Empirical Analysis of Social Insurance, Work Incentives and Employment Outcomes

Empirical Analysis of Social Insurance, Work Incentives and Employment Outcomes

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 24 januari 2024
klokke 13.45 uur

door

Heike Tammo Vethaak

geboren te Purmerend

in 1996

Promotores: prof.dr. C.L.J. Caminada
prof.dr. P.W.C. Koning (Vrije Universiteit Amsterdam)

Promotiecommissie: prof.dr. A.C. Gielen (Erasmus Universiteit Rotterdam)
prof.dr. E.L.W. Jongen
prof.dr. B. van der Klaauw (Vrije Universiteit Amsterdam)
prof.dr. O. van Vliet
dr. H.T. Wermink

The research in this book is sponsored by Instituut Gak.

Lay-out: AlphaZet prepress, Bodegraven
Printwerk: Ipskamp Printing

Preface

The years in which I wrote my dissertation were a fantastic experience for which I am very grateful. Therefore, I would like to thank those who provided me with this opportunity and supported me in the process. These people not only made the past years fantastic, but also make that I am confident that the coming years will be at least as good.

I am grateful for the support of my supervisors Koen Caminada and Pierre Koning. Koen, your trust helped me to stay calm and focused. Pierre, you not only contributed to the research we worked on together, but also taught me a lot during these years which will help in the rest of my career.

I would like to thank Anne Gielen, Egbert Jongen, Bas van der Klaauw, Olaf van Vliet and Hilde Wermink for taking place in my Ph.D. committee and providing me with useful comments to this dissertation.

Special thanks to my co-authors Bas, Ernst-Jan, Jim, Marike and Pierre. You have proven that writing a dissertation does not necessarily have to be a solitary process. I enjoyed all the in-depth discussions, both on research as on a broad range of other topics. Additionally, I have learned a great deal from you, which will help me in the rest of my career. For all of this, I am grateful to have worked with you and I look forward to working with you in the future.

Many thanks go to all who provided and helped me with the data, including CBS, Gemeente Rotterdam and UWV. Foremost, I would like to thank Peter Berkhout and Hans Terpstra for their truly indispensable assistance and comments. A special thanks to Patrick Hullegie and Lieke Kools for providing me with useful STATA code.

Further, I would like to thank the many participants at the conferences I presented, especially Patrick Arni, Chris Muris, Daniel van Vuuren and the discussants who spent their valuable time to read the papers thoroughly.

To all my colleagues of the Department of Economics of Leiden University, thank you for these years in which I felt at home as one of you.

Jim and Marco, my *paranimphs*, thank you for all the support and joy, both inside as outside of work. The memories we made and meaningful conversations we had are important to me.

Finally, I would like to thank my family and friends. My parents, Ada and Wouter, who have loved and supported me indefinitely. Mom, you are an example for me on how to be strong and caring at the same time. Dad, you showed me what is important in life. I still miss you. My sister and brother, Paulien who watches over me and Tom who keeps challenging me. Also thanks to the Zwier family. I thank my friends for all the necessary distractions. Above all Michael, who has always been a constant factor on whom I can fall back on and talk to. Yael, of so many things I thought it was impossible, but you make it possible.

Contents

Preface **v**

1 Introduction **1**

2 A panel data sample selection model to estimate life-cycle earning profiles: How important is selection into full-time and part-time employment? **11**

2.1 Introduction 12
2.2 Institutional background: Part-time employment in the Netherlands 15
2.3 Data 18
2.3.1 Data selection and variable definitions 18
2.3.2 Descriptive statistics 19
2.4 Model 27
2.4.1 Panel data sample selection model 28
2.4.2 Estimation 33
2.5 Estimation results 35
2.5.1 First stage: labor force participation 35
2.5.2 Second stage: wages 36
2.6 Binary versus ordered selection 41
2.7 Conclusion 42
2.A Wage descriptives over time 45
2.B Derivation of correction terms 46
2.C Estimation of the selection equation 49
2.C.1 Labor supply categories 49
2.C.2 Transitions in labor supply categories 50
2.C.3 First-stage regression results 52
2.D Robustness of the wage equation 54

3 Decomposing employment trends of disabled workers **57**

3.1 Introduction 57

3.2 Institutional background 61
3.2.1 DI in the Netherlands 61
3.2.2 Stricter screening: the GKP reform (2003) 107
3.2.3 The new disability law: WIA (2006) 64
3.3 Data 66
3.3.1 Data sources 66
3.3.2 Descriptive statistics 67
3.3.3 Comparing rejected and awarded applicants 71
3.4 Empirical strategy 73
3.4.1 Specification 73
3.4.2 Identification 76
3.5 Estimation Results 78
3.5.1 The Age-Period-Cohort model 78
3.5.2 Robustness 81
3.5.3 Decomposing incentive effects 83
3.5.4 Application cohort effects in more detail 85
3.6 Conclusions 89
3.A Additional tables and figures 92

4 Empirical evaluation of broader job search requirements for unemployed workers 99
4.1 Introduction 100
4.2 Background of the experiment 103
4.2.1 The Dutch UI system 103
4.2.2 The treatment 104
4.2.3 The experiment 105
4.3 Data and experimental evaluation 107
4.3.1 Data description 107
4.3.2 Evaluating the broader search program 110
4.4 Broader search task 115
4.4.1 Empirical approach and data 115
4.4.2 Justification of the IV assumptions 118
4.5 Effects of the broader job search task 124
4.5.1 Theoretical predictions 124
4.5.2 Estimated effects 127
4.5.3 Marginal treatment effects 134
4.5.4 Decomposing the effects of caseworker meetings . . 138
4.6 Conclusion 140
4.A Back-of-the-envelope costs-benefits analysis of the broader search program 143
4.B Robustness of the effects of the broader search task 145
4.C Heterogeneous effects of the broader search task 151
4.D Decomposition of the effects for the broader search program 156

5 The effects of application processing times and prepayments on welfare receipt and employment **157**
5.1 Introduction 157
5.2 Institutional background 162
5.2.1 Welfare benefits in the Netherlands 163
5.2.2 The application procedure 164
5.2.3 Benefit prepayments 165
5.2.4 Theoretical predictions 165
5.3 Data 167
5.3.1 Data sources 167
5.3.2 Descriptive statistics 168
5.4 Methodology 174
5.4.1 Empirical approach 174
5.4.2 Extended IV model 177
5.4.3 Justification of the IV assumptions 179
5.5 Results 187
5.5.1 Effects of fast application processing times 187
5.5.2 Results of the extended model 189
5.6 Conclusion 195
5.A Justification of the assumptions: Additional tables and figures 197
5.B Robustness of the results 202
5.C Results: additional tables 204

6 General discussion **207**
6.1 Aims 207
6.2 Theory on social insurance and employment 208
6.2.1 Theoretical objectives of social insurance 208
6.2.2 Employment effects of social insurance 210
6.3 The importance of selection in (part-time) work 211
6.4 Selection, targeting and welfare effects of social insurance . 212
6.5 Policy implications and future research 215

Bibliography **219**

Nederlandse samenvatting **235**

Author contributions **245**

Curriculum Vitae **247**

List of Tables

2.1 Selection and number of labor supply categories 41
2.2 Trends in participation and wages 46
2.3 Distribution of men and women over the 5 labor supply categories over time 50
2.4 Year-to-year transitions (fractions) in labor supply categories of men and women 51
2.5 Estimation results selection equation for men and women . 52

3.1 Employment, earnings and demographic characteristics for rejected, and partially and fully awarded DI applicant cohorts 69
3.2 DiD incentive effects of the Gatekeeper Protocol (GKP) and short-term and long-term incentive effect of the WIA reform 83
3.3 Estimated cohort differentials of rejected vs. awarded DI applicants 94

4.1 Caseworker services received by the treatment and control group 106
4.2 Descriptive statistics, balancing and compliance to the experiment 108
4.3 Effects of participating in the broader search program on cumulative outcomes - instrumental variable estimates . . . 114
4.4 Descriptive statistics, assignment of caseworker stringency and the broader search task 117
4.5 Use of other policy tools related to caseworker stringency and the broader search task 120
4.6 First-stage estimates by demographics 122
4.7 Effects of imposing the broader search task on cumulative outcomes - instrumental variable estimates 131
4.8 Effect of broader search task on job characteristics - instrumental variable estimates 132

4.9 Effects of the program, the broader search task and the caseworker meeting . . . 140
4.10 First-stage estimates using different sample selections on caseworkers and different controls . . . 145
4.11 First-stage estimates using different sample selections on caseworkers and different controls - split-sample approach 146
4.12 First-stage estimates for different groups of benefits recipients - reverse-sample approach . . . 147
4.13 Effects of imposing the broader search task on cumulative outcomes - split-sample approach . . . 148
4.14 Effects of imposing the broader search task on cumulative outcomes – instrumental variable estimates with quarter fixed effects . . . 149
4.15 Effects of imposing the broader search task on cumulative outcomes – instrumental variable estimates with additional controls for caseworker policy choices . . . 150
4.16 Effects of imposing the broader search task on cumulative outcomes after one year for different demographic groups - instrumental variable estimates . . . 151
4.17 Effects of imposing the broader search task on cumulative outcomes after one year for different demographic groups - reverse-sample instrumental variable estimates . . . 152
4.18 Marginal treatment effects coefficients of imposing the broader search task on cumulative outcomes - instrumental variable estimates . . . 155
4.19 Effects of imposing the broader search task on cumulative outcomes for compliers and always takers - instrumental variable estimates . . . 156

5.1 Descriptive statistics of applicants by application outcome . 170
5.2 Descriptive statistics, assignment of caseworker speed and the observed application processing time . . . 180
5.3 Effect of caseworker processing speed on award rates . . . 182
5.4 Effect of caseworker processing speed on the monthly level of welfare benefit payments . . . 183
5.5 First-stage estimates of caseworker speed on fast inflow by subgroups . . . 186
5.6 Effects of fast application processing time on cumulative outcomes - instrumental variable estimates . . . 189
5.7 Estimation results of the extended model on cumulative outcomes two years after application – instrumental variable estimates . . . 190
5.8 Sample selections and descriptives of the samples . . . 198

5.9 Testing for random assignment of caseworker instruments used in the extended model 200
5.10 First-stage estimates using different sample selections on caseworkers and different controls 201
5.11 Effects of fast application processing time on cumulative outcomes – instrumental variable estimates with additional caseworker stringency controls 202
5.12 Effects of fast application processing time on cumulative outcomes – instrumental variable estimates with quarter fixed effects 203
5.13 Estimation results of the extended model on cumulative outcomes one year after application – instrumental variable estimates 203
5.14 Effects of prepayments on cumulative outcomes – instrumental variable estimates on the subsample with processing times longer than 8 weeks 204
5.15 Effects of fast application processing time on cumulative outcomes after one year for different demographic groups – instrumental variable estimates 205

List of Figures

2.1 Incidence of part-time employment among (a) men and (b) women in OECD countries. 17
2.2 Life-cycle earnings of men (a) and women (b) 21
2.3 Percentage of men and women in full-time and part-time employment 23
2.4 Wage of men and women in full-time and part-time employment 25
2.5 Part-time (PT) and full-time (FT) regressions for men (a) and women (b) using first-differences (FD) and our model (BKV) 37
2.6 Part-time (PT) and full-time (FT) regressions using first-differences (FD) and our model (BKV) controlling for linear period effects and unemployment rates 54
2.7 Part-time (PT) and full-time (FT) regressions for men (a) and women (b) using levels and Dustmann and Schmidt (2000) approach 56

3.1 Annual DI application rate, inflow rate and claim denial rate of total insured working population, 1999-2013 62
3.2 GKP conditions in the sickness waiting period 64
3.3 Annual fraction employed DI applicants before and after the award decision, stratified by application year (1999-2013) 70
3.4 Annual average employment rates of rejected, partially and fully awarded DI applicant cohorts for three time periods, before and after the award decision. 72
3.5 Annual employment rates and Bound estimates for different application cohort samples between 1999 and 2013, measured three years after the DI decision 74
3.6 Elapsed time (‘age’), period and application cohort effects on the employment of DI applicants 79

3.7 Comparing implied absolute declines in cohort effects of three models, measured for 1999-2002, 2003-2005, and 2006-2013 . . . 82
3.8 Annual Bound-estimates for the unrestricted APC-DP models 85
3.9 Deaton-Paxson estimation results of elapsed time ('age') and cohort effects with step-wise inclusion of sets of control variables . . . 86
3.10 Fractions of awarded and rejected DI applicants by application cohort . . . 92
3.11 Cumulative distribution of the most important impairment groups of all applications for disability insurance between 1999-2013 by application cohort . . . 92
3.12 Annual average earnings of rejected, and partially and fully awarded applicant cohorts for three time regimes, before and after application for DI benefits . . . 93
3.13 Heterogeneous Deaton-Paxson estimates for age and cohort effects for employment . . . 95
3.14 Deaton-Paxson estimation results of age and cohort effects for earnings, probability of a permanent contract, UI benefit receipt, social assistance benefit receipt and mortality . . . 97

4.1 Effects of participating in the broader search program - instrumental variable estimates . . . 112
4.2 Distribution of caseworker stringency and demeaned by local UI office and month . . . 123
4.3 Effects of imposing the broader search task - instrumental variable estimates . . . 129
4.4 Marginal treatment effects of imposing the broader search task on cumulative outcomes after one year . . . 136
4.5 Intention to treat effects of the broader search program on cumulative UI benefits - OLS estimates . . . 144
4.6 Marginal treatment effects of imposing the broader search task on cumulative outcomes after 2 years . . . 153

5.1 Application processing times and received prepayments . . 169
5.2 Welfare receipt, welfare benefits, employment and earnings before and after month of application by application outcome . . . 172
5.3 Distribution of caseworker speed (a) and conditional on UI exhaustion and team and year fixed effects (b) . . . 185
5.4 Effects of fast application processing time – instrumental variable estimates . . . 193
5.5 Distribution of mean age of the applicants by caseworker . 197

5.6 Distribution of mean of the applicants who applied after exhaustion of UI benefits by caseworker 199
5.7 Caseworker speed in period t and $t-1$ 199

1 Introduction

In the design of social insurance programs, policy makers face the trade-off between providing insurance and incentives. This trade-off refers to the provision of income security in case of unemployment, sickness and disability, while maintaining incentives that reduce the risk of unemployment. The primary goals of social insurance are to ensure a minimum standard of living and to enable intertemporal consumption smoothing by protecting workers from temporary or permanent shocks in their income. At the same time, social insurance benefits should be accessible to those in need, while those who are able to work should be reasonably encouraged to do so.

The optimal design and targeting of social insurance ultimately depends on individual and societal preferences, political choices, and social norms. Still, economists can provide support to policy makers with empirical evidence on the effects of existing regulation in, and of changes to existing regulation in social insurance programs on various outcomes. These outcomes include labor market participation, earnings, benefit take-up, leisure time, and substitution between social insurance programs. This evidence can then assist policy makers in making evidence-based decisions that are more informed and, hopefully, welfare-improving.

The challenges faced by policy makers in designing socially optimal social insurance programs are complex, due to the multidimensionality of such programs. The level of insurance or generosity of the program depends on various policy conditions, including the replacement rate, the length of eligibility, and the eligibility conditions. The strength of the work incentives depends on the same conditions, which can be adjusted to balance the insurance and work incentives of the program with its goals to increase social welfare. For example, disability insurance programs typically have high thresholds for enrollment, including long waiting periods, so as to reduce the risk of moral hazard. However, upon entering the program, recipients are entitled to relatively generous benefits and are eligible for extensive entitlement periods, with few job search requirements. This matches the contributory nature of the program, the expected re-employability, and the source of unemployment. In comparison, welfare benefits are typically lower, have shorter waiting periods and more stringent job search requirements, as they function as a safety net.

In order to understand the effects of social insurance, it is crucial to analyze the incentives that drive these effects. These incentives can have heterogeneous effects, for instance, between poorer and richer workers. Similarly, the effectiveness of a particular treatment may vary, for example as workers may respond differently to mandatory or voluntarily program components. The incentives also vary on contextual factors such as economic conditions or employment protection legislation. As a result, new empirical findings can both support and challenge renowned theoretical models or previous empirical studies. Hence, it is crucial to continuously investigate the incentives in the labor market and in social insurance programs through empirics and increase the overall knowledge on these topics.

Empirical economists provide evidence of the policy-relevant relationships by employing various methods on real-world data. Many of these methods rely on exogenous variation of the independent variable to find a causal relation between the independent and dependent variable. The ideal method to obtain causal evidence is that of using a randomized controlled trial, as it allows a direct comparison between the treatment and control group without the risk of potential interfering factors (Chapter 4 of this thesis is partly based on such experimental methods). However, in many cases, it may not be feasible or ethical to exclude individuals from the treatment of interest. As a result, empirical economists often turn to quasi-experimental designs, exploiting (instantaneous) discontinuities over time – i.e. reforms – or in exogenous differences in policy parameters (as in Chapter 3). Another potential source of exogenous variation may follow from the random assignment of workers to caseworkers with different treatment intensities (as in Chapters 4 and 5). A strong advantage of these methods is that they often provide the policy-relevant outcomes, as they are based on the population that is responsive to the treatment, also known as "compliers". Even when non-random samples are compared, interesting estimates can still be obtained through correction for the selection using sample selection models (as in Chapter 2). While various research methods are employed in the chapters of this thesis, they all exploit large-scale administrative datasets. These datasets are advantageous of their representativeness, their large sample size, their accuracy, their panel structure, and their ability to be linked with data from other administrative sources.

This thesis contains four chapters. The first chapter investigates the labor supply decisions of men and women in the Netherlands. The chapter shows the importance of accounting for selection in the intensive margin of labor supply, i.e. the number of hours that are worked. This intensive margin decision is particularly important in the Netherlands, where there are substantial disparities between the labor supply of men and women and among different birth cohorts. The other chapters aim to gain a deeper understanding of the impact of social insurance and labor market programs on employment rates and other labor market outcomes in the Netherlands. The second chapter provides insight in the determinants of employment trends among disabled workers in the Netherlands, which have been largely impacted by reforms in the disability insurance program. The third chapter exploits a large-scale field experiment to investigate the effects of caseworker meetings and mandatory broader job search requirements on unemployment insurance recipients' outcomes. The fourth and final chapter investigates the effects of welfare application processing times and benefit prepayments on labor market outcomes of welfare applicants. The chapters can be read independently and contain an extensive introduction. The remainder of the introductory chapter provides a summary of the motivations, research questions, outcomes, and contributions to the literature of the chapters.

The importance of selection for labor supply decisions

A comprehensive understanding of employment decisions and the use of social insurance largely depends on the accurate estimation of earnings equations. Estimating earnings equations is however challenging, since earnings are only observed among individuals who work. Estimating earnings models without accounting for the non-random selection into work leads to serious inconsistent estimates of earnings (Heckman 1979), even in the case of panel data (Solon 1988). Several alternative methods have therefore been proposed to address the issue of selection bias, which would otherwise lead to an overestimation of the earnings of women and part-time working men. One frequently adopted alternative is to estimate earnings solely on prime-aged men, as they are most likely to work full-time and are less likely to self-select into work. However, the conclusions drawn from such estimates may not be generalizable to women, older men, and men for whom working full-time is less common. Therefore, Chapter 2 of this thesis aims to answer the question: *"How important is selection into full-/part-time employment?"*

To answer this question, Chapter 2 proposes a novel panel data sample selection model that combines two strands of literature. The first strand addresses the methodological challenge of dealing with the selection issue, but assumes that selection into earnings follows from selection at the extensive margin of labor supply. The second strand incorporates non-binary choices in the employment decision – allowing for differences differences between part-time and full-time employment – but these models are limited to cross-sectional data. The newly proposed model in Chapter 2 addresses this gap by integrating an ordered selection rule into the selection equation, without relying on parametric assumptions about the unobserved individual-specific heterogeneity in the equation of interest. The performance of this new model, as compared to a binary estimator, is evaluated using administrative data that are representative for the Netherlands. The Netherlands provides an ideal setting for this assessment, as the prevalence of part-time work is internationally high among both men and women.

The empirical application of the model reveals the significance of accounting for selection at the intensive margin of labor supply, as the outcomes differ substantially from those obtained with a binary selection rule. When correcting for the selection at the intensive margin, we find positive selection into part-time work for both genders. This means that individuals with higher productivity select into part-time employment, and failing to correct for this selection could result in an overestimation of part-time earnings. For full-time work, we find positive selection for women only. As a result, the generally assumed absence of selection into work among men in the literature is only true when considering full-time work. Furthermore, results based on prime-aged men cannot be directly translated to women, older men, and men who might work part-time. In conclusion, these findings confirm the need to take the selection at the intensive margin into account.

