time in occupation



Bibliography

Adda, J., C. Dustmann, and K. Stevens (2017): "The career costs of children," *Journal of Political Economy*, 125, 293–337.

Albrecht, J., M. A. Bronson, P. S. Thoursie, and S. Vroman (2018): "The career dynamics of high-skilled women and men: Evidence from Sweden," *European Economic Review*, 105, 83–102.

Albrecht, J., P. A. Gautier, and S. Vroman (2006): "Equilibrium directed search with multiple applications," *The Review of Economic Studies*, 73, 869–891.

Alesina, A., A. Ichino, and L. Karabarbounis (2011): "Gender-based taxation and the division of family chores," *American Economic Journal: Economic Policy*, 3, 1–40.

Alvarez, F., K. Borovicková, and R. Shimer (2016): "Decomposing Duration Dependence in a Stopping Time Model," *Unpublished manuscript*.

Angelov, N., P. Johansson, and E. Lindahl (2016): "Parenthood and the gender gap in pay," *Journal of Labor Economics*, 34, 545–579.

Attanasio, O., H. Low, and V. Sánchez-Marcos (2008): "Explaining changes in female labor supply in a life-cycle model," *American Economic Review*, 98, 1517–52.

Attanasio, O. P. and V. Lechene (2014): "Efficient responses to targeted cash transfers," *Journal of Political Economy*, 122, 178–222.

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The skill content of recent technological change: An empirical exploration," *The Quarterly Journal of Economics*, 118, 1279–1333.

Baily, M. N. (1978): "Some aspects of optimal unemployment insurance," *Journal of Public Economics*, 10, 379–402.

Bastani, S. (2013): "Gender-based and couple-based taxation," *International Tax and Public Finance*, 20, 653–686.

Bick, A. and N. Fuchs-Schündeln (2017): "Taxation and Labour Supply of Married Couples across Countries: A Macroeconomic Analysis," *The Review of Economic Studies*, 85, 1543–1576.

Black, S. E. and A. Spitz-Oener (2010): "Explaining women's success: technological change and the skill content of women's work," *The Review of Economics and Statistics*, 92, 187–194.

Blanchard, O. J. and P. Diamond (1994): "Ranking, unemployment duration, and wages," *The Review of Economic Studies*, 61, 417–434.

Blau, F. D. and L. M. Kahn (2017): "The gender wage gap: Extent, trends, and explanations," *Journal of Economic Literature*, 55, 789–865.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption inequality and partial insurance," *American Economic Review*, 98, 1887–1921.

Blundell, R. and I. Preston (1998): "Consumption inequality and income uncertainty," *The Quarterly Journal of Economics*, 113, 603–640.

Bronson, M. A. and M. Mazzocco (2019): "Taxation and Household Decisions: an Intertemporal Analysis," *Unpublished manuscript*.

Card, D., R. Chetty, and A. Weber (2007): "Cash-on-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market," *The Quarterly Journal of Economics*, 1511–1560.

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 128, 967–1015.

Carroll, C. D. (2006): "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics Letters*, 91, 312–320.

Cerina, F., A. Moro, and M. P. Rendall (2018): "The Role of Gender in Employment Polarization," *Unpublished Manuscript*.

Chade, H. and G. Ventura (2002): "Taxes and Marriage: A Two-Sided Search Analysis," *International Economic Review*, 43, 955–985.

——— (2005): "Income taxation and marital decisions," *Review of Economic Dynamics*, 8, 565–599.

Cherchye, L., S. Cosaert, B. De Rock, P. J. Kerstens, and F. Vermeulen (2018): "Individual welfare analysis for collective households," *Journal of Public Economics*, 166, 98–114.

Cherchye, L., B. De Rock, A. Lewbel, and F. Vermeulen (2015): "Sharing rule identification for general collective consumption models," *Econometrica*, 83, 2001–2041.

Cherchye, L., B. De Rock, and F. Vermeulen (2012): "Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information," *American Economic Review*, 102, 3377–3405.

Cherchye, L., T. Demuynck, B. De Rock, and F. Vermeulen (2017): "Household consumption when the marriage is stable," *American Economic Review*, 107, 1507–34.

Chetty, R. (2006): "A general formula for the optimal level of social insurance," *Journal of Public Economics*, 90, 1879–1901.

——— (2008): "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, 116, 173–234.

Chiappori, P.-A., M. C. Dias, and C. Meghir (2018): "The Marriage Market, Labor Supply, and Education Choice," *Journal of Political Economy*, 126, S26–S72.

Chiappori, P.-A. and M. Mazzocco (2017): "Static and intertemporal household decisions," *Journal of Economic Literature*, 55, 985–1045.

Chiappori, P.-A. and C. Meghir (2015): "Intrahousehold inequality," in *Handbook of Income Distribution*, Elsevier, vol. 2, 1369–1418.