Effects of disability insurance reforms on labor supply decisions

Over the last decades, many OECD countries have experienced a decline in employment rates among disabled individuals (OECD, 2010). This trend has been attributed to two primary factors (Autor and Duggan 2003, Bound et al. 2003, 2014, Maestas 2019, Von Wachter et al. 2011). Firstly, the selection of workers that entered disability insurance (DI) programs has shifted towards individuals with more vulnerable positions in the labor market. Secondly, after entering the program, DI recipients often experience a reduction in work incentives. The Netherlands was no exception to this trend, as it has also witnessed a strong reduction in the labor force participation among DI recipients. Two major reforms were implemented in the Netherlands during the period under investigation in Chapter 3, namely the Gatekeeper Protocol (Wet verbetering Poortwachter) in 2003 and the introduction of the new disability law (Wet Werk en Inkomen naar Arbeidsvermogen) in 2006. The reforms affected the population of disabled workers in two distinct ways. On the one hand, increases in screening stringency and eligibility thresholds led to increased targeting of the DI program, changing the selection of applicants towards workers with more severe disabilities and worse labor market characteristics (De Jong et al. 2011, Godard et al. 2022). On the other hand, the new disability law increased work incentives for new cohorts of DI recipients with residual earnings capacities (Koning and van Sonsbeek 2017). Therefore, Chapter 3 of this thesis aims to answer the question: *"To what extent is the employment of disabled workers in the Netherlands affected by targeting effects and incentive effects?"*

To answer this question, Chapter 3 assesses the employment trends of disabled workers between 1999 and 2013 in the Netherlands, while incorporating both selection and work incentive effects. For this, the chapter employs Age-Period-Cohort (APC) models, which allow to disentangle the application cohort specific effects from the calendar year ('period') effects and the elapsed time since application ('age') effects. Both selection and the work incentive effects are embodied in application cohort effects, since the new program rules exclusively apply to new applicants. Next, we further decompose the application cohort effects into selection effects and work incentive effects by combining APC models with difference-in-differences (DiD) models. The DiD models compare cohorts of rejected and awarded applicants before and after the reforms, assuming that the reforms changed the selection of applicants in the two groups equally. The resulting DiD estimates represent the work incentive effects of the reforms on the awarded applicants.

The results show that the declining employment rates of DI applicants in the Netherlands can be largely explained by application-cohort effects. The most recent cohort of DI applicants has a 30 percentage points lower employment rate compared to the first cohort in the data. In contrast, calendar year effects are negligible, suggesting that business cycle effects or secular time trends that affected all application cohorts equally were not that important. The changes in cohort effects are largely in tandem with the two reforms, although a gradual decline in cohort effects is observed following the second reform. The DiD analysis shows that the work incentive effects of the reforms are limited and, therefore, the substantial cohort effects that coincide with the reforms are almost entirely due to selection effects. In other words, the reforms have increased the self-screening among potential applicants, altering the targeting of the DI program towards a selection of applicants with unfavorable demographic and labor market characteristics.

Effects of broader job search requirements on unemployed workers
Unemployed workers are often too optimistic about their labor market prospects. These biased beliefs contribute to the slow exit out of unemployment and partly explain long-term unemployment (Mueller et al. 2021). In particular, unemployed workers anchor their reservation wage on their previous wage and search too often for work that resembles their previous job (Belot et al. 2019, Krueger and Mueller 2016). To offset these biased beliefs, stimulating unemployed workers to search more broadly may then positively affect labor market outcomes at a low cost to benefits administrations. An increasing number of OECD countries have therefore implemented policies requiring unemployed workers who are at risk of long-term unemployment to search and accept jobs beyond the occupation of their previous employment. A recent literature shows that *encouraging* unemployed workers to search more broadly indeed benefits them (Altmann et al. 2018, Belot et al. 2019). However, a question that remains is: *"What are the employment effects of mandatory broader job search requirements?"*

To answer this question, Chapter 4 evaluates the implementation of a program developed by the Dutch unemployment insurance (UI) administration that imposed broader job search. The program enforces broader job search requirements on unemployed workers who have been collecting UI benefits for at least six months. The program starts with an additional caseworker meeting, during which the worker's past job search behavior is evaluated. If the caseworker deems the past job search as too narrow, she has the discretion to mandate the unemployed worker to broaden their job search. The unemployed worker is obliged to comply with this requirement, which is monitored by the caseworker. In practice, it means that the unemployed worker should actively apply for jobs that are in different sectors, have a longer commuting distance, offer a lower wage, and require a lower level of education.

For the empirical evaluation, we use data from a large-scale field experiment conducted at the Dutch UI administration. A random subsample of about 130,000 unemployed workers has been invited to attend a mandatory caseworker meeting to discuss their job search strategies. The results from the experiment show, on average, that participation in the program increases employment and reduces the reliance on UI benefits, making the program cost effective for the UI administration. Next, we exploit the fact that caseworkers differ substantially in the rate at which they impose broader job search requirements and that unemployed workers are randomly assigned to caseworkers within local offices. Using this exogenous variation – i.e. caseworker stringency – as an instrumental variable, we can estimate the (isolated) causal effect of the broader search requirement on the unemployed workers that attend the caseworker meeting. The results show that imposing broader job search does not improve labor market outcomes. On the contrary, it reduces job finding and extends the period of collecting UI benefits. Additionally, the job characteristics are less favorable after the requirement, with individuals being less likely to have a permanent contract and working fewer hours per week. The implication is that the caseworker meeting in itself determines the positive effect of the broader search program, while the broader search requirement reduces the effectiveness of the program.

The adverse effects of imposing the broader search requirement appear to contradict the results from earlier studies that show positive effects of encouraging broader job search (e.g. Altmann et al. 2018, Belot et al. 2019, Skandalis 2019). However, an important difference is that the broader search requirement is mandatory, while other studies considered 'information treatments' that were not mandatory. It is likely that information treatments predominantly affect the beliefs about the returns to job search among a smaller and selective sample of unemployed workers who were too optimistic. In contrast, the mandatory nature of the broader search program implies that the treated population is larger. Furthermore, caseworkers will target the program to a different group than the respondents to an information treatment. Our results reveal that unemployed workers who are most likely to be imposed with the requirement, i.e. the unemployed workers who were searching narrowly, experience the largest adverse effects. These may be specialized workers who benefit most from a narrow job search and who were optimizing their job search before receiving the mandatory requirement. It is likely that specialized workers would be less responsive to an information treatment. Taken together, the difference in the targeting of the program can explain the adverse effects of the mandatory broader job search requirements, as opposed to the positive effects of information treatments found by other studies.

Effects of welfare application processing times and benefit prepayments
The trade-off of social insurance between ensuring income security and preserving work incentives is not limited to formal eligibility rules and benefit coverage. Instead, it also incorporates informal and administrative barriers that have been found to reduce the take-up and the ability of these programs to provide insurance (Currie 2006, Ko and Moffitt 2022). Concurrently, there is also a trade-off between providing timely income and ensuring benefit accuracy. Fast provision of insurance allows program applicants to improve consumption smoothing, but this may be at the cost of precision of the eligibility determination. This imprecision can be corrected at later stages, but may result in substantial repayments that can cause financial stress. Chapter 5 contributes to the literature as it aims to answer the question: *"What is the impact of welfare application processing times and benefit prepayments on the benefit and labor market outcomes of applicants in the Netherlands?"*

In the Netherlands, welfare benefits serve as a social safety net for all unemployed workers with insufficient means of subsistence. Eligibility for such benefits is determined based on the household income and household wealth. The applicants have to provide caseworkers with detailed information on their living situation, income and assets. The process of collecting and assessing this information can vary among individuals, particularly when the information provided was initially incomplete. Longer application processing times are an example of an informal administrative barrier that applicants might face when applying for welfare benefits. Despite the fact that processing times are ex ante unknown to the applicants, they might still change the targeting of the program or the labor market outcomes of applicants. In case of long processing times, applicants can request benefit prepayments to bridge the period without income.

The empirical analysis in Chapter 5 uses welfare benefits application data from Rotterdam, which has the highest take-up rate of welfare benefits in the Netherlands. The data reveal substantial variation in application processing times among applicants. This variation is partially due to differences in processing speed among the as-good-as randomly assigned caseworkers who act as gatekeepers. Similar to Chapter 4, the variation among caseworkers and the quasi-random assignment of applicants to caseworkers are exploited to estimate causal effects. The results show two offsetting effects of application processing times. On the one hand, some applicants are discouraged from continuing their welfare application due to longer processing times, and evidently therefore spend less time in welfare. These discouraged applicants tend to compensate for the lost benefits income through increased earnings, which suggests that the targeting of the program is increased. On the other hand, applicants who eventually receive welfare benefits after the longer processing times remain dependent on welfare for longer periods of time than those with fast applications. In other words, the findings suggest a trade-off, as improved targeting comes at the cost of worse labor market outcomes for the group of awarded applicants. Finally, the results indicate that prepayments of welfare benefits increase the employment of awarded applicants with long processing times. This suggests that the elimination of financial stress facilitates successful job search.

2 A panel data sample selection model to estimate life-cycle earning profiles: How important is selection into full-time and part-time employment?

Abstract

This paper proposes a new panel data sample selection model with 1) ordered discrete choices in the selection equation and 2) non-parametric unobserved heterogeneity in the equation of interest. This method is used to estimate life-cycle earnings profiles using high-quality administrative data. We compare conclusions regarding the existence and direction of selection into (part-time) work among men and women across different panel data sample selection techniques. The main conclusion is that our new approach is able to control for important unobserved heterogeneity from intensive labor supply choices with important consequences for the existence and direction of selection in (part-time) work.

A working paper version of this paper is published as Been et al. (2023). This chapter is co-authored by Jim Been and Marike Knoef. We gratefully acknowledge Netspar and Instituut Gak for their financial contribution to this project. We thank all participants of the Netspar Workshop on Pensions, Retirement, and the Financial Position of the Elderly, the CPB Netherlands Bureau for Economic Policy Analysis Seminar and the conference participants at the International Association for Applied Econometrics, the Econometric Society European Meetings, the Netspar Pension Day, the International Workshop on Pensions, Insurance and Savings, the International Panel Data Conference, and the KVS New Paper Sessions. More specifically, we would like to thank Hippolyte d'Albis, Rob Alessie, Jonneke Bolhaar, Jan Bonenkamp, Lans Bovenberg, Pavel Cizek, Bart Cockx, Stefan Hochguertel, Egbert Jongen, Adriaan Kalwij, Pierre Koning, Jordy Meekes, Chris Muris, Pedro Raposo, Peter van Santen, Joanna Tyrowicz, Ola Vestad, Daniel van Vuuren, Bas ter Weel, Jeffrey Wooldridge, and Bram Wouterse for providing us with valuable comments at different stages of the paper.

2.1 Introduction

Estimating earnings profiles is crucial for understanding earnings dynamics and life-cycle consumption and savings decisions. Since earnings are only observed among those who work, simply estimating an earnings model without taking into account the non-random selection into work leads to serious inconsistent estimates of earnings (Heckman 1979), even in the case of panel data (Solon 1988). In light of this selection issue, many of the earnings processes estimated in the literature focus on prime age males as it can be argued that this group is most likely to work (full time) and least likely to self-select into work.[1] This also holds for recent estimates of life-cycle wages (Lagakos et al. 2018), which are estimated solely on full-time public sector male workers. As a consequence, conclusions from such estimates may not be generalizable to women[2] and older men[3] for whom working (full time) is less self-evident. Hence, it is important to derive models that correct for sample selection with panel data and test the assumption of no selection into (full-time) work among both men and women to get an impression of the generalizability of results for prime age males. In this paper, we test if there is additional information hidden in selection into part-time versus full-time employment compared to selection in employment at the extensive margin to estimate selection-corrected earnings profiles.

The first panel data sample selection models are derived by Wooldridge (1995), Kyriazidou (1997), and Rochina-Barrachina (1999) who build upon the sample selection model of Heckman (1979).[4] The three methods differ in the assumptions and estimation of the first-stage and second-

[1] See, for example, Baker (1997), Baker and Solon (2003), Daly et al. (2022), Gottschalk and Moffitt (1994), Guvenen (2009), Heathcote et al. (2010), Lillard and Weiss (1979), Lillard and Willis (1978), Meghir and Pistaferri (2004, 2010), Moffitt and Gottschalk (2012), Pischke (1995), Storesletten et al. (2004).

[2] Ermisch and Wright (1993), for example, find positive selection of women into full-time work in the UK.

[3] Myck (2010), for example, shows that lower paid older men are more likely to remain in employment than higher paid older men in the UK, i.e. negative selection. This is consistent with evidence from Hanoch and Honig (1985) for American men and women.

[4] A newer strand of literature extends these models in the direction of making fewer parametric assumption (Semykina and Wooldridge 2018), allowing for endogenous regressors (Charlier et al. 2001, Dustmann and Rochina-Barrachina 2007, Semykina and Wooldridge 2010), and dynamic models (Semykina and Wooldridge 2013).

stage of the model.[5] Both Wooldridge (1995) and Rochina-Barrachina (1999) propose parametric estimators of the linear panel data model under sample selection when the explanatory variables are strictly exogenous. Kyriazidou (1997) derives a semi-parametric estimator for such models. Wooldridge (1995) proposes estimation in levels and makes parametric assumptions on the unobserved individual-specific heterogeneity in both the first- and second-stage. Rochina-Barrachina (1999) proposes estimation in first-differences and makes no parametric assumptions on the unobserved individual-specific heterogeneity in the second-stage and exploits the autoregressive nature of participation to condition on unobserved individual-specific heterogeneity.

All aforementioned estimators assume that selection into earnings is a matter of selecting into work versus non-work (i.e. extensive labor supply decisions) and, therefore, use a binary selection rule. A different strand of literature has not extended the model of Heckman (1979) in the direction of panel data, but by using non-binary choices in the selection equation. Extending selection into work beyond a binary selection rule and allowing for labor supply decisions at the intensive margin may add important unobserved information to the wage equation, such as leisure-time preferences. Only few papers, like Zabalza et al. (1980), Nakamura and Nakamura (1983), Hotchkiss (1991), and Ermisch and Wright (1993), have argued to use an ordered selection rule[6] to capture self-selection into full-time and part-time work. Unlike the first-mentioned strand of literature, these models are only applicable to cross-sectional data and not to panel data.

To be able to distinct between age- and cohort effects in the estimation, it is important to use a panel data sample selection model to estimate the earnings over the life-cycle. The first attempt to combine panel data with adjustments for self-selection into work, and thereby extend the canonical sample selection model of Heckman (1979) to panel data, is by Hanoch and Honig (1985) although their model only uses cohort- and period fixed

[5]Dustmann and Rochina-Barrachina (2007) show how these different assumptions affect the application to real world panel data.

[6]Using an ordered selection rule is consistent with Averett and Hotchkiss (1997), Tummers and Woittiez (1991), Van Soest (1995) who argue that labor supply is semi-continuous.

effects and no individual fixed effects. The first paper to bridge the gap between the two extensions of the Heckman (1979) sample selection model is Dustmann and Schmidt (2000). Dustmann and Schmidt (2000) is the first to use an ordered selection rule in a panel data sample selection model by extending the approach in Wooldridge (1995) from a binary to an ordered selection rule. Like Wooldridge (1995), both the first- and second stage make parametric assumptions about the unobserved effects (Dustmann and Schmidt 2000).

In this paper, we propose a new panel data sample selection model with an ordered selection rule. Compared to Dustmann and Schmidt (2000), we make no parametric assumptions on the unobserved individual-specific heterogeneity in the wage equation and allow to condition on the unobserved individual-specific heterogeneity in participation by exploiting the autoregressive nature of labor supply decisions like Rochina-Barrachina (1999). Compared to Rochina-Barrachina (1999), we use an ordered instead of binary selection rule.

Using administrative panel data that are representative for the Netherlands in the period 2001-2014, we show how an ordered selection rule in the framework of Rochina-Barrachina (1999) can provide additional information for the estimation of earnings over the life-cycle compared to a binary estimator. This may especially hold for the Netherlands where the prevalence of part-time work is internationally high among both men (2020: 28.5%) and women (2020: 73.8%) (OECD 2020). Furthermore, rich administrative data allows us to use very flexible functional forms, such as semi-parametric age effects like in Kalwij and Alessie (2007).

The empirical application of our panel data sample selection model to estimating life-cycle earnings shows that it is important to take self-selection in the intensive margin of labor supply into account. When correcting for the labor supply decision on the intensive margin, we find positive selection into part-time work for both men and women. This means that men and women with more affluent characteristics self-select into part-time employment. Not correcting for such selection leads to an overestimation of part-time earnings. For full-time work, we find positive selection for women only. For full-time men, we find no statistical evidence for selection. Hence, the generally assumed absence of selection into work among men in the literature is only true if full-time work is considered. Our findings regarding the existence and direction of selection are in stark contrast with conclusions based on applying the Rochina-Barrachina (1999) method – with a binary selection rule, which show negative selection into part-time work for men (and none for women) and full-time work for both men and women. Hence, our new approach exploits important unobserved information that stays hidden otherwise and which has implications for understanding who selects into (part-time) work.

Applying our method to estimate life-cycle earnings profiles, we show that correcting for selection changes the earnings estimates significantly and results in different shapes of the earnings-age curve over the life-cycle compared to regular first-differences estimates. With our proposed method, we find that earnings in full-time employment peak later in the life-cycle than earnings in part-time employment. This is true for both men and women. Additionally, these differences are amplified when correcting for selection into full-time and part-time employment.

The remainder of the paper is organized as follows. In the next section, we show the importance of part-time employment in the Netherlands by describing the institutional setting. In section 2.3 contains a description of the data and shows the employment, earnings and wages over the life-cycle. section 2.4 describes the new model and explains the empirical specification. section 2.5 reports the main estimation results. In section 2.6, we investigate the importance of an ordered selection rule compared to a binary rule (the estimator proposed by Rochina-Barrachina 1999). Finally, section 2.7 concludes.

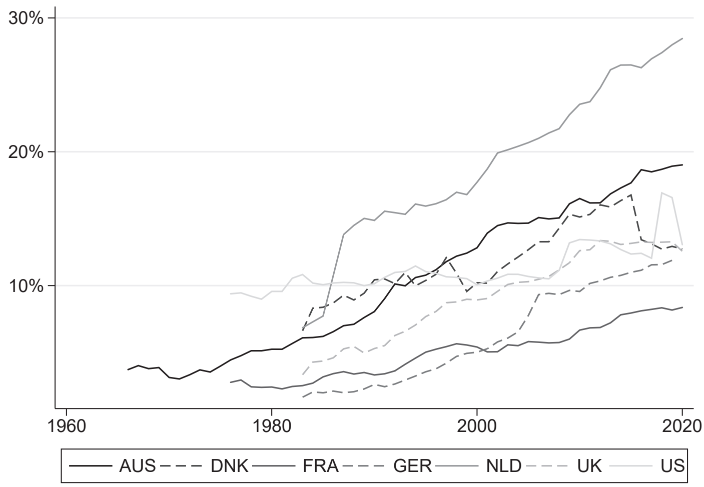
2.2 Institutional background: Part-time employment in the Netherlands

In Figure 2.1, we show the development of part-time employment for a selection of OECD countries for men and women, respectively. From the figures, four general conclusions stand out. First, the incidence of part-time employment is substantial in OECD countries and has been steadily increasing since the late 1960s. Second, part-time employment has in all countries a higher incidence among women than among men. In 2020, the OECD average of part-time employment as a percentage of total employment was 12.4% for men and 31.3% for women. Third, much of the increase in part-time employment across countries is largely due to increasing part-time employment among men (who have higher overall employment rates). Between 1966 and 2020, the incidence of male and female part-time employment grew with 235% (from 3.7% to 12.4%) and 30% (from 24.0% to 31.2%), respectively. Fourth, part-time employment is much more prevalent in the Netherlands than in any of the other (reported and non-reported) OECD countries. This applies to both men (28.5% in 2020) and women (73.8%). These statistics show the relevance of analyzing the selection effects in the intensive margin as the popularity of part-time employment has widely increased and is no longer specific to women only.

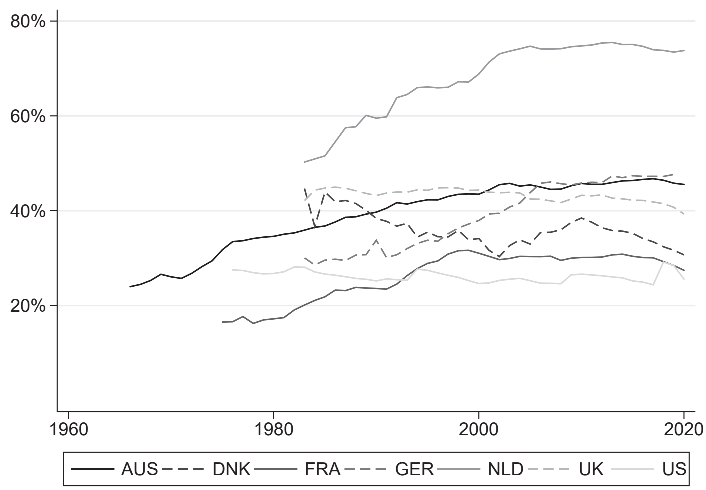
Unlike other countries, most of the part-time employment is on a voluntary basis in the Netherlands (Visser et al. 2004). In the Netherlands, employers are in principle obliged to accept a request for part-time employment of an employee. According the labor law (*Wet Aanpassing Arbeidsduur, WAA*), employees are allowed to request for a decrease (or increase) in their contractual employment hours without any further specification as to the reason why. This only applies to employers with more than 10 employees, employees working at the employer for at least one year, and has a two-month notice. Such a request can be made once a year. The *WAA* implies that part-time employment is highly institutionalized in the Netherlands. Prior to the *WAA*, which was introduced in February 2000, many collective bargaining agreements included the possibility for part-time employment requests. Since January 2016, the flexibility of choosing the number of hours has been extended to flexibility in the daily work hours and location by a law stimulating flexible work (*Wet Flexibel Werken, WFW*). To summarize, these labor laws indicate that flexible work, including part-time work is highly facilitated and accepted in the Netherlands. Additionally, part-time work of couples is facilitated through the tax system, including child care subsidies.

Figure 2.1: Incidence of part-time employment among (a) men and (b) women in OECD countries.

(a) Men (OECD 2020).



(b) Women (OECD 2020).



2.3 Data

2.3.1 Data selection and variable definitions

We use two data sets for our analysis: (i) administrative tax records from the Dutch Income Panel Study from the Netherlands (IPO) for the years 2001-2014, and (ii) data on working hours from the Dutch payroll administration for the years 2001-2014. The IPO data set consists of an administrative panel data set for a representative sample from the Dutch population of, on average, 95,000 selected individuals per year who are followed longitudinally.[7] The data set contains detailed information on personal and household income, labor market status and demographics.

The main advantages of using these administrative data sets compared to using survey data for our analysis are the large sample size, the long panel aspect of the data, the accuracy of tax data compared to self-reported survey answers, and representativeness. Interestingly, the data include a "part-time employment factor", that measures the proportion of work a person has undertaken in relation to a full-time job over the course of a year. A factor of 1 indicates that a person worked full time for the entire year. However, a factor of 0.5 can have two different interpretations: (i) the person worked half of a full-time contract throughout the entire year, or (ii) the person worked full-time for half of the year. We are particularly interested in (i) and not in (ii). Appendix 2.C.2 describes year-to-year transitions in labor supply categories and shows that most individuals stay in the same category from year to year. The dependent variable in our analysis is the full-time equivalent (before tax) wage expressed in (log) 2015 euros. To construct the full-time equivalent wages, we divide yearly earnings by the part-time employment factor mentioned above. Inevitably, we do not observe wages for people who are not wage employed.[8]

[7]Sampling is based on individuals' national security number, and the selected individuals are followed together with their household members for as long as they are residing in the Netherlands on December 31 of the sample year. Individuals born in the Netherlands enter the panel for the first time in the year of their birth, and immigrants to the Netherlands in the year of their arrival.

[8]This includes the self-employed. Following Bardasi and Gornick (2008) we categorize all persons in non-paid employment as 'unemployed.'

In this study we select individuals between the ages of 24 and 64 (387,841 observations for men and 385,298 observations for women). To reduce measurement error, we restrict the sample in the following ways. First, per year, we regard observations below the minimum wage and in the top 1% of the wage distribution as outliers and exclude these from the analysis. Second, per year, observations with the 1% largest decreases or increases in relative year-to-year-changes in the full-time equivalent wage rate are considered outliers and removed. It is likely that such substantial changes in year-to-year wages are a consequence of measurement error in the part-time employment factor (due to the definition of this measure defined by Statistics Netherlands, as explained above). Third, since people who leave employment as a result of a disability might result in measurement error of the part-time employment factor, we drop observations of workers who received disability benefits during (part of) the year. Fourth, we exclude individuals who worked less than one-twelfth of a full-time year. We argue they worked too little to calculate a reliable (full-time equivalent) wage. Fifth, we restrain the sample to individuals who remain in the same labor supply category.[9] This reduces our sample to 266,950 males and 265,305 females. Finally, we use population weights to account for representativity with respect to age, gender, marital status, province, household size and the age of the head of the household.

Descriptive statistics 2.3.2

Earnings

Figure 2.2 presents average earnings profiles for men and women (including those who do not work), with eminent differences between them. The earnings profile of men depicts the typical inverted U-shape moderately well as the wages grow over the life-cycle and only declines sharply in the years in which people retire. For men, average earnings are about 25,000

[9]Appendix 2.C.2 shows that most people remain in the same labor participation category. A change is often caused by individuals becoming unemployed or starting a job during the calendar year and in this case we can not determine for all years the actual labor supply category during the part of the year that people are at work.

euros per year at the age of 25 and grow up to just over 40,000 euros per year around the age of 50. After the age of 50 we observe a decline in average yearly earnings, with the largest drop in earnings around the age of 60. The decline in average earnings among older men may be explained by several phenomena: (i) early retirement, (ii) drops in hours worked preceding retirement (partial retirement), (iii) older workers receiving lower wages and (iv) birth-cohort effects. Negative selection into work at older ages might strengthen this decline (Casanova 2010, Myck 2010).

For women, we see that the earnings are declining after the age of 30. We observe that a 25 year-old female earns about 22,000 euros per year on average. Around the age of 35 (when most women raise their children) earnings are relatively low, probably because of a drop in the labor force participation and/or the number of hours work. Thereafter, earnings remain fairly stable and as from the age of 50 earnings decrease again. However, we should keep in mind that there are profound cohort effects among women. These cohort effects – namely the increased labor force participation and higher educational attainment among younger generations of women – can likely explain the substantial vertical differences between the cohorts among women (which we do not see for men).

Participation

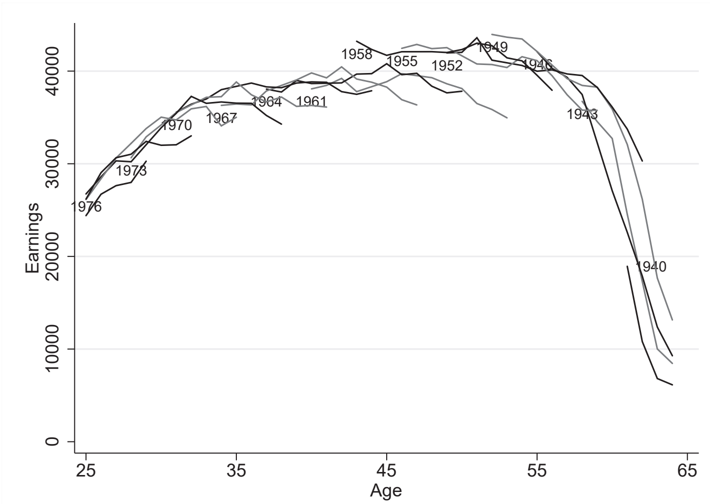
Unemployment and part-time employment shape the earnings profiles as shown in figure 2.2.[10] Figure 2.3 therefore shows the percentage in full-time and part-time employment over the life-cycle for different cohorts for men and women separately. In 2001 about 70% of all men in all cohorts seem to work full-time until the age of 55.[11] However, between 2001 and 2014 it seems at all ages about 10% of the men moved from a full-time to a part-time job. Most men seem to leave the labor market at older ages. About 20% is unemployed at the age of 55 and this increases to about 80% at the age of 64. These changes in employment are almost entirely

[10] Recall that in this paper we define people to be unemployed when they do not earn labor income from paid employment.

[11] We assume persons to be working full-time if the part-time employment factor is equal to one. Every person with a part-time employment factor of smaller than one is considered to be working part-time or unemployed.

Figure 2.2: Life-cycle earnings of men (a) and women (b)

(a) Mean earnings men



(b) Mean earnings women

confined to transitions from full-time employment into unemployment. As expected, younger cohorts of men retire later.

The employment patterns for women are different than those for men, with lower employment rates and more part-time work, especially among older women. This is also depicted in Table 2.2, where we show how participation has evolved over time. Whereas participation of men is fairly stable or even declined over time, we observe a substantial increase in women's participation (10%-points in 15 years). Although the literature generally suggests that women's labor supply is largely affected by changes in child care subsidies, see among others (Berger and Black 1992), such effects are found to be small in the Netherlands (Bettendorf et al. 2015).

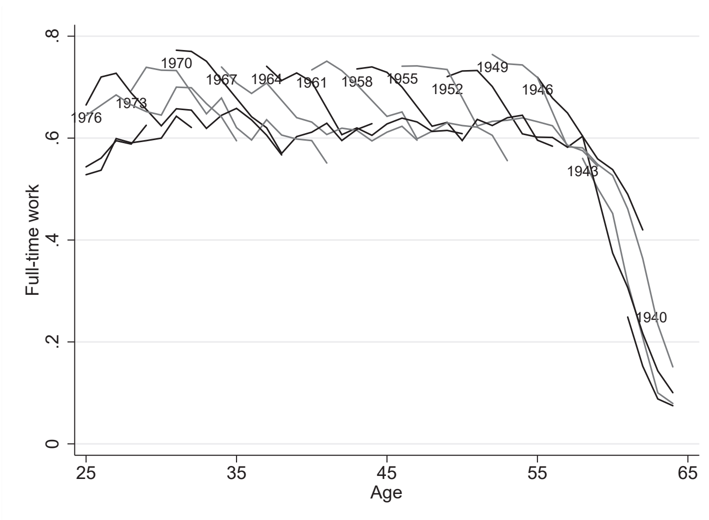
For women, we observe a substantial drop in full-time employment around the age at which they raise children. Before the age of 30 about 40-50% of women work full-time and this drops to about 20% at the age of 40, after which it stays constant until the age of 55. This is in line with the findings of Bosch et al. (2010). Part-time work, on the other hand, increases between the age of 30 and 40 from about 40 to 55%. The large shift from full-time employment to part-time or unemployment also largely explains the earnings decline as depicted in panel (b) of Figure 2.2. Similarly to men, women leave the labor market at older ages. Finally, employment is much higher for younger cohorts than for older cohorts of women.

Wages

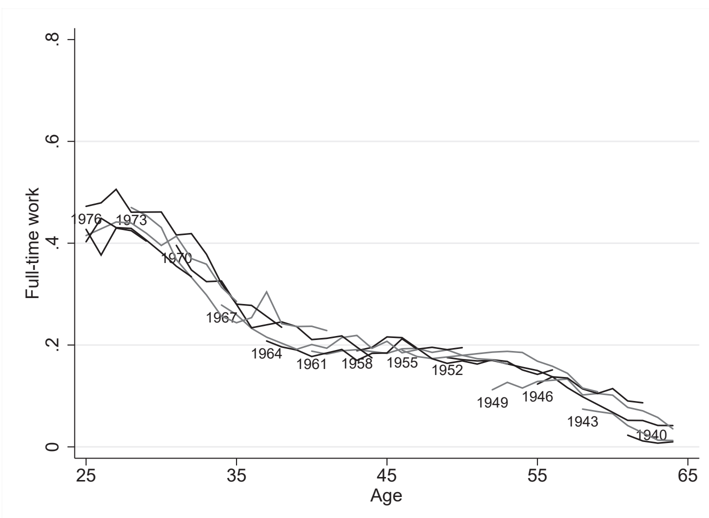
Figure 2.4 shows the average yearly wage (on a full-time basis) for men and women in full-time and part-time employment. Although we found an inverted U-shape for life-cycle earnings of men, wages are increasing over the life-cycle. Average yearly wages are approximately 33,000 euros at the age of 25 for men in full-time employment, and about 30,000 euros in part-time employment. Both full-time and part-time wages increase with age, with the largest changes in the beginning of the career. Full-time wages are on average 53,000 euros before retirement, while part-time wages end around 50,000 euros.

Figure 2.3: Percentage of men and women in full-time and part-time employment

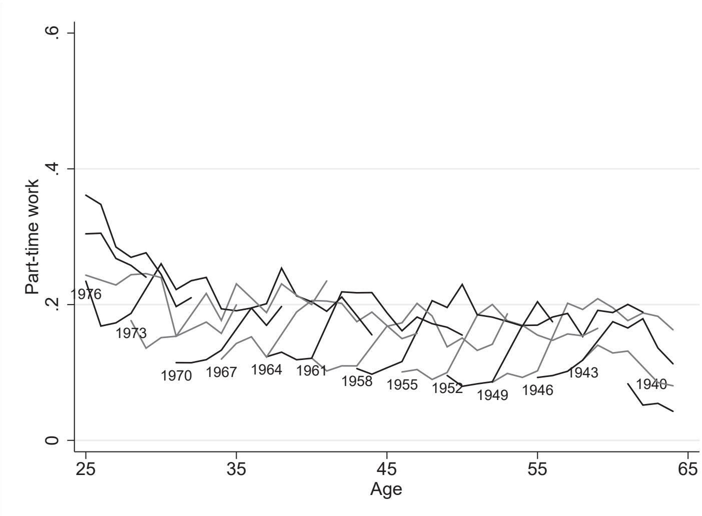
(a) Full-time employment (%) of men



(b) Full-time employment (%) of women



(c) Part-time employment (%) of men



(d) Part-time employment (%) of women

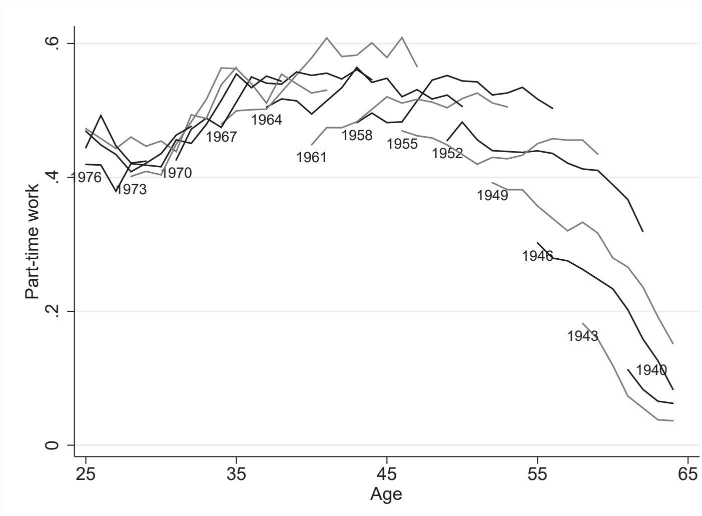
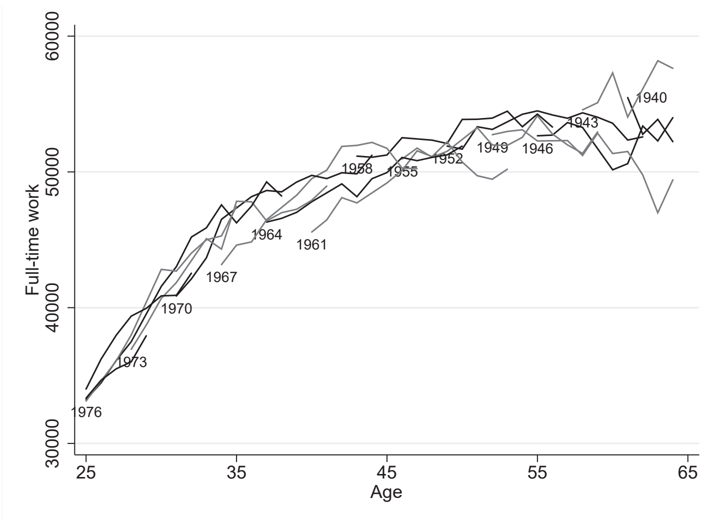
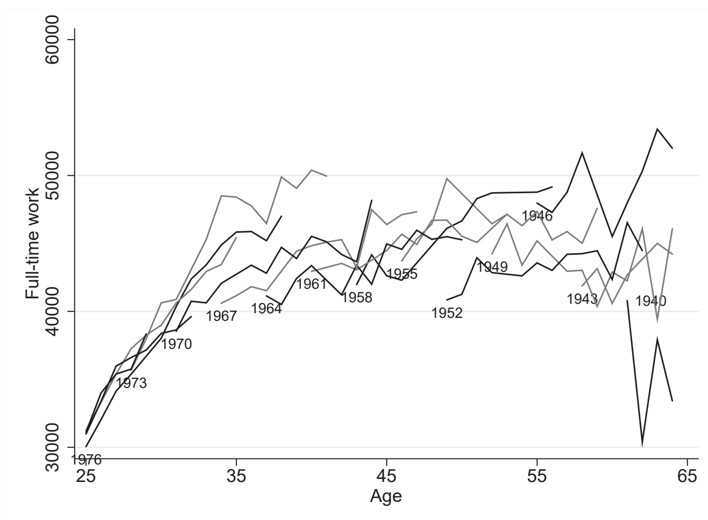


Figure 2.4: Wage of men and women in full-time and part-time employment

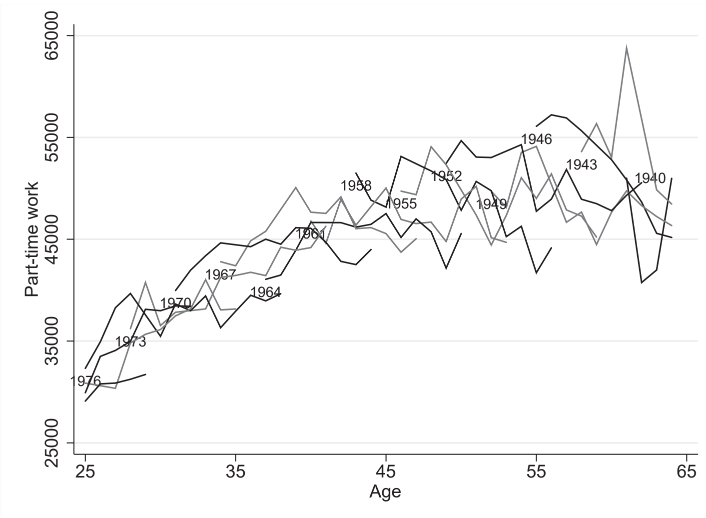
(a) Full-time wages of men



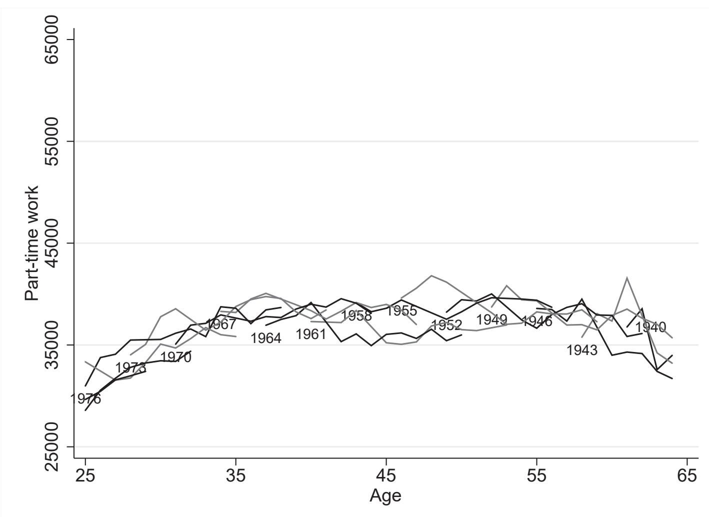
(b) Full-time wages of women



(c) Part-time wages of men



(d) Part-time wages of women



Female yearly average wages also show large increases over the life-cycle. Although part-time and full-time wages both increase with age, full-time wages show a larger and more persistent growth. This results in an increase from around 30,000 euros at the age of 25 (for both full-time and part-time work) to 45,000 euros at the age of 45 for women in full-time employment, and less than 40,000 euros for those in part-time employment. Thereafter wages remain relatively constant. Appendix 2.A also shows the trends in (part-time) participation and wages over the time period 2001-2014 for both men and women.[12]

Model 2.4

The previous section showed that full-time wages are higher than the full-time equivalent of part-time wages and that wages grow over the life-cycle. However, to be able to correctly estimate the life-cycle earnings profiles, we should take into account that we only observe wages for those individuals who are working and that workers might select into (part-time) employment. As a result, these workers might differ in both observed as well as unobserved characteristics. Accordingly, the goal of the remainder of the paper is to estimate life-cycle wage profiles for men and women in both full-time and part-time employment while controlling for selection on observed and unobserved heterogeneity. To do so, we first introduce our panel data sample selection model with an ordered selection rule and no parametric assumptions on the individual-specific heterogeneity in the wage equation.