Ciscato, E. (2019): "The Changing Wage Distribution and the Decline of Marriage," *Unpublished manuscript*.

Conesa, J. C. and D. Krueger (2006): "On the optimal progressivity of the income tax code," *Journal of Monetary Economics*, 53, 1425–1450.

de Boer, H.-W., E. Jongen, P. Koot, et al. (2018): "Optimal Taxation of Secondary Earners in the Netherlands: Has Equity Lost Ground?" Tech. rep., CPB Netherlands Bureau for Economic Policy Analysis.

DellaVigna, S., A. Lindner, B. Reizer, and J. F. Schmieder (2017): "Reference-dependent job search: Evidence from Hungary," *The Quarterly Journal of Economics*, 132, 1969–2018.

Doppelt, R. (2016): "The Hazards of Unemployment: A Macroeconomic Model of Job Search and Résumé Dynamics," *Unpublished manuscript*.

Dustmann, C., J. Ludsteck, and U. Schönberg (2009): "Revisiting the German wage structure," *Quarterly Journal of Economics*, 124, 843–881.

Felfe, C. (2012): "The motherhood wage gap: What about job amenities?" *Labour Economics*, 19, 59–67.

Fernandez, R., N. Guner, and J. Knowles (2005): "Love and money: A theoretical and empirical analysis of household sorting and inequality," *The Quarterly Journal of Economics*, 120, 273–344.

Fernández, R. and R. Rogerson (2001): "Sorting and long-run inequality," *The Quarterly Journal of Economics*, 116, 1305–1341.

Fernandez, R. and J. C. Wong (2017): "Free to leave? A welfare analysis of divorce regimes," *American Economic Journal: Macroeconomics*, 9, 72–115.

Fernandez-Blanco, J. and E. Preugschat (2018): "On the effects of ranking by unemployment duration," *European Economic Review*, 104, 92–110.

Fernández-Kranz, D., A. Lacuesta, and N. Rodríguez-Planas (2013): "The Motherhood Earnings Dip: Evidence from Administrative Records," *Journal of Human Resources*, 48, 169–197.

Fitzenberger, B. (2012): "Wages and employment across skill groups: an analysis for West Germany," .

Fitzenberger, B., R. Schnabel, and G. Wunderlich (2004): "The gender gap in labor market participation and employment: A cohort analysis for West Germany," *Journal of Population Economics*, 17, 83–116.

Galichon, A., S. D. Kominers, and S. Weber (2018): "Costly concessions: An empirical framework for matching with imperfectly transferable utility," *Journal of Political Economy (Forthcoming).*

Gathmann, C. and U. Schönberg (2010): "How General is Human Capital? A Task-based Approach," *Journal of Labor Economics*, 28, 1–49.

Gautier, P. A., J. L. Moraga-González, and R. P. Wolthoff (2016): "Search costs and efficiency: Do unemployed workers search enough?" *European Economic Review*, 84, 123–139.

Gayle, G.-L. and A. Shephard (2019): "Optimal taxation, marriage, home production, and family labor supply," *Econometrica*, 87, 291–326.

Goldin, C. (2014): "A grand gender convergence: Its last chapter," *American Economic Review*, 104, 1091–1119.

Goldin, C. and J. Mitchell (2017): "The new life cycle of women's employment: Disappearing humps, sagging middles, expanding tops," *Journal of Economic Perspectives*, 31, 161–82.

Goussé, M., N. Jacquemet, and J.-M. Robin (2017): "Marriage, labor supply, and home production," *Econometrica*, 85, 1873–1919.

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2016): "Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation," *American Economic Journal: Macroeconomics*, 8, 1–41.

Groes, F., P. Kircher, and I. Manovskii (2014): "The U-shapes of occupational mobility," *The Review of Economic Studies*, 82, 659–692.

Gruber, J. (1997): "The Consumption Smoothing Benefits of Unemployment Insurance," *The American Economic Review*, 87, 192–205.

Guner, N., R. Kaygusuz, and G. Ventura (2011): "Taxation and household labour supply," *The Review of Economic Studies*, 79, 1113–1149.

——— (2012): "Taxing women: A macroeconomic analysis," *Journal of Monetary Economics*, 59, 111–128.

Guvenen, F. and M. Rendall (2015): "Women's emancipation through education: A macroeconomic analysis," *Review of Economic Dynamics*, 18, 931–956.

Hagedorn, M., F. Karahan, I. Manovskii, and K. Mitman (2019): "Unemployment Benefits and Enemployment in the Great Recession: the Role of Equilibrium Effects," *Unpublished manuscript.*

Hall, R. E. and A. I. Mueller (2015): "Wage dispersion and search behavior," *Unpublished manuscript.*

Hardoy, I., P. Schøne, and K. M. Østbakken (2017): "Children and the gender gap in management," *Labour Economics*, 47, 124–137.