[12]We observe a discontinuity in the hours worked around 2006, which especially affects our part-time employment variable. This discontinuity is also addressed by De Nardi et al. (2021), who show similar patterns in (part-time) employment for the 2001-2014 and the post-2006 periods. We test the robustness of our results using a dummy for the post-2006 period in the wage equation. The dummy is significant, however, with a coefficient of 0.016 the effect is not substantial. Our main conclusions remain the same when adding this dummy.

2.4.1 Panel data sample selection model

Suppose that we have two individuals A and B with the same observed characteristics. A is working part-time and B is working full-time. B most likely has more favorable unobserved characteristics (like ability and motivation) which lead both to more hours worked and a higher wage rate. As long as these unobserved characteristics are time-invariant we can use a individual fixed-effects data model to take this into account. However, it is likely that there are also time variant unobserved characteristics such as time variant unobserved ability or health that influence both participation, the number of hours worked, and the wage rate of individual i in period t. To take this into account we use a panel data sample selection model that models both wages and labor force participation at the extensive and intensive margin. The model can be written as follows:

$$y_{it}^* = x_{it}\beta + \alpha_i + u_{it} \quad i = 1, ..., N \quad t = 1, ..., T \tag{2.1}$$

$$h_{it}^* = z_{it}\gamma_t + \eta_i + v_{it} \tag{2.2}$$

$$y_{it} = \begin{cases} y_{it}^* & \text{if } h_{it}^* > \delta_{1t} \\ \text{unobserved} & \text{otherwise} \end{cases} \tag{2.3}$$

$$h_{it} = \begin{cases} 0 \quad \text{(no participation)} & \text{if } h_{it}^* \leq \delta_{1t} \\ 1 \quad \text{(part-time)} & \text{if } \delta_{1t} < h_{it}^* \leq \delta_{2t} \\ 2 \quad \text{(part-time)} & \text{if } \delta_{2t} < h_{it}^* \leq \delta_{3t} \\ \vdots & \\ J \quad \text{(full-time)} & \text{if } \delta_{Jt} < h_{it}^* \end{cases} \tag{2.4}$$

where y_{it} is the observed wage for individual i in period t. h_{it} is the observed labor force participation containing J categories of labor (no labor force participation, several categories of part-time labor force participation, and full-time labor force participation). h_{it}^* indicates the latent equivalent. x_{it} and z_{it} are vectors of individual's observed characteristics. For identification, z_{it} includes variables that do not appear in x_{it}. β and γ_t are unknown parameter vectors to be estimated and α_i and η_i are unobserved individual-specific effects, which are possibly correlated with x_{it} and z_{it}.

They capture education, time-invariant ability, and cohort effects that incorporate participation and productivity differences between generations (Kapteyn et al. 2005). δ_{jt} with $j = \{1, .., J\}$ are time-specific thresholds to be estimated. Finally, u_{it} and v_{it} are unobserved disturbances which are assumed to follow a normal distribution with mean zero and variances $\sigma_{u,t}$ and $\sigma_{v,t}$.

Presumably, u_{it} and v_{it} are correlated and therefore we need to incorporate selection into the wage equation. Furthermore, because u_{it} is likely to be serially correlated, we use the first difference (FD) estimator in the main equation.[13] FD requires a weaker form of exogeneity than what is required for FE. Namely $E(x_{it}u_{is}) = 0$ for $s = t, t-1$ instead of $E(x_{it}u_{is}) = 0$ for $s = 1, 2, ...T$. Thus, FD allows that past wage shocks affect the explanatory variables later in life ('feedback effects'), which may be necessary as the selection correction terms (which we explain below) may not be strictly exogenous in the main equation.

We use a discrete choice model to model the allocation to part-time and full-time jobs. Discrete choice models have been used repeatedly in the literature to model the allocation to part-time and full-time jobs.[14] In this way we account for mass points in the number of hours worked (e.g. because of work hour restrictions) like Van Soest (1995). A drawback is the incomplete use of available data, however, the number of labor supply categories J can be increased to allow for more differentiation in labor supply, but increasing J goes at the cost of statistical power per category. The optimal number of categories J is found to be arbitrary (Franses and Cramer 2010).

We can only observe wage differences for those observations for which an individual has worked at both time t and $t-1$:

$$y_{it} - y_{it-1} = \begin{cases} y^*_{it} - y^*_{it-1} & \text{if } h^*_{it-1} > \delta_{1,t-1} \text{ and } h^*_{it} > \delta_{1,t} \\ \text{unobserved} & \text{otherwise} \end{cases} \tag{2.5}$$

[13]We find evidence of serial correlation in wages in our data. The Wooldridge (2002) test for autocorrelation in panel data rejects the null-hypothesis of no first-order autocorrelation for men (*F-stat=7.22*, p-value=0.0072) and women (*F-stat=56.46*, p-value=0.0000).

[14]See, for example, Duncan and Weeks 1997, Dustmann and Schmidt 2000, Ermisch and Wright 1993, Hotchkiss 1991, Nakamura and Nakamura 1983, Zabalza et al. 1980.

where

$$y_{it}^{*} - y_{it-1}^{*} = (x_{it} - x_{it-1})\beta + (u_{it} - u_{it-1}) \tag{2.6}$$

Since the first-difference in wages ($y_{it} - y_{it-1}$) is only observed if a person actually worked in both periods ($h_{it-1}^{*} > \delta_{1,t-1}$ and $h_{it}^{*} > \delta_{1,t}$), estimating Equation (2.6) by OLS would yield inconsistent estimates of β as the conditional expectation of the error term is unlikely to be zero due to correlation between u_{it} and v_{it}. Therefore, we need to calculate the expectation conditional on participation. We do not only know whether someone is participating, but also whether someone is participating full-time ($h_{it}^{*} > \delta_{Jt}$) or whether someone is in some part-time labor supply category ($\delta_{jt} < h_{it}^{*} \leq \delta_{j+1,t}$ where $j = 1, 2, ...J - 1$). This gives us additional information about the unobserved characteristics. The conditional expectation of the first differences can be written as follows:

$$\begin{aligned} & E[y_{it} - y_{it-1} | \, x_{it}, x_{it-1}, z_{it}, z_{it-1}, \delta_{j,t} < h_{it}^{*} \leq \delta_{j+1,t}, \delta_{j,t-1} < h_{it-1}^{*} \leq \delta_{j+1,t-1}] \\ & = (x_{it} - x_{it-1})\beta \\ & + E[u_{it} - u_{it-1} | \, x_i, z_i, \delta_{j,t} < h_{it}^{*} \leq \delta_{j+1,t}, \delta_{j,t-1} < h_{it-1}^{*} \leq \delta_{j+1,t-1}] \end{aligned} \tag{2.7}$$

where j is the working hours category of individual i at time t. For persons who do not work at time t, we define $\delta_{0,t} = -\infty$. Similarly, for persons engaged in full-time work at time t, $\delta_{J+1,t} = \infty$.

Following Mundlak (1978) we parameterize the individual specific effect in the selection equation (2.2) as a linear function of the average explanatory variables over time plus a random individual specific effect that is assumed to be independent of the explanatory variables:

$$\eta_i = \overline{z_i}\theta + c_i \tag{2.8}$$

where θ is an unknown parameter vector to be estimated and c_i is assumed to be a normally distributed random variable with mean zero and variance σ_c. Substituting (2.8) into (2.2) yields:

$$h_{it}^{*} = z_{it}\gamma_t + \overline{z_i}\theta + \mu_{it} \tag{2.9}$$

where $\mu_{it} = c_i + v_{it}$. Given the distributional assumptions it holds that $\mu_{it} \sim N(0, \sigma_{\mu,t})$, where $\sigma^2_{\mu,t} = \sigma^2_c + \sigma^2_{v,t}$. Furthermore, μ_{it} is allowed to be serially dependent (this is necessary, because of the term c_i). Denote the correlation coefficient of μ_{it-1} and μ_{it} by ρ_t. Substituting (2.9) into the last term of (2.7) gives us

$$\begin{aligned}
&\mathrm{E}[u_{it} - u_{it-1} | x_i, z_i, \delta_{j,t} < h^*_{it} \leq \delta_{j+1,t}, \delta_{j,t-1} < h^*_{it-1} \leq \delta_{j+1,t-1}] \\
&= \mathrm{E}[u_{it} - u_{it-1} | x_{it}, x_{it-1}, z_{it}, z_{it-1}, a_{it-1} \leq \frac{\mu_{it-1}}{\sigma_{\mu,t-1}} < b_{it-1}, a_{it} \leq \frac{\mu_{it}}{\sigma_{\mu,t}} < b_{it}]
\end{aligned} \tag{2.10}$$

where

$$a_{it-1} = (-\delta_{j+1,t-1} + z_{it-1}\gamma_t + \overline{z_i}\theta)/\sigma_{\mu,t-1} \tag{2.11}$$

$$b_{it-1} = (-\delta_{j,t-1} + z_{it-1}\gamma_t + \overline{z_i}\theta)/\sigma_{\mu,t-1} \tag{2.12}$$

$$a_{it} = (-\delta_{j+1,t} + z_{it}\gamma_t + \overline{z_i}\theta)/\sigma_{\mu,t} \tag{2.13}$$

$$b_{it} = (-\delta_{j,t} + z_{it}\gamma_t + \overline{z_i}\theta)/\sigma_{\mu,t} \tag{2.14}$$

The errors $[(u_{it} - u_{it-1}), \mu_{it-1}, \mu_{it}]$ are assumed to be trivariate normally distributed conditional on x_{it-1}, x_{it}, z_{it-1} and z_{it}.

Following the method of the two-step approach proposed by Heckman (1976, 1979), we work out (2.10) to obtain correction terms, that can be added as additional regressors to the main equation (the wage equation). Rochina-Barrachina (1999) also extends Heckman's sample selection technique to the case where one correlated selection rule in two different time periods generates the sample. We extend this further by allowing for an ordered selection indicator.

In order to work out (2.10), we take the derivative of the moment generating function of the doubly truncated trivariate normal distribution with respect to $t - 1$ and evaluate this function in $t = 0$. For details regarding the derivation, we refer to Appendix 2.B. The derivation gives us

$$\mathrm{E}(u_{it} - u_{it-1} | x_{it}, x_{it-1}, z_{it}, z_{it-1}, a_{it-1} \leq \frac{\mu_{it-1}}{\sigma_{\mu,t-1}} < b_{it-1}, a_{it} \leq \frac{\mu_{it}}{\sigma_{\mu,t}} < b_{it}) = \tag{2.15}$$

$$\pi_1\lambda_{1it}(\rho_t, a_{it-1}, a_{it}, b_{it-1}, b_{it}) + \pi_2\lambda_{2it}(\rho_t, a_{it-1}, a_{it}, b_{it-1}, b_{it})$$
$$+ \pi_3\lambda_{3it}(\rho_t, a_{it-1}, a_{it}, b_{it-1}, b_{it}) + \pi_4\lambda_{4it}(\rho_t, a_{it-1}, a_{it}, b_{it-1}, b_{it})$$

where

$$\lambda_{1it}(\rho_t, a_{it-1}, a_{it}, b_{it-1}, b_{it}) = \frac{\phi(b_{it-1})\left[\Phi\left((b_{it} - \rho_t b_{it-1})/\sqrt{1-\rho_t^2}\right) - \Phi\left((a_{it} - \rho_t b_{it-1})/\sqrt{1-\rho_t^2}\right)\right]}{\Phi^2(b_{it-1}, b_{it}; \rho_t) - \Phi^2(a_{it-1}, a_{it}; \rho_t)} \tag{2.16}$$

$$\lambda_{2it}(\rho_t, a_{it-1}, a_{it}, b_{it-1}, b_{it}) = \frac{\phi(a_{it-1})\left[\Phi\left((b_{it} - \rho_t a_{it-1})/\sqrt{1-\rho_t^2}\right) - \Phi\left((a_{it} - \rho_t a_{it-1})/\sqrt{1-\rho_t^2}\right)\right]}{\Phi^2(b_{it-1}, b_{it}; \rho_t) - \Phi^2(a_{it-1}, a_{it}; \rho_t)} \tag{2.17}$$

$$\lambda_{3it}(\rho_t, a_{it-1}, a_{it}, b_{it-1}, b_{it}) = \frac{\phi(b_{it})\left[\Phi\left((b_{it-1} - \rho_t b_{it})/\sqrt{1-\rho_t^2}\right) - \Phi\left((a_{it-1} - \rho_t b_{it})/\sqrt{1-\rho_t^2}\right)\right]}{\Phi^2(b_{it-1}, b_{it}; \rho_t) - \Phi^2(a_{it-1}, a_{it}; \rho_t)} \tag{2.18}$$

$$\lambda_{4it}(\rho_t, a_{it-1}, a_{it}, b_{it-1}, b_{it}) = \frac{\phi(a_{it})\left[\Phi\left((b_{it-1} - \rho_t a_{it})/\sqrt{1-\rho_t^2}\right) - \Phi\left((a_{it-1} - \rho_t a_{it})/\sqrt{1-\rho_t^2}\right)\right]}{\Phi^2(b_{it-1}, b_{it}; \rho_t) - \Phi^2(a_{it-1}, a_{it}; \rho_t)} \tag{2.19}$$

$\xi_{it} \equiv (u_{it} - u_{it-1}) - (\pi_1\lambda_{1it} + \pi_2\lambda_{2it} + \pi_3\lambda_{3it} + \pi_4\lambda_{4it})$ has a conditional expectation of zero by construction. This means that when we assume that we can form consistent estimates of the λ's, we can consistently estimate β as well.

Intuitively, it makes sense that we have four correction terms since the selection indicator in the panel data sample selection model is a combination of the ordered probit model of Dustmann and Schmidt (2000) (leading to a doubly truncated bivariate normal distribution with two selection terms for the lower- and upper threshold) and the bivariate probit model of Rochina-Barrachina (1999) (leading to a singly truncated trivariate normal distribution with two selection terms for the thresholds at time t and $t-1$). The bivariate ordered probit model in our method leads to a doubly truncated trivariate normal distribution with two selection terms for the lower- and upper threshold and two selection terms for the thresholds at time t and $t-1$.

Estimation 2.4.2

In the first step of the estimation procedure we deal with the selection equation. For each $s = \{t, t-1\}$ we estimate the following bivariate ordered probit model

$$h^*_{it-1} = z_{it-1}\gamma_{t-1} + \overline{z_i}\theta_{t-1} + \mu_{it-1} \tag{2.20}$$

$$h^*_{it} = z_{it}\gamma_t + \overline{z_i}\theta_t + \mu_{it} \tag{2.21}$$

$$h_{is} = \begin{cases} 0 \quad \text{(no participation)} & \text{if } h^*_{is} \leq \delta_{1s} \\ 1 \quad \text{(part-time)} & \text{if } \delta_{1s} < h^*_{is} \leq \delta_{2s} \\ 2 \quad \text{(part-time)} & \text{if } \delta_{2s} < h^*_{is} \leq \delta_{3s} \\ \vdots & \\ J \quad \text{(full-time)} & \text{if } \delta_{Js} < h^*_{is} \end{cases} \quad \text{for } s = \{t, t-1\} \tag{2.22}$$

where we choose the number of categories J=5 as our baseline specification.[15] Van Soest (1995) argues that mass points in the number of hours worked exist, because of work hour restrictions in contractual agreements. With J=5, we account for such bunching at full-time work (i.e. 40 hours per week), large part-time work (i.e. 32 hours per week), and small part-time work (i.e. 8-16 hours per week). We provide sensitivity analyses regarding the number of categories in section 2.6. z_{it} includes age dummies for a semi-parametric specification of age effects. Furthermore, we follow Blank (1990b), Ermisch and Wright (1993), Manning and Robinson (2004) and use information regarding marital status, children and other household char-

[15]Franses and Cramer (2010) show that there is no formal statistical testing method for the number of categories in an ordered regression model.

acteristics as exclusion restrictions (z_{it}).[16] By estimating separate models for each s, the age effects are allowed to differ across periods and cohorts. The model takes into account correlation between μ_{it} and μ_{it-1}, denoted in (2.16) to (2.19) by ρ_t. This is important because of the time-constant individual component c_i in $\mu_{it} = c_i + v_{it}$ (explained above).

In the second step we construct the correction terms (2.16) to (2.19) by using the estimates $\hat{a}_{it}$, $\hat{a}_{it-1}$, $\hat{b}_{it}$, $\hat{b}_{it-1}$, and $\hat{\rho}_t$. Next, $\hat{\lambda}_{1it}$, $\hat{\lambda}_{2it}$, $\hat{\lambda}_{3it}$ and $\hat{\lambda}_{4it}$ are used as additional regressors in the wage equation to obtain consistent estimates of β by OLS on the sample of first differences in wages that are observed in t and $t-1$. In order to avoid issues with discontinuous jumps in wages due to labor supply decisions, we select only those with $\Delta h_s = 0$ for the estimation of wages.[17] Similar to Dustmann and Schmidt (2000) we estimate separate wage equations for full-time and part-time work. Furthermore, following Kalwij and Alessie (2007), x includes a flexible semi-parametric specification of age-effects. To avoid the issue with age, period, and cohort effects (captured by the individual-specific effect), as these cannot be identified empirically because the calendar year is equal to the year of birth plus age thereby spanning up the vector space, we leave out period effects in the baseline specification of the wage equation. We leave out period effects as we argue that period effects are less important than age and cohort effects. In the robustness analysis in Section 2.5.2, we show how the results are affected when including period effects, parameterized as a linear time trend or as a function of the unemployment rate. Finally, we use block bootstrapped standard errors clustered at the individual level for inference in the two-stage approach as suggested by Wooldridge (2002).

[16] Our main conclusions are robust to using different exclusion restrictions.
[17] Our main conclusions are robust to allowing for $|\Delta h_s| \leq 1$.

Estimation results 2.5

First stage: labor force participation 2.5.1

The first-stage bivariate ordered probit model is estimated for every combination of t and $t-1$ for $t = \{2002, ..., 2014\}$ for men and women separately. We choose the number of labor supply categories $J = 5$ (as argued in subsection 2.4.2). For men, the bulk of the observations is in the full-time (62%) or non-working category (21%). If men are working part-time, they are often included in the highest part-time category (12%, part-time employment factor ≥ 0.75 and < 1.00). Women are more evenly spread over the different categories. 34% is in the non-working category, 11% in the smallest part-time category (part-time employment factor > 0 and < 0.50), 16% in the third category (≥ 0.50 and < 0.75), 18% in the largest part-time category (≥ 0.75 and < 1.00) and only 21% of women fall in the full-time category.[18]

In Table 2.5 in Appendix 2.C.3 we report the estimation results of the selection equation for the combination of 2001 and 2002 for men and women, respectively. Apart from the direction and significance, the reported coefficients have no direct interpretation and should be interpreted with respect to the estimated parameters $\delta_{j-1,t}$ and $\delta_{j-1,t-1}$ that indicate the thresholds between the J labor supply categories for time t and $t-1$, respectively.

Beginning with the exclusion restrictions, the results show that these variables have large predictive power for both men ($\chi^2 = 216$) and women ($\chi^2 = 1,942$).[19] This holds for both men and women, although we observe differences in which variables are important. For men, we find that only the average individual specific effects or 'contextual effects' predict the labor market participation.[20] Men without children and married men are more likely to work (full-time). For women, we observe that both within and contextual variation predict the labor force participation. For women,

[18]For a complete overview of the labor supply categories for all years see Appendix 2.C.1.

[19]The exclusion restrictions are predictive for all combinations of t and $t-1$.

[20]For an explanation of the decomposition into within, between and contextual effects see Bell et al. (2019).

having children, being married or widowed and having a partner past the early retirement age (ERA)[21] are associated with a lower labor force participation. The results show that the likelihood of participation, and especially full-time work, decreases with age. This is true for both men and women and all combinations of t and $t-1$.

Finally, the estimates suggest that the autoregressive nature of labor supply decisions ρ_t is important. Since ρ_t controls for unobserved heterogeneity in the first-stage in the approaches of Rochina-Barrachina (1999) and ours, a high and significant ρ_t can partially explain different results in the application of our approach and Dustmann and Schmidt (2000). Next to first-differences estimation and (non-)parametric assumptions about unobserved heterogeneity. For differences in results between our approach and that of Dustmann and Schmidt (2000), we refer to the robustness analysis in Section 2.5.2.

2.5.2 Second stage: wages

Main estimation results

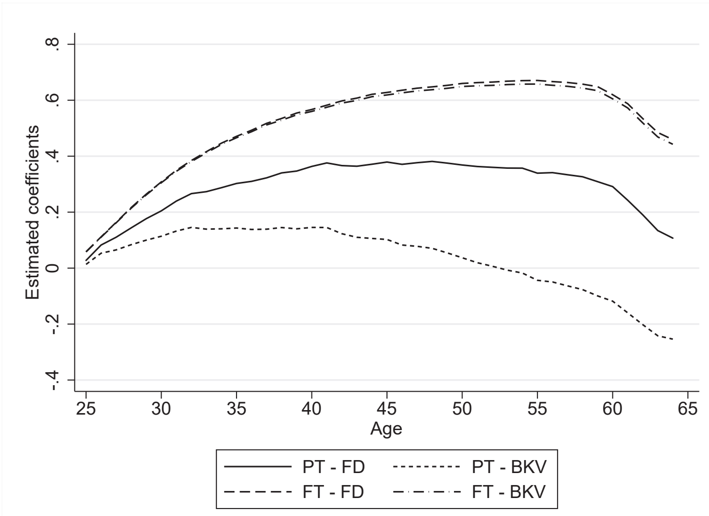
Figure 2.5 presents the age profile of the wages for men and women in part-time and full-time employment. We show the age coefficients without (FD) and with (BKV) the correction terms for selection, which are obtained in the first stage. The first-difference model takes the observed and time-invariant unobserved heterogeneity into account, while our model additionally controls for time-variant unobserved heterogeneity that is related to full-time and part-time work decisions. Taking into account the selection based on time-variant unobserved heterogeneity into part-time and full-time work changes the earnings estimates significantly.

The wage profiles using the FD estimator match the wage descriptives from the previous section well up to the final years prior to retirement. Wages grow over the life-cycle, with the largest increases at younger ages.

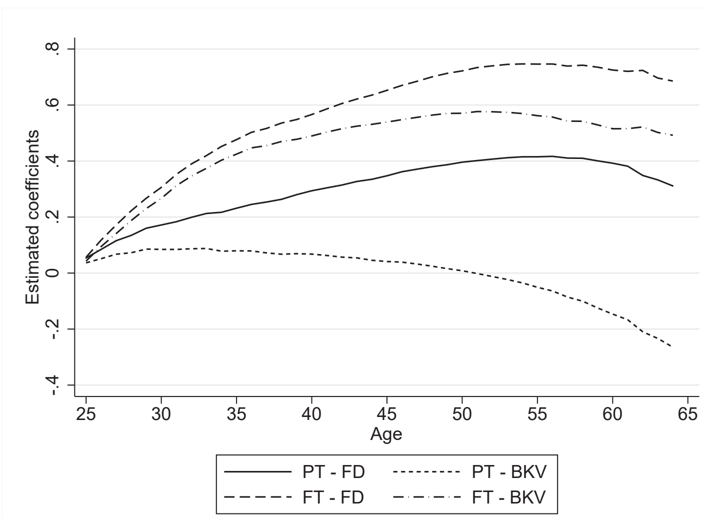
[21] We include a dummy for whether a person's spouse has reached the early retirement age (ERA), because prior empirical literature has shown that reaching the ERA affects own and spouses' labor supply decisions (Been et al. 2021, Stancanelli and Van Soest 2012). The ERA is 62 in many mandatory occupational pension schemes in the Netherlands.

Figure 2.5: Part-time (PT) and full-time (FT) regressions for men (a) and women (b) using first-differences (FD) and our model (BKV)

(a) Estimated coefficients men



(b) Estimated coefficients women



Thereafter, wages only start to decline in the final years prior to retirement. This phenomenon is not observed in the descriptive data and shows the importance of using an FD model. The maximum wage growth is larger for those working in full-time employment (more than 60%) than those in part-time employment (about 40%). These findings are similar for men and women.

Next, we move to the wage profiles obtained using the model that controls for selection into (part-time) work on both observed and unobserved heterogeneity. For men we find positive selection on unobserved individual characteristics in part-time work (p-value=0.0000), but no significant selection into full-time work (p-value=0.9397). This means that men with more affluent characteristics self-select into part-time employment. When men work full-time between the ages of 25 and 64 their estimated wage growth is 69% at the age of 55 (peak) and still about 49% at the age of 64. If, instead, they work part-time, their estimated wage growth is not significantly different from zero. Over the life-cycle, the wage growth of men working full-time is significantly higher than for those working part-time (F-test shows a p-value=0.0000).

For women, we find positive selection based on unobserved characteristics into both part-time work (p-value=0.0000) and full-time work (p-value=0.0000). After correcting for selection, the estimated wage growth is 59% at the age of 51 (peak) and still 51% at the age of 64 for women working full-time. Similar to the results for men, the wage growth when taking selection into account is not significantly different from zero for women working part-time. Over the life-cycle, the wage growth of women in full-time employment is significantly higher than for those in part-time employment (F-test shows a p-value of 0.0000). Overall, the results are comparable for men and women, with the exception that women not only positively select into part-time work, but also positively select into full-time work.

Robustness of the wage profiles

As discussed in section 2.4, when estimating the wage equation we have to make one additional assumption to deal with the collinearity problem of having age, period and cohort effects. Therefore, in our main analysis above we left out the period effects.[22] The robustness of these results are tested by re-estimating models with period effects parameterized as a linear time trend or as a function of the unemployment rate. Next, we test how our model compares to the method proposed by Dustmann and Schmidt (2000). This provides us with insight into the importance of estimation in first-differences with non-parametric assumptions on the unobserved heterogeneity in the wage equation and autoregressive nature of labor supply decisions.[23]

Figure 2.6 in Appendix 2.D shows the estimated age coefficients for the models with period effects for both men and women.[24] We begin with the model with linear period effects.[25] Albeit that age, period and cohort effects cannot be fully identified, this specification enables us to estimate the linear trend in period effects. Both for men and women, we find that the trend in year-to-year changes in wages is negligible, as the concerning coefficients are statistically insignificant. Accordingly, the main conclusions regarding the estimated life-cycle earnings and selection effects for both men and women remain the same. Next, we consider a parametric specification of the model where period effects are a function of the unemployment rate.[26] We find one percentage point increase in the unemployment rate compared to the previous period to be associated with at most a 0.6 percent change in the wages.[27] Again, the general

[22] In Appendix 2.A, we show clear trends in participation, the incidence of part-time work and wages. These trends are probably correlated with age and, therefore, in the case of a model without period effects (partly) absorbed by the age effects.

[23] Dustmann and Schmidt (2000) also use an ordered selection rule in the first-stage, but make parametric assumptions on the unobserved heterogeneity in both the first- and second stage. In comparison, the model proposed in this paper makes no parametric assumptions on the unobserved heterogeneity in the second stage and exploits the autoregressive nature of labor supply decisions similar to Rochina-Barrachina (1999).

[24] To allow for more flexibility, we allow the period effects to differ for those working in part-time and full-time employment in both models.

[25] Since we are estimating first-difference models, this means that we assume the wages to have a constant growth rate over time.

[26] To be more precise, we include the differenced unemployment rate as we are estimating first-difference models in the second stage.

[27] For men in full-time employment, we find one percentage point increase in the unemployment rate compared to the previous period to be associated with a significant 0.6 percent decrease in the wages and no significant association for men in part-time employment. For women, we find one percentage point increase in the unemployment

conclusions regarding the direction and significance of selection remain the same.

Figure 2.7 in Appendix 2.D presents the age estimates for men and women in part-time and full-time employment using the Dustmann and Schmidt (2000) method.[28] From the figure, three general observations stand out. First – as opposed to the previous results – we observe the increases in wages to be largely comparable for men in part-time and full-time employment, even with the correction terms of Dustmann and Schmidt (2000) included. As a result, the life-cycle difference between full-time and part-time wages is negligible. Second, albeit the age estimates of those in both part-time and full-time employment without correction terms show the typical inverted U-shape, the wage profiles of men change drastically when including the correction terms. The wage profiles of men in both part-time and full-time employment continue to go up after the age of 40, indicating substantial negative selection at older ages. Instead of the substantial decreases in wages in the years prior to retirement, the results using the Dustmann and Schmidt (2000) model suggest that wages of men in both part-time and full-time work continue to grow up till retirement. Third, the results for women obtained using the Dustmann and Schmidt (2000) model are comparable to those of the model proposed in this paper, although the magnitude of the selection on unobserved characteristics is smaller. The inclusion of the correction terms using the Dustmann and Schmidt (2000) method shows positive selection on unobserved characteristics into both full-time (p-value=0.0018) as well as part-time employment (p-value=0.0000). The different selection effects, especially for men, show the importance of the estimation in first-differences with non-parametric assumptions on the unobserved heterogeneity in the wage equation and autoregressive nature of labor supply decisions.

rate compared to the year before to be associated with a 0.3 percent decrease (increase) in wages for those in full-time (part-time) employment.

[28]The Dustmann and Schmidt (2000) model is in levels and not in first-differences as it does not exploit the autoregressive nature of participation. Because of this, the age profiles without the correction terms also differ from those estimated using first-differences. However, investigation of this model is still a useful exercise as our main interest lies in how the inclusion of correction terms and the selection effects of the unobserved heterogeneity are affected by the parametric assumptions in the wage equation.

Table 2.1: Selection and number of labor supply categories

	Part-time			Full-time		
J	Selection	χ^2	P-value	Selection	χ^2	P-value
	Men					
RB	Negative	51.4	0.0000	Negative	44.1	0.0000
3	Positive	31.4	0.0000	–	0.3	0.8696
4	Positive	27.9	0.0000	–	0.1	0.9686
5	Positive	30.9	0.0000	–	0.1	0.9397
6	Positive	20.0	0.0000	–	0.1	0.9436
7	Positive	13.0	0.0112	–	0.3	0.8766
8	Positive	24.5	0.0001	–	0.3	0.8645
	Women					
RB	–	0.2	0.9198	Negative	15.1	0.0005
3	Negative	24.9	0.0001	Positive	61.3	0.0000
4	Negative	63.5	0.0000	Positive	66.3	0.0000
5	Positive	93.1	0.0000	Positive	80.1	0.0000
6	Positive	88.7	0.0000	Positive	81.7	0.0000
7	Positive	69.5	0.0000	Positive	77.8	0.0000
8	Positive	125.7	0.0000	Positive	96.8	0.0000

Binary versus ordered selection 2.6

In this section, we investigate the importance of taking the selection in the intensive margin of labor supply into account, as compared to a binary selection rule (i.e. as proposed by Rochina-Barrachina (1999)). As argued in section 2.4, the choice of the number of labor supply categories in our model (J) is arbitrary to some extent and is a trade-off between more categories versus more observations per category. Hence, to get an idea of how important the choice for J is for conclusions regarding selection effects, we present Table 2.1 in which we show the direction and significance of the selection terms for different choices of J. We restrict our analysis to $2 \leq J \leq 8$ to make sure we have a sufficient number of observations per category. In theory, $J > 8$ should be possible as long as there is a sufficient number of observations per category. Recall, in our main analysis we use J=5, allowing for three different part-time employment categories.

We find two interesting patterns regarding selection and choices for J in Table 2.1. Firstly, for $J > 2$ (ordered selection), our proposed method produces different conclusions regarding the existence and direction of selection than for $J = 2$ (binary selection). Hence, including unobserved information regarding the intensive labor supply decision is important compared to information on selection in the extensive margin of labor supply. The results with $J = 2$ suggest negative selection among both part-time and full-time employed men, whereas we find positive or no selection effects among these groups for $J > 2$, respectively. For women, the results of $J = 2$ show no selection effects for part-time employed women whereas we find evidence in favor of selection for $J > 2$, albeit the direction of the selection bias depends on J. For women working full-time, we find negative selection for $J = 2$ and positive selection for $J > 2$.

Secondly, we find that conclusions regarding selection are consistent across $J > 2$ among men, but not among women. For men, we find that adding information beyond $J = 3$ does not change the results for both part-time and full-time employed men. Among full-time employed women, conclusions regarding selection are consistent across $J > 2$. For part-time employed women, however, a less consistent picture arises when analyzing selection for $J > 2$. For $3 \leq J \leq 4$, we find negative selection. For $5 \leq J \leq 8$, we find positive selection. This switching of the direction of selection from $J = 4$ to $J = 5$ is most likely a consequence of the increased unobserved information allowed for by a larger J. Logically, this tends to be especially important among part-time employed women since there are relatively many women working part-time, both in relatively small and large part-time jobs (see Table 2.3 in the appendix). In contrast, part-time working men can often be found in relatively large part-time jobs which makes the additional information from $J > 3$ less important than for women.

Given the analyses in Table 2.1, we conclude that allowing for part-time employment is important for conclusions regarding selection, but choosing the number of categories $J > 2$ is of less importance as results are largely consistent. However, applied researchers should be aware that the additional information from a larger J is most likely important for the analysis of women in part-time employment.

2.7 Conclusion

To estimate correct earnings profiles over the life-cycle, we argue that non-random selection into full-time and part-time work contains relevant information on unobserved heterogeneity. Therefore, we propose a new panel data sample selection model that conditions on selection into both full-time and part-time work. We build on the method of Rochina-Barrachina (1999) and extend her method by allowing for an ordered instead of binary selection rule which allows us to differentiate between full-time and part-time work. In this way, we extend the method by Rochina-Barrachina (1999) in a similar way as Dustmann and Schmidt (2000) extended Wooldridge (1995). The main advantage of Rochina-Barrachina (1999) over Wooldridge (1995) is that no parametric assumptions about the unobserved heterogeneity in wages and the decision to work are needed.

Using administrative data from the Netherlands, where part-time work is highly prevalent, we show that taking into account non-random selection into (part-time) work changes the earnings estimates significantly. For men, we find no selection into full-time work which suggests that selecting full-time working prime age males in models of earnings dynamics does not lead to biased estimates. Hence, using full-time working men without selection correction, as in Lagakos et al. (2018) for example, is justified by our results (though we particularly focus on Dutch men, who are not considered by Lagakos et al. (2018)). However, results are unlikely to be representative for other groups among which part-time working men for whom we find positive selection in part-time work. This implies that men with relatively affluent characteristics choose part-time work and that part-time wages are overestimated if such selection is not taken into account. For women, we find positive selection into both part-time and full-time work. Moreover, we show with our new panel data estimator that it is important to distinguish between part-time and full-time employment, as taking into account labor supply decisions at the extensive margin only – like in Rochina-Barrachina (1999) – leads to different conclusions with respect to the existence and direction of selection. Hence, we conclude that part-time employment entails additional information on unobserved characteristics that are import in the estimation of wage profiles.

Applying our method to estimate life-cycle earnings profiles, we show that correcting for selection also results in different shapes of the earnings profiles compared to regular first-differences estimates. With our proposed method, we find that earnings in full-time employment peak later in the life-cycle than earnings in part-time employment. This is true for both men and women. Additionally, these differences are amplified when correcting for selection into full-time and part-time employment.

Our study has important implications for both academics and policy. For academics, our proposed method is useful for several applications, such as 1) the estimation of part-time wage penalties and 2) testing for the existence of selection among full-time working prime age men who are generally selected in earnings models.[29] Additionally, our model is also useful in other contexts where the selection decision is ordered, e.g. the number of children or subjective health outcomes. For policy, applying our method to administrative earnings data from the Netherlands, we show that part-time work has large effects on life-time earnings and, hence, on the accumulation of savings, pensions, and wealth.

[29]Among others, Baker (1997), Baker and Solon (2003), Daly et al. (2022), Gottschalk and Moffitt (1994), Guvenen (2009), Heathcote et al. (2010), Lagakos et al. (2018), Lillard and Weiss (1979), Lillard and Willis (1978), Meghir and Pistaferri (2004, 2010), Moffitt and Gottschalk (2012), Pischke (1995), Storesletten et al. (2004).

Wage descriptives over time 2.A

Columns 3 to 10 in Table 2.2 present full-time and part-time wage rates, and (part-time) participation rates for men and women, respectively. As expected, participation rates are higher for men than for women. However, both declining participation rates for men and increasing participation rates for women make the difference in participation rates between men and women smaller over time, from 20%-points in 2001 to 5%-points in 2014. For both men and women part-time employment (conditional on participation) has increased over time, with the most substantial growth among men. Despite this, men still had much lower part-time employment rates (27%) than women (71%) in 2014.

Next we look at full-time and part-time wages, where we increased the part-time wage using the part-time employment factor to match the full-time wages. From the wage statistics, four general observations stand out. First, wages are on average higher for men than for women. This holds for both full-time and part-time wages in all sample years. Second, full-time wages are on average higher than part-time wages. Similarly to the previous observation, this holds for both men and women in all sample years. Third, the gender wage gap (column 2) has declined between 2001 and 2014.[30] In turn, this is the result of the faster increase in part-time employment of men compared to women, declining (part-time) wages for men and increasing (full-time) wages for women.

[30] Again, we observe a discontinuity around 2006. When we focus on the two separate time period, i.e. 2001-2005 and 2006-2014, the cumulative decline is even more pronounced.

Table 2.2: Trends in participation and wages

Year	Men				Women			
	Average FT wage	Average PT wage	Part-time[a] (%)	Partici-pation (%)	Average FT wage	Average PT wage	Part-time[a] (%)	Partici-pation (%)
2001	46,536	43,765	14	80	38,850	36,923	64	60
2002	46,646	43,915	15	80	39,375	37,764	64	61
2003	47,255	44,793	15	79	39,922	37,731	65	62
2004	47,549	44,715	16	78	40,596	38,212	66	62
2005	47,397	46,033	17	78	40,680	38,409	67	62
2006	48,092	43,570	26	78	40,968	35,502	68	64
2007	47,610	42,763	26	79	41,102	35,972	68	66
2008	48,318	42,252	28	79	41,587	35,861	67	67
2009	48,419	42,848	27	78	42,205	36,629	68	67
2010	48,636	41,721	24	76	42,297	37,072	68	67
2011	48,120	41,857	26	77	42,663	36,619	69	69
2012	47,596	41,278	26	77	42,250	36,433	69	70
2013	47,282	39,716	26	76	42,095	36,048	70	70
2014	47,805	40,147	27	75	43,024	36,171	71	70

[a] For persons who actually work.

2.B Derivation of correction terms

Following the method of the two-step approach proposed by Heckman (1976, 1979), we work out (2.10) to obtain correction terms, that can be added as additional regressors to the main equation (the wage equation). Rochina-Barrachina (1999) also extends Heckman's sample selection technique to the case where one correlated selection rule in two different time periods generates the sample. In addition, we allow an ordered selection rule instead of a binary selection indicator.

Equation 2.10 contains the first moment of a doubly truncated trivariate normal distribution (where $(u_{it} - u_{it-1})$ is not truncated[31] and $\frac{\mu_{it-1}}{\sigma_{\mu t-1}}$ and $\frac{\mu_{it}}{\sigma_{\mu t}}$ are doubly truncated). For the sake of convenience, in the remainder of this Appendix we denote $w_1 = u_{it} - u_{it-1}$, $w_2 = \frac{\mu_{it-1}}{\sigma_{\mu t-1}}$ and $w_3 = \frac{\mu_{it}}{\sigma_{\mu t}}$. Following Manjunath and Wilhelm (2012), the trivariate truncated normal density is defined as

$$\phi_{\alpha\Sigma}(w_1, w_2, w_3) = \begin{cases} \frac{\phi_\Sigma(w_1,w_2,w_3)}{\alpha} & \text{for } a_{it-1} \leq w_2 < b_{it-1} \text{ and } a_{it} \leq w_3 < b_{it} \\ 0 & \text{otherwise} \end{cases}$$

[31] Boundaries of $-infty$ and $infty$

where $a_{it-1}, a_{it}, b_{it-1}$ and b_{it} are defined in (2.11) to (2.14). α denotes the fraction after truncation ($= P(a_{it-1} \leq w_2 < b_{it-1}$ and $a_{it} \leq w_3 < b_{it})$), and $\phi_{\boldsymbol{\Sigma}}$ the normal density with expectations of zero and covariance matrix Σ.