Heathcote, J., K. Storesletten, and G. L. Violante (2017): "Optimal tax progressivity: An analytical framework," *The Quarterly Journal of Economics*, 132, 1693–1754.

Holter, H. A., D. Krueger, and S. Stepanchuk (2019): "How does tax progressivity and household heterogeneity affect Laffer curves?" *Quantitative Economics (Forthcoming).*

Hopenhayn, H. A. and J. P. Nicolini (1997): "Optimal Unemployment Insurance," *Journal of Political Economy*, 105, 412–438.

Hunt, J. (1995): "The effect of unemployment compensation on unemployment duration in Germany," *Journal of Labor Economics*, 88–120.

Jarosch, G. and L. Pilossoph (2018): "Statistical Discrimination and Duration Dependence in the Job Finding Rate," *Review of Economic Studies (Forthcoming).*

Judd, K. L. (1998): *Numerical Methods in Economics*, MIT press.

Kaas, L. (2010): "Variable Search Intensity with Coordination Unemployment," *The BE Journal of Macroeconomics*, 10.

Kalai, E. (1977): "Proportional solutions to bargaining situations: interpersonal utility comparisons," *Econometrica: Journal of the Econometric Society*, 1623–1630.

Kircher, P. (2009): "Efficiency of simultaneous search," *Journal of Political Economy*, 117, 861–913.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2019): "Child Penalties Across Countries: Evidence and Explanations," *American Economic Review (Papers and Proceedings, forthcoming).*

Kleven, H., C. Landais, and J. E. Søgaard (2018): "Children and Gender Inequality: Evidence from Denmark," *American Economic Journal: Applied Economics.*

Knowles, J. A. (2012): "Why are married men working so much? An aggregate analysis of intra-household bargaining and labour supply," *Review of Economic Studies*, 80, 1055–1085.

Kolsrud, J., C. Landais, P. Nilsson, and J. Spinnewijn (2018): "The optimal timing of unemployment benefits: Theory and evidence from sweden," *American Economic Review*, 108, 985–1033.

Kreyenfeld, M. and T. Mika (2008): "Erwerbstätigkeit und Fertilität: Analysen mit der Versicherungskontenstichprobe der deutschen Rentenversicherung," *DRV-Schriften Band 79.*

Kroft, K., F. Lange, and M. J. Notowidigdo (2013): "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment," *The Quarterly Journal of Economics*, 128, 1123–1167.

Krueger, A. B. and A. Mueller (2011): "Job search, emotional well-being, and job finding in a period of mass unemployment: evidence from high-frequency longitudinal data," *Brookings Papers on Economic Activity*, 2011, 1–57.

Krueger, A. B. and A. I. Mueller (2016): "A contribution to the empirics of reservation wages," *American Economic Journal: Economic Policy.*

Krueger, D. and F. Perri (2006): "Does income inequality lead to consumption inequality? Evidence and theory," *The Review of Economic Studies*, 73, 163–193.

Kuziemko, I., J. Pan, J. Shen, and E. Washington (2018): "The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?" NBER Working Paper 24740, National Bureau of Economic Research, Cambridge, MA.

Lalive, R., C. Landais, and J. Zweimüller (2015): "Market externalities of large unemployment insurance extension programs," *The American Economic Review*, 105, 3564–3596.

Landais, C., P. Michaillat, and E. Saez (2016): "A Macroeconomic Approach to Optimal Unemployment Insurance: Theory," *American Economic Journal: Economic Policy, forthcoming.*

——— (2018): "A macroeconomic approach to optimal unemployment insurance: Applications," *American Economic Journal: Economic Policy*, 10, 182–216.

Lehr, B. (2017): "Optimal Unemployment Insurance with Endogenous Negative Duration Dependence," *Public Finance Review*, 45, 395–422.

Lentz, R. (2009): "Optimal unemployment insurance in an estimated job search model with savings," *Review of Economic Dynamics*, 12, 37–57.

Lichter, A. (2016): "Benefit duration and job search effort: Evidence from a natural experiment," *Unpublished manuscript.*

Lise, J. and J.-M. Robin (2017): "The macrodynamics of sorting between workers and firms," *American Economic Review*, 107, 1104–35.

Lise, J. and S. Seitz (2011): "Consumption inequality and intra-household allocations," *The Review of Economic Studies*, 78, 328–355.

Lise, J. and K. Yamada (2018): "Household Sharing and Commitment: Evidence from Panel Data on Individual Expenditures and Time Use," *The Review of Economic Studies (Forthcoming).*

Lockwood, B. (1991): "Information externalities in the labour market and the duration of unemployment," *The Review of Economic Studies*, 58, 733–753.