To calculate the first moment of w_1, we use the moment generating function (*m.g.f*) of the doubly truncated trivariate normal distribution. We take the derivative with respect to t_1 and evaluate the function in $\boldsymbol{t} = 0$. The moment generating function is defined as the threefold integral of the form

$$m(\boldsymbol{t}) = E(e^{\boldsymbol{t}'\boldsymbol{w}}) \tag{2.23}$$

$$= \frac{1}{\alpha(2\pi)^{3/2}|\Sigma|^{1/2}} \int_{\mathbf{a}}^{\mathbf{b}} exp\left(-\frac{1}{2}\boldsymbol{w}'\boldsymbol{\Sigma}^{-1}\boldsymbol{w} - 2\boldsymbol{t}'\boldsymbol{w}\right) d\boldsymbol{w} \tag{2.24}$$

For the derivation of the first derivative of the m.g.f. with regard to t_1 we refer to (7)–(10) in Manjunath and Wilhelm (2012):

$$\frac{\partial m(\boldsymbol{t})}{\partial t_1} = e^{\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}} \frac{\partial \Phi_{\alpha\boldsymbol{\Sigma}}}{\partial t_1} + \Phi_{\alpha\boldsymbol{\Sigma}} \frac{\partial e^{\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}}}{\partial t_1} \tag{2.25}$$

where

$$\Phi_{\alpha\boldsymbol{\Sigma}} = \frac{1}{\alpha(2\pi)^{3/2}|\Sigma|^{1/2}} \int_{\boldsymbol{a}-\boldsymbol{\Sigma}t}^{\boldsymbol{b}-\boldsymbol{\Sigma}t} \exp\left(-\frac{1}{2}\boldsymbol{w}'\boldsymbol{\Sigma}^{-1}\boldsymbol{w}\right) d\boldsymbol{w}. \tag{2.26}$$

In (2.26) $\boldsymbol{a} = (-\infty, a_{it-1}, a_{it})$ and $\boldsymbol{b} = (\infty, b_{it-1}, b_{it})$. In (2.25) the last term can be simplified as

$$\frac{\partial e^{\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}}}{\partial t_1} = e^{\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}} \left(t_1\sigma_1^2 + t_2\sigma_{12} + t_3\sigma_{13}\right) \tag{2.27}$$

Furthermore, the last part of the first term of (2.25) can be rewritten as

$$\frac{\partial \Phi_{\alpha\boldsymbol{\Sigma}}}{\partial t_1} = \frac{\partial}{\partial t_1} \int\limits_{\boldsymbol{a}-\boldsymbol{\Sigma}t}^{\boldsymbol{b}-\boldsymbol{\Sigma}t} \phi_{\alpha\boldsymbol{\Sigma}}(\boldsymbol{w})d\boldsymbol{w} \tag{2.28}$$

After applying the Leibniz's rule for differentiation under the integral sign and rewriting the equation this becomes

$$\begin{aligned}
&\frac{\partial \Phi_{\alpha\boldsymbol{\Sigma}}}{\partial t_1} = \\
&-\sigma_1^2 \int\limits_{a_2^*}^{b_2^*} \int\limits_{a_3^*}^{b_3^*} \phi_{\alpha\boldsymbol{\Sigma}}(b_1^*, w_2, w_3) dw_3 dw_2 + \sigma_1^2 \int\limits_{a_2^*}^{b_2^*} \int\limits_{a_3^*}^{b_3^*} \phi_{\alpha\boldsymbol{\Sigma}}(a_1^*, w_2, w_3) dw_3 dw_2 \\
&-\sigma_{12} \int\limits_{a_1^*}^{b_1^*} \int\limits_{a_3^*}^{b_3^*} \phi_{\alpha\boldsymbol{\Sigma}}(w_1, b_2^*, w_3) dw_3 dw_1 + \sigma_{12} \int\limits_{a_1^*}^{b_1^*} \int\limits_{a_3^*}^{b_3^*} \phi_{\alpha\boldsymbol{\Sigma}}(w_1, a_2^*, w_3) dw_3 dw_1 \\
&-\sigma_{13} \int\limits_{a_1^*}^{b_1^*} \int\limits_{a_2^*}^{b_2^*} \phi_{\alpha\boldsymbol{\Sigma}}(w_1, w_2, b_3^*) dw_2 dw_1 + \sigma_{13} \int\limits_{a_1^*}^{b_1^*} \int\limits_{a_2^*}^{b_2^*} \phi_{\alpha\boldsymbol{\Sigma}}(w_1, w_2, a_3^*) dw_2 dw_1
\end{aligned} \tag{2.29}$$

where $[a_1^* \quad a_2^* \quad a_3^*]' = \boldsymbol{a}^* = \boldsymbol{a} - \boldsymbol{\Sigma t}$ and $[b_1^* \quad b_2^* \quad b_3^*]' = \boldsymbol{b}^* = \boldsymbol{b} - \boldsymbol{\Sigma t}$. Taking the terms together and evaluating the derivative $\frac{\partial m(\boldsymbol{t})}{\partial t_1}$ in $\boldsymbol{t} = 0$ gives us the first moment of w_1

$$\begin{aligned}
&\mathrm{E}(w_1 | a_{it-1} \leq w_2 < b_{it-1} \text{ and } a_{it} \leq w_3 < b_{it}) = \\
&\quad -\sigma_{12}\frac{\phi(b_{it-1})}{\alpha}\left[\Phi\left(\frac{b_{it}-\rho b_{it-1}}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{a_{it}-\rho b_{it-1}}{\sqrt{1-\rho^2}}\right)\right] \\
&\quad +\sigma_{12}\frac{\phi(a_{it-1})}{\alpha}\left[\Phi\left(\frac{b_{it}-\rho a_{it-1}}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{a_{it}-\rho a_{it-1}}{\sqrt{1-\rho^2}}\right)\right] \\
&\quad -\sigma_{13}\frac{\phi(b_{it})}{\alpha}\left[\Phi\left(\frac{b_{it-1}-\rho b_{it}}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{a_{it-1}-\rho b_{it}}{\sqrt{1-\rho^2}}\right)\right] \\
&\quad +\sigma_{13}\frac{\phi(a_{it})}{\alpha}\left[\Phi\left(\frac{b_{it-1}-\rho a_{it}}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{a_{it-1}-\rho a_{it}}{\sqrt{1-\rho^2}}\right)\right]
\end{aligned} \tag{2.30}$$

where ρ is the correlation coefficient of w_2 and w_3, and $\alpha = \Phi_2(b_{it-1}, b_{it}, \rho) - \Phi_2(a_{it-1}, a_{it}, \rho)$.

Estimation of the selection equation 2.C

Labor supply categories 2.C.1

Here we describe the distribution of workers over the five labor supply categories ($J = 5$) for men and women, respectively. For men, the bulk of the observations is in the full-time (62%) or the non-working category (21%). Only 2% and 3% of the men fall in the two smallest part-time categories (part-time employment factor > 0 and < 0.50, and ≥ 0.50 and < 0.75, respectively) and 12% in the highest part-time category (part-time employment factor ≥ 0.75 and < 1.00). The share of men in the full-time category is declining over time from 70 percent in 2001 to 56 percent in 2014. The categories that consequently show the largest increases are the non-working and the largest part-time work categories.

Women are more evenly spread over the different categories. 34% is in the non-working category, 11% in the smallest part-time category (part-time employment factor > 0 and < 0.50), 16% in the third category (part-time employment factor ≥ 0.50 and < 0.75), 18% in the largest part-time category (part-time employment factor ≥ 0.75 and < 1.00). Only 21% of women work full-time and fall in the final category. As opposed to men, the share of women in the full-time employment category is relatively stable over time. The largest changes for women are observed in the non-working and the larger part-time work categories. The share of women that is non-working has decreased from 40 percent in 2001 to 30 percent in 2014, which resulted in more women in the two largest part-time categories.

Table 2.3: Distribution of men and women over the 5 labor supply categories over time

Year	Men: Non-working	Men: 0 < fte <0.5	Men: 0.5 ≤ fte <0.75	Men: 0.75 ≤ fte <1	Men: Full-time	Women: Non-working	Women: 0 < fte <0.5	Women: 0.5 ≤ fte <0.75	Women: 0.75 ≤ fte <1	Women: Full-time
2001	0.185	0.021	0.026	0.070	0.699	0.397	0.110	0.131	0.143	0.218
2002	0.189	0.021	0.026	0.071	0.693	0.381	0.113	0.139	0.147	0.220
2003	0.196	0.024	0.025	0.074	0.681	0.378	0.112	0.143	0.147	0.220
2004	0.205	0.023	0.029	0.076	0.668	0.377	0.105	0.153	0.153	0.212
2005	0.211	0.027	0.030	0.076	0.656	0.372	0.112	0.148	0.158	0.210
2006	0.206	0.022	0.030	0.151	0.591	0.347	0.110	0.154	0.179	0.210
2007	0.198	0.021	0.029	0.156	0.596	0.335	0.103	0.160	0.192	0.210
2008	0.197	0.023	0.028	0.174	0.578	0.321	0.104	0.159	0.195	0.221
2009	0.209	0.022	0.029	0.159	0.581	0.316	0.105	0.162	0.198	0.219
2010	0.220	0.022	0.030	0.136	0.592	0.315	0.102	0.168	0.192	0.222
2011	0.218	0.023	0.033	0.150	0.576	0.307	0.102	0.171	0.209	0.212
2012	0.219	0.025	0.031	0.150	0.575	0.298	0.101	0.177	0.209	0.215
2013	0.229	0.024	0.035	0.139	0.572	0.295	0.101	0.184	0.211	0.210
2014	0.232	0.022	0.033	0.152	0.562	0.294	0.098	0.181	0.220	0.208
Total	0.208	0.023	0.030	0.123	0.616	0.338	0.106	0.159	0.182	0.215

2.C.2 Transitions in labor supply categories

Table 2.4 describes the year-to-year transitions in labor supply categories for $J = 5$. The diagonal of the transition matrix represents individuals who remained in the same labor supply category from time $t-1$ to t (i.e. $\Delta h_s = 0$). Both men and women exhibit strong persistence in certain categories. Specifically, the probability of staying in non-employment ($h_t = 1$) is approximately 0.98. Similarly, the probability of staying in full-time employment ($h_t = 5$) is very high. Among men, the persistence in full-time work is particularly strong at 0.91, while among women it is also notable at 0.84. Persistence in the part-time categories ($h_t = 2, 3, 4$) is lower compared to non-employment and full-time employment but still substantial, especially among women.

As elaborated in more detail in Section 2.3, the administrative records provide comprehensive information regarding (labor) income and the part-time factor but do not include details about the distribution of working hours throughout the calendar year. This paper compares part-time and full-time wages. Transitions between different labor supply categories (i.e., $\Delta h_s \geq 1$), however, are likely driven by changes in the extensive margin rather than the intensive margin (e.g. people becoming unemployed or starting a job during the calendar year). Given this and the small absolute and relative numbers, we exclude them from the main analysis.

Table 2.4: Year-to-year transitions (fractions) in labor supply categories of men and women

Men	t					
$t-1$	Non-working	$0 <$ fte <0.5	$0.5 \leq$ fte <0.75	$0.75 \leq$ fte <1	Full-time	N
Non-working	0.979	0.003	0.003	0.005	0.010	49,507
$0 <$ fte <0.5	0.064	0.580	0.147	0.110	0.010	7,761
$0.5 \leq$ fte <0.75	0.039	0.086	0.478	0.233	0.164	7,086
$0.75 \leq$ fte <1	0.023	0.013	0.046	0.593	0.325	32,393
Full-time	0.013	0.003	0.007	0.067	0.911	173,203
Women	t					
$t-1$	Non-working	$0 <$ fte <0.5	$0.5 \leq$ fte <0.75	$0.75 \leq$ fte <1	Full-time	N
Non-working	0.989	0.005	0.003	0.002	0.002	87,297
$0 <$ fte <0.5	0.027	0.775	0.152	0.033	0.014	26,318
$0.5 \leq$ fte <0.75	0.015	0.079	0.761	0.118	0.028	43,104
$0.75 \leq$ fte <1	0.010	0.013	0.102	0.737	0.139	49,167
Full-time	0.010	0.005	0.022	0.125	0.839	59,420

2.C.3 First-stage regression results

Table 2.5: Estimation results selection equation for men and women

	men				women			
	t=2002		t-1=2001		t=2002		t-1=2001	
	Coef.	S.e.	Coef.	S.e.	Coef.	S.e.	Coef.	S.e.
Age 25	–		-0.37**	0.17	–		-0.05	0.16
Age 26	-0.44***	0.17	-0.35**	0.15	-0.15	0.16	-0.10	0.11
Age 27	-0.41***	0.15	-0.52***	0.15	-0.13	0.12	-0.37***	0.14
Age 28	-0.58***	0.15	-0.61***	0.15	-0.40***	0.14	-0.39***	0.13
Age 29	-0.64***	0.15	-0.44***	0.15	-0.54***	0.13	-0.45***	0.14
Age 30	-0.56***	0.16	-0.66***	0.15	-0.54***	0.15	-0.63***	0.14
Age 31	-0.79***	0.15	-0.63***	0.15	-0.70***	0.15	-0.67***	0.14
Age 32	-0.75***	0.16	-0.98***	0.16	-0.80***	0.15	-0.91***	0.15
Age 33	-1.08***	0.16	-1.03***	0.16	-1.03***	0.16	-0.96***	0.15
Age 34	-1.13***	0.17	-1.11***	0.17	-1.10***	0.16	-1.25***	0.16
Age 35	-1.28***	0.17	-1.27***	0.17	-1.33***	0.17	-1.33***	0.17
Age 36	-1.37***	0.17	-1.52***	0.17	-1.47***	0.17	-1.49***	0.18
Age 37	-1.60***	0.18	-1.61***	0.18	-1.59***	0.18	-1.48***	0.18
Age 38	-1.73***	0.18	-1.55***	0.19	-1.61***	0.19	-1.54***	0.19
Age 39	-1.68***	0.19	-1.70***	0.19	-1.70***	0.20	-1.55***	0.19
Age 40	-1.87***	0.19	-1.93***	0.20	-1.71***	0.20	-1.71***	0.20
Age 41	-2.03***	0.20	-2.00***	0.20	-1.86***	0.21	-1.73***	0.20
Age 42	-2.19***	0.20	-2.10***	0.20	-1.86***	0.21	-1.58***	0.21
Age 43	-2.27***	0.20	-2.22***	0.21	-1.73***	0.22	-1.63***	0.22
Age 44	-2.34***	0.21	-2.21***	0.21	-1.79***	0.22	-1.75***	0.22
Age 45	-2.32***	0.22	-2.50***	0.21	-1.94***	0.23	-1.78***	0.23
Age 46	-2.59***	0.22	-2.48***	0.22	-1.99***	0.23	-1.74***	0.23
Age 47	-2.62***	0.22	-2.58***	0.23	-2.01***	0.24	-1.84***	0.24
Age 48	-2.72***	0.23	-2.66***	0.23	-1.99***	0.25	-1.80***	0.24
Age 49	-2.86***	0.23	-2.85***	0.23	-2.06***	0.25	-1.77***	0.24
Age 50	-2.99***	0.24	-2.76***	0.24	-1.97***	0.25	-1.70***	0.25
Age 51	-2.96***	0.24	-2.80***	0.24	-1.95***	0.26	-1.88***	0.25
Age 52	-2.98***	0.25	-2.80***	0.25	-2.08***	0.26	-1.81***	0.26
Age 53	-2.98***	0.25	-2.96***	0.25	-1.97***	0.26	-1.93***	0.26
Age 54	-3.07***	0.25	-3.13***	0.25	-2.15***	0.27	-1.98***	0.26
Age 55	-3.25***	0.25	-3.16***	0.26	-2.17***	0.27	-1.94***	0.27
Age 56	-3.36***	0.26	-3.28***	0.26	-2.16***	0.27	-2.00***	0.27
Age 57	-3.43***	0.26	-3.36***	0.26	-2.29***	0.28	-2.19***	0.28
Age 58	-3.59***	0.26	-3.67***	0.28	-2.46***	0.29	-2.28***	0.29
Age 59	-3.89***	0.28	-3.78***	0.28	-2.55***	0.29	-2.26***	0.30
Age 60	-4.06***	0.28	-4.22***	0.29	-2.50***	0.31	-2.27***	0.32
Age 61	-4.60***	0.29	-4.39***	0.30	-2.69***	0.33	-2.71***	0.33
Age 62	-4.83***	0.30	-4.83***	0.31	-3.28***	0.34	-3.27***	0.38
Age 63	-5.09***	0.32	-5.02***	0.36	-3.52***	0.39	-3.65***	0.47
Age 64	-4.98***	0.36			-3.74***	0.48		
Children	0.02	0.02	0.01	0.02	-0.27***	0.02	-0.25***	0.02
Single	–		–		–		–	
Married	0.00	0.04	0.02	0.03	-0.30***	0.04	-0.29***	0.04
Divorced	-0.03	0.06	-0.05	0.06	-0.15***	0.06	-0.16***	0.05
Widowed	-0.04	0.16	-0.06	0.16	-0.39***	0.09	-0.27***	0.09
Partner ERA	0.10*	0.06	0.05	0.06	0.12***	0.04	0.03	0.04
Children (average)	-0.15***	0.04	-0.14***	0.04	-0.40***	0.04	-0.39***	0.04
Single (average)	–		–		–		–	

Continued on next page

Table 2.5 – continued from previous page

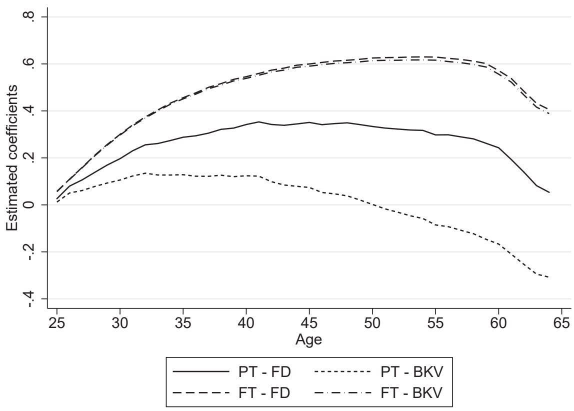
	Men				Women			
	t=2002		t-1=2001		t=2002		t-1=2001	
	Coef.	S.e.	Coef.	S.e.	Coef.	S.e.	Coef.	S.e.
Married (average)	0.36***	0.05	0.38***	0.04	-0.34***	0.05	-0.33***	0.05
Divorced (average)	0.02	0.02	0.06	0.07	-0.05	0.07	-0.05	0.06
Widowed (average)	0.13	0.16	0.23	0.17	-0.66***	0.11	-0.76***	0.10
Partner ERA (average)	-0.18**	0.09	0.18**	0.09	-0.60***	0.07	-0.52***	0.06
χ^2-stat $\bar{z}_i$	2,011***				2,847***			
χ^2-stat $\bar{z}_i$ excl. age dummies	103***				390***			
χ^2-stat exclusion restrictions	216***				1,942***			
δ_{1s}	3.35**	1.43	3.09**	1.37	1.68	1.48	0.83	1.51
δ_{2s}	3.43**	1.43	3.18**	1.37	2.07	1.48	1.23	1.51
δ_{3s}	3.53**	1.43	3.29**	1.37	2.51*	1.48	1.65	1.51
δ_{4s}	3.74***	1.43	3.50**	1.37	2.99**	1.48	2.12	1.51
ρ_t	0.97***	0.00			0.95***	0.00		
Obs.	20,985				19,510			
χ^2	2,037				4,458			

Note: $\bar{z}_i$ includes individual time averages of all age dummies, marital status dummies, children dummy and the variable indicating whether having a partner past the early retirement age (ERA). The different parameters for δ_{Js} indicate the thresholds between the $J = 5$ labor supply categories. ρ_t indicates the correlation between the error terms at time t and $t - 1$. Standard errors are robust and clustered at the individual level. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

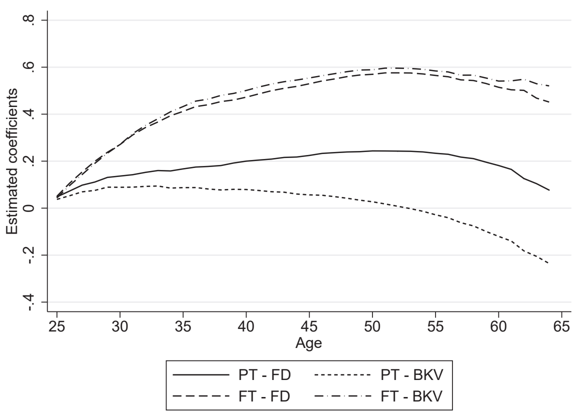
2.D Robustness of the wage equation

Figure 2.6: Part-time (PT) and full-time (FT) regressions using first-differences (FD) and our model (BKV) controlling for linear period effects and unemployment rates