Low, H., C. Meghir, L. Pistaferri, and A. Voena (2018): "Marriage, Labor Supply and the Social Safety Net," *Unpublished manuscript.*

Lundberg, S. J., R. A. Pollak, and T. J. Wales (1997): "Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit," *Journal of Human Resources*, 463–480.

Marinescu, I. (2017): "The general equilibrium impacts of unemployment insurance: Evidence from a large online job board," *Journal of Public Economics*, 150, 14–29.

Mazzocco, M., C. Ruiz, and S. Yamaguchi (2013): "Labor supply, wealth dynamics, and marriage decisions," *Unpublished manuscript.*

McGrattan, E. R., R. Rogerson, et al. (2004): "Changes in hours worked, 1950-2000," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28, 14–33.

Michaillat, P. (2012): "Do matching frictions explain unemployment? Not in bad times," *The American Economic Review*, 102, 1721–1750.

Ngai, L. R. and B. Petrongolo (2017): "Gender gaps and the rise of the service economy," *American Economic Journal: Macroeconomics*, 9, 1–44.

Pavoni, N. (2009): "Optimal Unemployment Insurance With Human Capital Depreciation And Duration Dependence," *International Economic Review*, 50, 323–362.

Pendakur, K. (2018): "Welfare analysis when people are different," *Canadian Journal of Economics/Revue canadienne d'économique*, 51, 321–360.

Pertold-Gebicka, B., F. Pertold, and N. Datta Gupta (2016): "Employment adjustments around childbirth," IZA Discussion Paper 9685, Institute of Labor Economics, Bonn.

Piketty, T. and E. Saez (2013): "Optimal labor income taxation," in *Handbook of Public Economics*, Elsevier, vol. 5, 391–474.

Rendall, M. (2018): "Female market work, tax regimes, and the rise of the service sector," *Review of Economic Dynamics*, 28, 269–289.

Reynoso, A. (2019): "The impact of divorce laws on the equilibrium in the marriage market," *Unpublished manuscript.*

Santaeulàlia-Llopis, R. and Y. Zheng (2017): "Why Is Food Consumption Inequality Underestimated? A Story of Vices and Children," *Unpublished manuscript.*

Schmieder, J. F. and T. Von Wachter (2016): "The effects of unemployment insurance benefits: New evidence and interpretation," *Annual Review of Economics*, 8, 547–581.

Schmieder, J. F., T. von Wachter, and S. Bender (2010): "The effects of unemployment insurance on labor supply and search outcomes: Regression discontinuity estimates from Germany," *Unpublished manuscript.*

——— (2012): "The Effects of Extended Unemployment Insurance Over the Business Cycle: Evidence from Regression Discontinuity Estimates Over 20 Years," *The Quarterly Journal of Economics*, 127, 701–752.

——— (2016): "The effect of unemployment benefits and nonemployment durations on wages," *The American Economic Review*, 106, 739–777.

Shephard, A. (2019): "Marriage market dynamics, gender, and the age gap," *Unpublished manuscript.*

Shimer, R. (2004): "Search intensity," *Unpublished manuscript.*

Shimer, R. and I. Werning (2006): "On the Optimal Timing of Benefits with Heterogeneous Workers and Human Capital Depreciation," *Unpublished manuscript.*

——— (2008): "Liquidity and Insurance for the Unemployed," *The American Economic Review*, 98, 1922–1942.

Spitz-Oener (2006): “Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure,” *Journal of Labor Economics*, 24, 235–270.

Villena-Roldan, B. (2012): “Aggregate implications of employer search and recruiting selection,” *Unpublished manuscript.*

Wolthoff, R. (2017): “Applications and Interviews: Firms’ Recruiting Decisions in a Frictional Labour Market,” *The Review of Economic Studies*, 85, 1314–1351.

Yamaguchi, S. (2012): “Tasks and heterogeneous human capital,” *Journal of Labor Economics*, 30, 1–53.

——— (2018): “Changes in returns to task-specific skills and gender wage gap,” *Journal of Human Resources*, 53, 32–70.

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation selbstständig angefertigt und mich keiner anderen als der in ihr angegebenen Hilfsmittel bedient zu haben. Insbesondere sind sämtliche Zitate aus anderen Quellen als solche gekennzeichnet und mit Quellenangaben versehen.

Mannheim, 28.6.2019

Curriculum Vitae, Tim Obermeier

2015 – 2019	PhD Studies in Economics, Center for Doctoral Studies in Economics, University of Mannheim. Employee at the Chair for Public Finance and Economic Policy, Prof. Eckhard Janeba, University of Mannheim.
2013 – 2015	M.Sc. in Economic Research, University of Mannheim.
2010 – 2013	B.Sc. in Economics, University of Mannheim.
2010	Abitur, Immanuel-Kant-Gymnasium, Bad Oeynhausen.