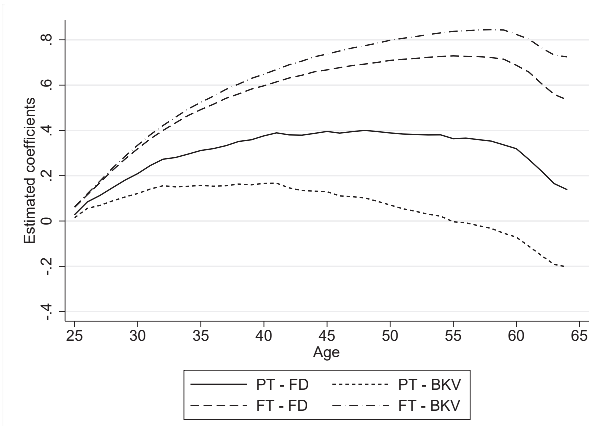
(a) Linear period effects, men



(b) Linear period effects, women



(c) Unemployment rates, men



(d) Unemployment rates, women

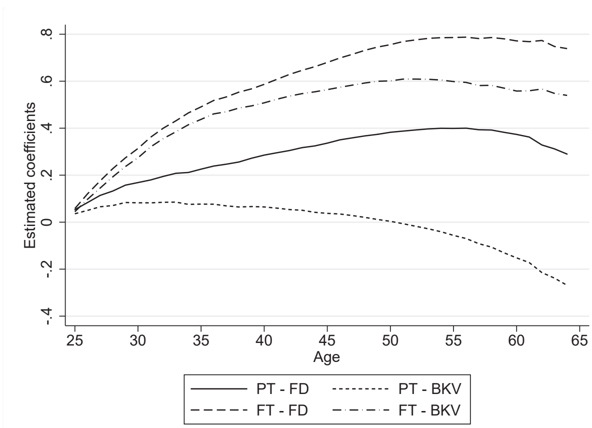
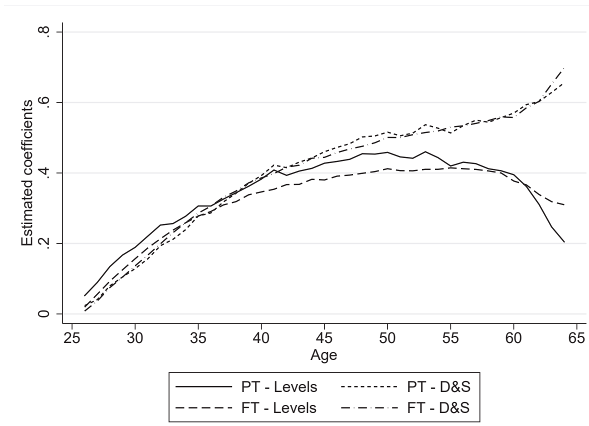
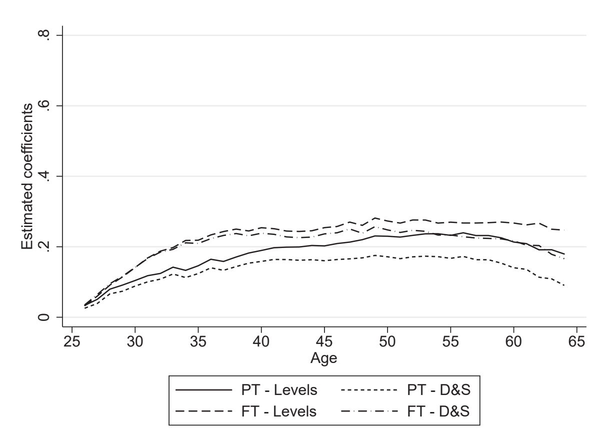


Figure 2.7: Part-time (PT) and full-time (FT) regressions for men (a) and women (b) using levels and Dustmann and Schmidt (2000) approach

(a) Estimated coefficients men



(b) Estimated coefficients women



3 Decomposing employment trends of disabled workers

Abstract

This paper estimates Age-Period-Cohort models on employment rates of Dutch Disability Insurance (DI) applicants. We find that the substantial decrease in employment between 1999 and 2013 is explained by year-of-application cohort effects and that period effects are negligible. In turn, application cohort effects partly stem from increasing shares of applicants without permanent contracts. Changes in application cohort effects are largely confined to the years following two DI reforms that increased self-screening among workers. We next analyze changes in employment rates of awarded and rejected applicants and follow a Difference-in-Differences approach. Assuming common compositional cohort effects, we infer negligible effects of changes in benefit conditions.

3.1 Introduction

Over the last decades, many OECD countries have shown declining employment rates of disabled individuals (OECD 2010). For the US, there is

A journal version of this paper is published in *The B.E. Journal of Economic Analysis & Policy* as Koning and Vethaak (2021). This chapter is co-authored by Pierre Koning. The authors are grateful to Patrick Hullegie for his indispensable work on the data. They also thank Nicole Maestas, Jan-Maarten van Sonsbeek, Eric French, Astrid Grasdal, Lawrence Katz, Marike Knoef, Carla van Deursen, Barend Barentsen and seminar participants at the Dutch Economists Day of 2019, the University of Bergen, the University of Nantes, the KVS New Paper Sessions of 2019 and the EALE/SOLE/AASLE World Meeting of 2020 for useful comments to presentations and earlier drafts of the paper. This research is sponsored by Instituut Gak.

strong evidence that the Social Security Disability Insurance (SSDI) program has become a more attractive scheme for low-skilled workers (Autor and Duggan 2003, Bound et al. 2003, 2014, Maestas 2019, Von Wachter et al. 2011). Since the mid-eighties, the expansion of the SSDI program coincided with higher fractions of applicants with weak labor market positions for whom the receipt of benefits has discouraged them from working. This means that overall decline in employment among SSDI recipients can be attributed both to changes in the composition of applicants – with vulnerable labor market positions – and a lack of work incentives for those awarded benefits. This raises two fundamental questions that are inherent with the design of Disability Insurance (DI) schemes. The first is on the targeting of benefits: who should be eligible to DI? In this respect, DI benefits can either be restricted to workers with severe disabilities or expanded to vulnerable workers with mild impairments that cannot engage in substantial gainful activities. The second question concerns the design of work incentives: how can benefit recipients be encouraged to exploit their remaining earnings capacity? For the overall assessment of DI reforms, it thus is of key importance to both address targeting effects – i.e., compositional changes – and incentive effects.

This paper provides such a broad assessment of the employment trends of DI applicants in the Netherlands, a country that also experienced strong decreases in the labor force attachment of claimants. In the Netherlands, drastic reforms have been implemented to curb the inflow into DI as well as increase work incentives for disabled workers. On the one hand, increases in screening stringency and eligibility thresholds changed the composition of new applicant cohorts. This increased the severity of new claims and decreased their employment rates (De Jong et al. 2011, Godard et al. 2022).[1] On the other hand, the new disability that started in 2006 increased work incentives for new DI recipients with residual earnings capacities. Koning and van Sonsbeek (2017) show that this increased the individual employment rates of awarded applicants. Taken together, the

[1]Contributions of Campolieti (2006) for Canada, Deshpande and Li (2019) for the US, Markussen et al. (2018) for Norway and Liebert (2019) for Switzerland suggest that increased scrutiny and increased application costs have the potential to substantially lower DI inflow rates.

reforms have both changed the targeting efficiency as well as the work incentives of the DI scheme.

To incorporate both selection and incentive effects in the assessment of employment trends of disabled workers, this paper is the first to estimate Age-Period-Cohort (APC) models on administrative applicant data. We focus on DI applicant cohorts between 1999 and 2013 which are followed up to 2016. In the context of our model, 'age' corresponds to the elapsed duration since application, period effects capture business cycle and other calendar time effects, and cohort effects resemble changes in employment rates that are specific to annual DI application cohorts. Using a Deaton-Paxson (DP) specification, we first disentangle application cohort effects from period and age effects. These application cohort effects represent the joint effect of: (i) compositional changes induced by secular cohort-specific time trends in the demand for and health conditions of the insured population of workers with disabilities; (ii) compositional changes induced by disability reforms that affected self-screening before application; and (iii) individual changes in the employment rate of awarded applicants – or: 'incentive effects' – induced by cohort-specific changes in benefit conditions. With reforms in the Netherlands that affected new applicant cohorts only, both changes in the targeting and in incentive effects are embodied in year-of-application cohort effects.

Our second aim is to provide a further decomposition of the application cohort effects into changes stemming from compositional changes ('targeting') and changes in the individual's employment probability stemming from DI reforms ('incentive effects'). In the spirit of Bound (1989), we follow a Difference-in-Differences (DiD) approach with awarded and rejected DI applicants as treatment and control groups.[2] Assuming that compositional effects – both induced by reforms and gradual changes in the labor market – affected treatment and control cohort groups equally, the DiD estimates of the reforms indicate changes in the individual em-

[2]To estimate the discouraging impact of SSDI benefits, Bound (1989), Chen and Van der Klaauw (2008), Von Wachter et al. (2011), Maestas et al. (2013) and French and Song (2014) compare accepted and denied SSDI applicants. Following the seminal article by Bound (1989), the resulting estimates form an upper bound of the employment rates of awarded applicants, since rejected applicants are considered to have more labor market attachment than accepted applicants.

ployment probability of awarded applicants. Stated differently, we assume that changes in employment probabilities caused by compositional differences across application cohorts are equal in magnitude for awarded and rejected applicants and are captured by an additive application cohort effect common to both awarded and rejected applicants. The residual change in employment rates for awarded applicants can then be characterized as the 'incentive' effects of the reforms on awarded applicants.

Our main research findings can be summarized as follows. First, application cohort effects of DI applicants are the main contributor to their observed decline in employment, amounting to about 30 percentage points in total. Contrasting to this, the effect of calendar time effects is negligible, suggesting that both business cycle effects or secular time trends that affected all application cohorts equally were not important. Second, changes in application cohort effects are largely in tandem with the disability reforms of 2003 and 2006; it is only for the years after the 2006 reform that we observe a gradual and substantial further decline in cohort effects. Third, our DiD-analysis provides limited evidence for employment rates to respond to changes in the work incentives of awarded applicants. This implies that the substantial changes in application cohort effects are almost entirely driven by compositional changes of applicants. Again, this highlights the importance of self-screening among potential applicants as a driver of the observed changes in employment rates. Finally, a substantial part of application cohort effects is explained by changes in demographic variables and the initial labor market position of applicants. As far as we can infer from the inclusion of observed controls, there is a general worsening in the labor market position of application cohorts that are more likely to have flexible contracts. This finding resembles e.g. Autor and Duggan (2003), Von Wachter et al. (2011) and Maestas et al. (2013) who argue there is a declining demand for low-skilled workers with health conditions in the US.

The remainder of the paper is organized as follows. Section 3.2 describes the Dutch DI system, together with the relevant reforms and their expected effects. Section 3.3 provides a description of the selected data and section 3.4 contains the estimation strategy. Section 3.5 presents the results of the analysis before section 3.6 concludes.

Institutional background 3.2

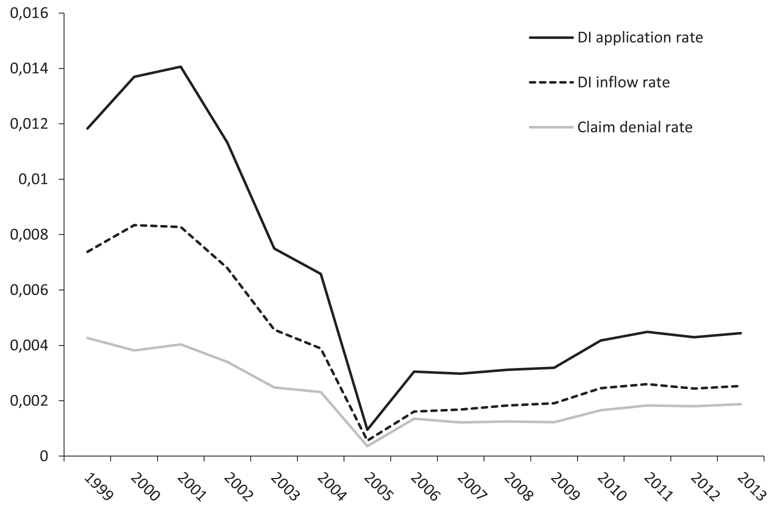
This section describes the main characteristics of the Dutch DI system and the two major disability reforms since 1999: the Gatekeeper Protocol (in Dutch: *Wet verbetering Poortwachter*) and the WIA (in Dutch: *Wet Werk en Inkomen naar Arbeidsvermogen*). From now on, we refer to these reforms as the GKP and WIA, respectively. When explaining the effects of the reforms, a particular interest lies in the distinction between compositional effects and incentive effects. We define compositional effects as changes in the average employment rates that result from changes in the composition of new cohorts of DI applicants. For DI applicants, these changes stem from changes in self-screening and work resumption in the waiting period before the DI decision. Incentive effects are defined as changes in individual employment rates as a response to changes in the work incentives for awarded DI applicants, measured after the DI award decision.

DI in the Netherlands 3.2.1

The Dutch DI program covers income losses resulting from both occupational and non-occupational injuries of all employed workers. Sick-listed workers apply for DI benefits at the end of the waiting period of absence. The employer is obliged to continue full wage payments in this period. The waiting period was extended from one to two years in 2004.

After application, the National Social Insurance Institute (NSII) determines the degree of disability of workers. To this end, medical examiners assess the limitations of applicants and vocational experts subsequently select occupations with corresponding wages to determine the residual potential earning capacity. The degree of disability then equals the lost potential earning capacity as a fraction of pre-disability earnings. Until 2006, the applicant was awarded DI benefits if the degree of disability exceeded the threshold of 15%. This threshold was increased to 35% as part of the WIA reform in 2006. Workers with a degree of disability between 35 and 80% are awarded partial DI benefits and those with losses of more than 80% receive full benefits. Partially disabled receive 70 percent

Figure 3.1: Annual DI application rate, inflow rate and claim denial rate of total insured working population, 1999-2013



Source: Statistics Netherlands

of their loss of earnings capacity and fully disabled receive 70 percent of their pre-disability earnings.

With its broad coverage, generous benefits and limited self-screening, the Dutch DI system laid the ground for a continuous increase in DI enrollment. Around the turn of the century, DI enrollment peaked at about 12% of the insured working population (Koning and Lindeboom 2015). Figure 3.1 shows that annual DI application rates then ranged between 1.2 and 1.4% of the working population. The first substantial drop in both DI application and awards occurred in 2003, at the start of the GKP reform. Using a discontinuity-in-time regression, Godard et al. (2022) find that the effect amounted to a 40 percent reduction in the DI applicant rate. The second major decrease in DI application and award rates is observed since 2005. While this drop initially demarcates the mechanical effect of the extension of the sickness period to two years in 2005, the new disability law (WIA) led to persistently lower DI inflow rates.[3] In what follows, we discuss both the GKP reform and the WIA reform in more detail.

Stricter screening: the GKP reform (2003) 3.2.2

The GKP reform has affected the screening process for new DI application cohorts since 2003.[4] The GKP stipulates the responsibilities of both the worker and the employer for sickness spells lasting at least six weeks. This means the responsibility of reintegrating sick workers during the waiting period was removed from the NSII, which since then acts as a gatekeeper at the moment of DI claim. Figure 3.2 provides an overview of the steps of the application process towards entering DI under the GKP.[5] After six weeks of absence, the worker and the employer must draft a rehabilitation plan together which is based on an assessment of cause of disability, functional limitations and the likelihood of work resumption. The rehabilitation plan should be approved by a caseworker of the NSII in the eighth week of absence, after which it is binding for both parties. The worker can apply for DI benefits if work resumption is not established before the end of the waiting period and when all requirements of the GKP have been met. If not, the wage continuation period may be extended with one year at maximum.

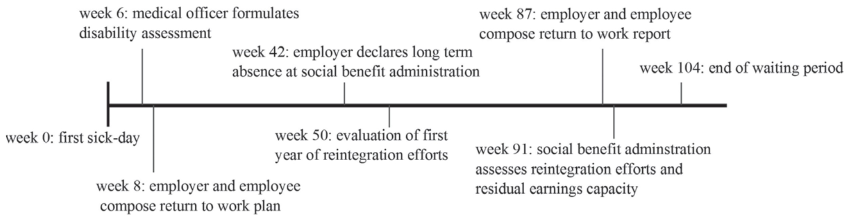
There is strong evidence that the GKP changed the composition of DI applicants. The increased rehabilitation efforts did not only increase the likelihood of work resumption in the absence period that precedes DI claims for workers with better employment prospects, but also induced self-screening among those workers with less severe health conditions

[3]Albeit that the reform extended the waiting period to two years, some workers still applied for DI in 2005 due to (administrative) delays. For instance, waiting periods could be extended due to a lack of integration efforts.

[4]In 2003, the start of the GKP went together with the abolishment of DI experience rating for smaller firms – see De Groot and Koning (2016). As this reform affected only a small share of the working force in the Netherlands, the overall effects on DI applications are small.

[5]Note that the figure is relevant under the (current) disability scheme with an absence period of two years. In the year the GKP came into force, the waiting period was one year.

Figure 3.2: GKP conditions in the sickness waiting period



(De Jong et al. 2011, Godard et al. 2022).[6] Both these mechanisms have resulted in a sample of DI applicants that are probably more deserving, with worse health conditions.[7]

3.2.3 The new disability law: WIA (2006)

The main goal of the WIA reform of 2006 was to stimulate workers with less-severe impairments to exploit their residual earnings capacity. The idea was that three policy changes would contribute to this: (i) increased self-screening through an extension of the waiting period from one to two years; (ii) stricter eligibility, as the threshold for DI receipt was increased to 35%; and (iii) differentiated benefits for severely disabled and applicants with sufficient remaining earnings capacity.

First, the extension of the waiting period from one to two years implied another increase in the costs of wage continuation and all other costs inherent with the GKP.[8] Following similar arguments as for the introduction

[6]De Jong et al. (2011) evaluate a large-scale experiment in the Netherlands to study the effects of increased screening. They find that this induces employers to increase reintegration activities, which in turn increases work resumption rates during sickness absenteeism. They argue that those higher rates are induced by self-screening among the potential applicants.

[7]Koning and Lindeboom (2015) argue that the increased application costs of the GKP may also have had adverse effects on the individual employment rates of disabled workers. The increased responsibilities and the risk of extension of wage sanctions – i.e., the increase of the wage continuation period – may have discouraged employers to hire workers with disabilities (see also Hullegie and Koning 2018).

[8]Godard et al. (2022) argue that GKP costs vary between 0.23 and 0.43 of the total wage sum of firms in the Netherlands.

of the GKP, one would expect this extension to increase work resumption and self-screening in the waiting period before DI application.

As a second part of the WIA, the threshold of the degree of disability for eligibility was increased from 15 to 35 percent of pre-disability earnings. Van Sonsbeek and Gradus (2012) argue that this implied a drop in DI inflow rates of roughly 20 percentage points. With a substantially lower share of beneficiaries with partial benefits, it is expected that the average employment rate among the total group of beneficiaries has declined. This compositional effect may have been strengthened by increased self-screening among (potential) applicants with mild health conditions.

Third, the WIA differentiates between fully and permanently disabled workers (IVA) and partially and/or temporary disabled workers (WGA) for which strong financial incentives were introduced. Workers in the WGA scheme receive 70 percent of their lost earnings during the first period of benefit receipt ('wage-related related benefits'). Depending on the work history, this period lasts 38 months at maximum. Next, WGA beneficiaries continue receiving the same benefit level if and only if they exploit at least 50 percent of their earnings capacity; if not, the benefit is based on the statutory minimum wage. Benefits for partially disabled workers thus function as a wage subsidy that incentivizes them to work.[9] Koning and van Sonsbeek (2017) find that the incentive change for partially disabled workers increases the employment incidence with 2.6 percentage points.[10] Still, the overall effect of the increase in incentives is probably smaller than this, as wage subsidies are targeted at partially disabled workers – constituting about one quarter of the total DI inflow – and are relevant in the second period of benefit receipt only (Koning and Lindeboom 2015).[11]

[9]For a detailed explanation of the functioning and consequences of the wage subsidy, we refer to Koning and van Sonsbeek (2017).

[10]Kantarci et al. (2023) find somewhat smaller employment effects, comparing sick-listed worker cohorts that fell under the old and new disability scheme, respectively. In their study, the effect estimate of work incentives can be interpreted as an upper bound, as it also captures the effect of the waiting period extension from one to two years.

[11]The wage subsidy may have induced perverse work incentives for fully and temporary disabled workers in the WGA scheme, as switches to the partial scheme inhibit the risk of sizable declines in benefits (Koning and Lindeboom 2015).

Overall, the GKP and the WIA reform most likely affected the composition of the pool of new DI applicant cohorts. Increases in self-screening and increases in work resumption in the absence period probably have resulted in a smaller sample of DI applicants with more severe health conditions and lower employment rates. Since the reforms affected new cohorts, we expect that these effects are mirrored by discrete jumps in cohort effects, rather than gradual changes stemming from secular labor market and health trends. For applicants who were awarded benefits, the WIA reform also changed the incentive to work. Accordingly, positive changes in relative cohort employment effects of awarded and rejected applicants may be indicative of incentive effects. Taken together, our interest thus lies in employment changes stemming from both selection and incentive effects of the reforms.

3.3 Data

3.3.1 Data sources

We use individual-level data on all DI applications between 1999 and 2013 from the administrative records of the NSII. Application cohorts from these years are followed between 1999 and 2016, containing information on the award decision and date, the diagnosed impairment and the assessed degree of disability.[12] Medical diagnoses are grouped by impairment type (mental, musculoskeletal, respiratory, endocrine, cardiovascular, nervous system and other impairments).[13] The degree of disability is given by intervals ($<$15%, 15-34%, 35-44%, 45-54%, 55-64%, 65-79%, $\geq$80%).

We merge the application data with administrative data of Statistics Netherlands of the full Dutch population between 1999 and 2016. This yields individual-year data covering a sufficiently long period to assess

[12] After 2007 we observe a shift from rejections due to insufficient degree of disability to rejections for 'unknown' reasons (see Figure 3.10). This probably reflects administrative changes, as the medical assessment was unchanged and rejection rates remained more or less constant. Our analysis therefore does not differentiate between different reasons for rejection.

[13] The distribution of impairment groups by application cohorts is shown in Figure 3.11.

the long-term effects of both the GKP and WIA reform. The Census Register contains information on the personal characteristics, such as gender, month of birth and death, and nationality. The tax records provide information (in 2015 Euros) on annual gross earnings and receipt of unemployment, disability, and social assistance benefits. We define an individual as employed in a specific year when he or she received positive wage earnings. For employed individuals we also observe the contract type (permanent or temporary) and sector of employment (70 in total).

In total, we observe 1,183,186 individual applications between 1999 and 2013. For our empirical analysis, we exclude reapplications, workers that are younger than 18, older than 65 or deceased at the time of application and workers for which the year of application or award decision was unknown.[14] This reduces our sample to 962,356 observations. Attrition from this longitudinal sample stems from the occurrence of deaths and migration.

Descriptive statistics 3.3.2

Table 3.1 presents descriptive statistics of employment and earnings of rejected and awarded applicant cohorts before and after the DI decision. We separate the total sample of applicants in three sub-samples: (i) application cohorts unaffected by the reforms, 1999-2002; (ii) application cohorts covered by the GKP but not by the WIA, 2003-2005; and (iii) application cohorts subject both to the GKP and the WIA, 2006-2013.[15] The table shows that rejected and awarded DI applicants with full benefits have similar pre-disability employment rates two years before the DI assessment. Inherent with the eligibility conditions for DI, these rates are close to 100%. Applicants awarded partial benefits have higher pre-disability earnings and have more often a permanent contract than those rejected and those awarded full benefits. As expected, awarded applicants experience drops

[14]The vast majority of the omitted applications is excluded due to the year of application being unknown (74,761) or the observation of multiple records for the same individual (134,015). In the latter case, we only selected the first application.

[15]The GKP affected sick-listed workers as from 2002. Hence, DI applicants of 2002 are not affected. Likewise, the extension of the waiting period from one to two years affected workers that became sick from 2004 onwards.

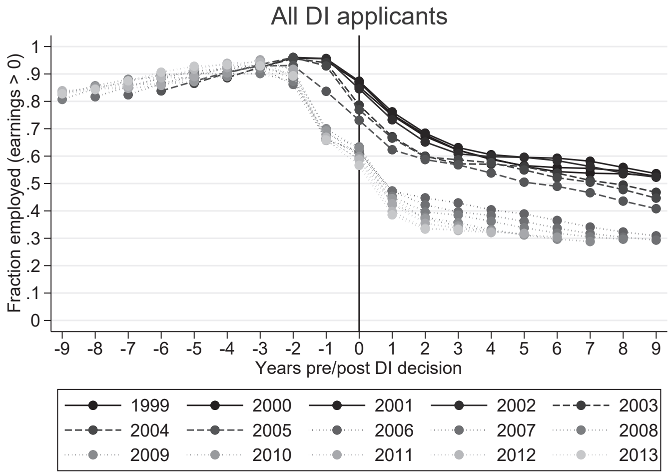
in income from earnings that are more sizable than for rejected applicants. Awarded applicants tend to be more often male, older and show higher mortality rates than rejected applicants. Over the years, we also observe substantial changes in the employment rates and the composition of DI applicants. Most notably, in the last time frame (2006-2013) applicant cohorts show markedly lower employment rates two years after application. This drop is most sizable for applicants awarded full DI benefits.

Table 3.1: Employment, earnings and demographic characteristics for rejected, and partially and fully awarded DI applicant cohorts

Application cohort	1999-2002			2003-2005			2006-2013		
Application status	Rejected applicants	Awarded partial benefits	Awarded full benefits	Rejected applicants	Awarded partial benefits	Awarded full benefits	Rejected applicants	Awarded partial benefits	Awarded full benefits
Labor supply and earnings 2 years before application									
Percent positive covered earnings	91.1	95.6	90.2	95.2	97.2	94.5	90.0	91.2	89.3
Average annual earnings (€1,000)	25,477	34,404	26,901	25,220	35,304	27,071	23,104	34,086	26,278
Median positive annual earnings (€1,000)	23,000	33,000	24,000	24,000	34,000	25,000	21,000	33,000	24,000
Percent permanent contract	—	—	—	72.9	86.4	79.3	71.8	78.4	76.8
Labor supply and earnings in year of application									
Percent positive covered earnings	88.2	91.4	76.8	81.6	86.6	68.7	64.9	70.5	57.8
Average annual earnings (€1,000)	24,149	32,353	22,957	23,441	32,105	22,479	20,648	26,953	18,709
Median positive annual earnings (€1,000)	22,000	32,000	20,000	22,000	31,000	19,000	18,000	25,000	15,000
Percent permanent contract	—	—	—	66.8	78.8	59.6	53.1	62.5	52.5
Labor supply and earnings 2 years after application									
Percent positive covered earnings	74.5	78.6	41.8	69.2	74.1	38.5	58.0	59.5	18.4
Average annual earnings (€1,000)	25,301	29,006	19,467	23,916	29,051	19,893	21,475	23,556	17,408
Median positive annual earnings (€1,000)	24,000	28,000	16,000	22,000	28,000	17,000	19,000	21,000	13,000
Percent permanent contract	67.3	74.9	38.1	60.3	69.4	33.3	47.5	54.1	16.0
Demographics									
Average age at application	40	44	43	41	45	42	43	46	46
Age at application < 40	49.0	32.3	39.9	46.7	30.6	40.4	39.1	27.5	29.3
Age at application 40 - 50	27.5	32.0	27.2	28.9	31.9	27.7	30.3	29.2	26.0
Age at application $50 \leq$	23.5	35.8	32.9	24.5	37.5	31.9	30.6	43.3	44.7
Percent male	37.3	51.6	40.8	41.4	53.4	46.6	46.1	54.3	48.8
Percent Dutch	77.1	83.0	76.7	73.9	81.2	74.4	70.3	77.3	72.6
Percent death after application									
Percent deceased 2 years after application	1.7	1.0	4.5	1.2	1.3	5.9	1.0	1.2	6.2
Percent deceased 4 years after application	2.3	1.9	6.1	2.2	2.4	7.8	1.8	2.4	8.9
Observations	126,323	115,639	158,558	55,909	41,496	52,019	145,677	44,709	142,835

Note: The contract type is only observed from 2001 onward.

Figure 3.3: Annual fraction employed DI applicants before and after the award decision, stratified by application year (1999-2013)



To shed light on longitudinal patterns, Figure 3.3 depicts the evolution of employment rates of applicant cohorts before and after the award decision. Figure 3.4 shows a similar graph for separate samples of rejected, partially awarded and fully awarded applicants, with separate panels for the three regimes as in Table 3.1. From the figures, four general observations stand out. First, employment rates generally increase up to two years before the award decision and decline thereafter. While the initial increase follows from the eligibility conditions inherent to the Dutch DI system, the subsequent decline follows from the start of the absence period that precedes the award decision. Second, we observe large jumps in employment rates in the years the two reforms were implemented, but employment rates are roughly constant *within* the time periods of 1999-2002 and 2003-2005. This suggests that changes in employment rates until 2006 can largely be linked to the GKP and WIA reform. Third, we observe changes in the employment patterns of new application cohorts after 2006, the year the new disability law came into force. Since then, a large share of the decline in employment is already observed in the absence period, two years before the disability decision. Finally, since 2006 the employment rates of successive cohorts gradually decreased with virtually constant between-cohort employment differences. Following the eyeball test suggested by Voas and Chaves (2016), it is unlikely that elapsed duration effects and period effects cancel each other out in such a way that there are constant employment differentials. The driving factor behind the declining employment rates are therefore most plausibly the presence of cohort effects and not combined period and elapsed duration effects.

Comparing rejected and awarded applicants 3.3.3

Following Bound (1989), we proxy the discouraging impact of DI benefits by the difference in employment rates of rejected and awarded applicants. Since the severity of health impairments is likely stronger among accepted applicants, the difference in employment rates – the 'Bound estimate' – probably provides an upper bound of the discouraging impact. Figure 3.5 presents annual changes in the Bound estimate for the Netherlands for annual application cohorts, measured three years after the award decision.[16] Panel A shows the annual Bound estimates that follows from comparing rejected and *all* awarded applicants. Rejected applicants show a gradual decline in the employment rates three years after application, contrasting to the change in employment rates for awarded applicants shows a dramatic decline in 2006, when the WIA came into force. After 2006, the Bound estimate is about 30 percentage points, which is in the ballpark of estimates obtained for SSDI benefits.[17]

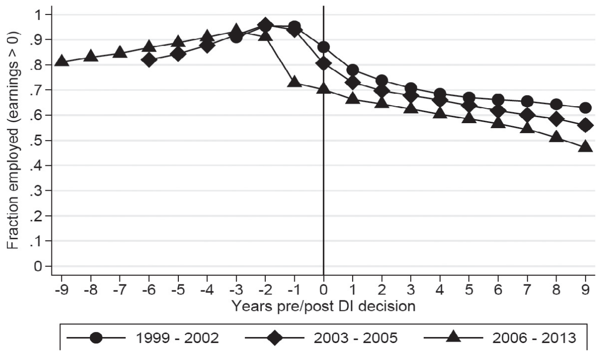
To reduce the supposedly positive bias stemming from differences in the severity of impairments, we next limit the sample of awarded applicants to those with partial DI benefits. Panel B of Figure 3.5 shows that these two groups have very similar downward employment patterns. The corresponding Bound estimate becomes small and even negative, ranging between -2 and -5 percentage points. This negative sign originates

[16]This gives a sufficiently long time delay to assess long-term employment rates of these cohorts.

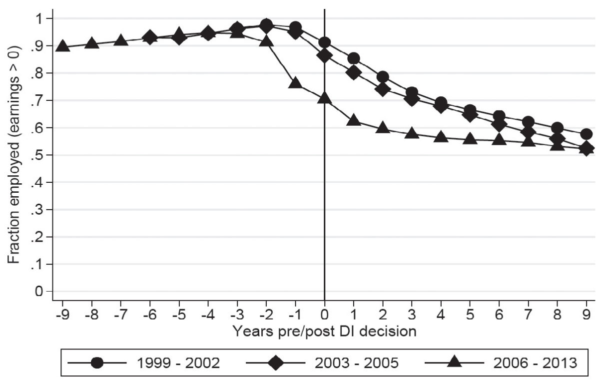
[17]Bound (1989) finds a difference in employment rates one year after application of between 26 and 30 percentage points for applicants aged 45-64. Von Wachter et al. (2011) shows that the Bound estimate amounts to more than 35 percentage points for applicants aged 30-44. Bound et al. (2003) estimates a difference three years after application of 20 percentage point. These results are similar to Chen and Van der Klaauw (2008) who show a reduction of the labor force participation of 15-18 percentage points.

Figure 3.4: Annual average employment rates of rejected, partially and fully awarded DI applicant cohorts for three time periods, before and after the award decision.

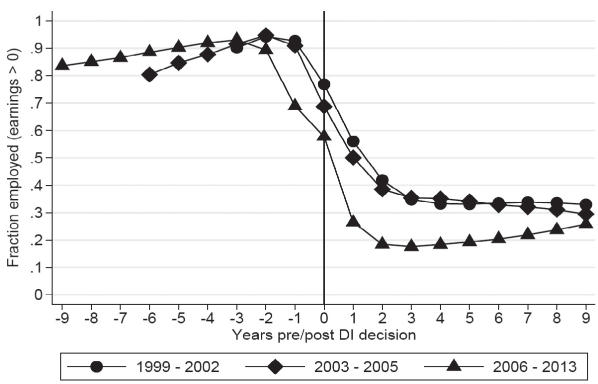
Panel A. Annual fraction employed of rejected DI applicants



Panel B. Annual fraction employed of applicants awarded partial DI benefits



Panel C. Annual fraction employed of applicants awarded full DI benefits



from the fact that applicants with higher pre-disability earnings are more likely to have a strong labor force attachment and experience a higher percentage drop in earning capacity. Similar arguments are put forward by Maestas et al. (2013), who show that rejected SSDI applicants typically have lower pre-employment rates. Finally, Panel C shows the employment rates for application cohort samples that are classified by degree of disability (below 35% or between 35 and 80%) and not by benefit outcome. Until 2006, employment rates of both groups are virtually equal to each other. Thereafter, the patterns are the same as in Panel B.

Empirical strategy 3.4

Specification 3.4.1

The aim of this paper is to decompose the mechanisms underlying the substantial decline in the employment rates of DI applicants. To this end, we propose a two-step analysis with Age-Period-Cohort (APC) models. First, we decompose employment trends into changes in the effect of the elapsed duration since application (the 'age' effect), time period effects and application cohort effects. Second, we further decompose application cohort effects into compositional and incentive effects, using a Difference-in-Differences (DiD) approach with distinct effects for awarded and rejected applicants.

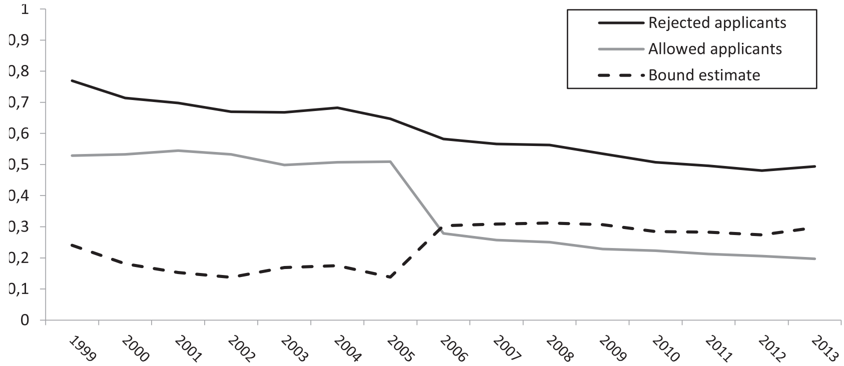
We specify the APC model for the incidence of employment E for all DI applicants in our sample, measured for post-application years. E is equal to one while working and zero otherwise.

$$E_{it,\tau} = \alpha_{t-\tau} + \pi_t + \gamma_\tau + \epsilon_{it}, \tag{3.1}$$

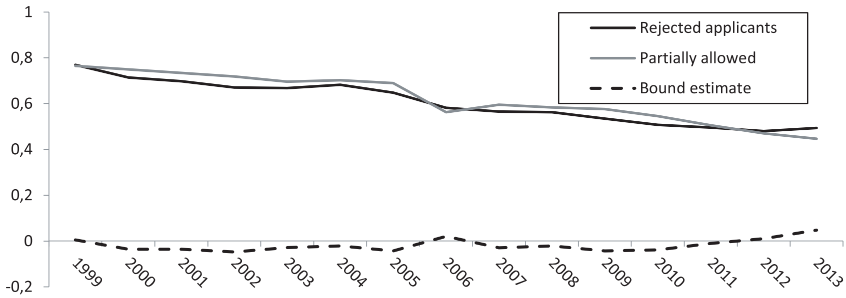
with $t \geq \tau$. In Equation (3.1), the employment status E of individual i (i = 1,..,N) in year t (t = 1,..,T) with a DI decision in year τ (τ = 1,..,$\mathcal{T}$) is determined by the number of years after application (i.e., the 'age' effect), a calendar year ('period') effect and an application cohort effect. Note that we have T = 18 years (1999-2016) and $\mathcal{T}$ = 15 application cohorts

Figure 3.5: Annual employment rates and Bound estimates for different application cohort samples between 1999 and 2013, measured three years after the DI decision

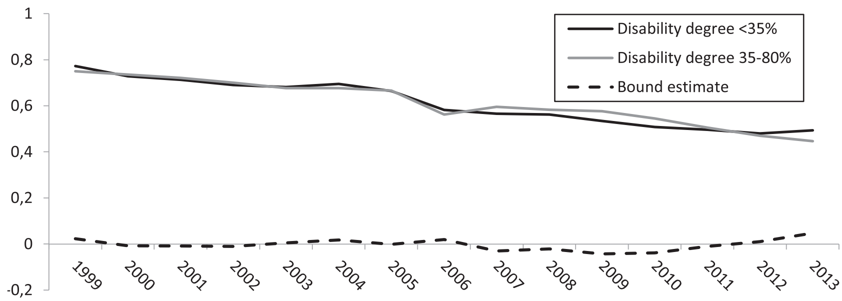
Panel A. Employment rates of rejected applicants and awarded applicants



Panel B. Employment rates of rejected and partially awarded applicants



Panel C. Employment rates of applicants with degree of disability of <35% and 35-80%



(1999-2013) in our sample. Age, period and cohort effects are denoted by the vectors α, π and γ, respectively. Regarding cohort effects, it is important to stress that reforms affected new application cohorts only. Discontinuous changes in application cohort effects may therefore indicate reform effects. Without age as a control variable, the 'age' effect equals the effect of aging and the elapsed duration since application.[18] Finally, ϵ is an error term.

In Equation (3.1), application cohort effects estimates represent both compositional changes among applicant cohorts and incentive changes among the sub-sample of awarded applicants. To further disentangle these two effects, we therefore extend the APC model with distinct age, period and cohort effects for awarded and rejected DI applicants. Specifically, we define $A_{i,\tau}$ as a dummy that is equal to one if DI applicant i in the application year cohort τ is awarded benefits, and zero otherwise. This yields the following specification:

$$\begin{aligned} E_{it,\tau} = & (1 - A_{i,\tau}) \left\{ {\alpha^0}_{t-\tau} + \pi_t^0 \right\} + A_{i,\tau} \left\{ {\alpha^1}_{t-\tau} + \pi_t^1 \right\} \\ & + \gamma_\tau + (1 - A_{i,\tau})\ \tilde{\gamma}_\tau + \epsilon_{it}, \end{aligned} \tag{3.2}$$

with α^0 and π^0 denoting age and period effects for the rejected applicants, respectively; α^1 and π^1 denoting age and period effects for the awarded applicants, respectively; and $\tilde{\gamma}_\tau$ as the application cohort effect that is interacted with the award indicator.[19] $\tilde{\gamma}_\tau$ can be interpreted as the Bound estimate for a specific application cohort τ. This estimate controls for the fact that age and period effects may differ between awarded and rejected applicants. Increases in the Bound estimate ($\tilde{\gamma}$) indicate equal decreases in incentive effects; this follows from the fact that the Bound estimate takes awarded applicants as a reference group. Since there is a specific Bound estimate for each cohort year, we refer to γ_τ as the unrestricted Bound estimates.

[18] We will also estimate model specifications that control for the age of applicants.

[19] Similar to Equation 3.1, note that we impose orthogonality restrictions on α^0 and α^1 to estimate all parameters of Equation (3.2).

To provide a more structured view on the effect of the reforms in our sample, we next impose the following restrictions on $\tilde{\gamma}$:

$$\begin{aligned}\tilde{\gamma}_{\tau} =& \tilde{\gamma}_0 + I(\tau \geq 2003)\ \tilde{\gamma}_{gkp} \\ & + I(2006 \leq \tau \leq 2009)\ \tilde{\gamma}_{wia,st} + I(\tau \geq 2010)\ \tilde{\gamma}_{wia,lt}\end{aligned} \tag{3.3}$$

with $\tilde{\gamma}_{gkp}$, $\tilde{\gamma}_{wia,st}$ and $\tilde{\gamma}_{wia,lt}$ denoting the effect of the GKP reform and the short-term and long-term effect of the WIA reform on the Bound estimate.[20] We refer to the combined Equation (3.2) and Equation (3.3) as the restricted DiD model.

3.4.2 Identification

To estimate the age, period and cohort parameters in our specifications, we essentially build on two sets of usual identifying model assumptions and one additional identifying assumption that is relevant for the specification with distinct effects for awarded and rejected DI applicants. First, it is well-known that the identification of all APC parameters requires a constraint on the linear relationship between age, period and cohort effects. We do so by following Deaton and Paxson (1994), who assume that the average effect of period effects is equal to zero ($\Sigma_1^T\ \pi_t = 0$) and that there is no trend in period effects ($\Sigma_1^T\ t\pi_t = 0$). This resembles the idea that time effects reflect transitory business cycle effects. We will challenge this hypothesis in two ways. Most importantly, we will parameterize time effects as a function of business cycle indicators to investigate the robustness of our findings. We also provide a test on non-stationarity of period effects that allows for quadratic time period effects. If quadratic time effects matter, the assumption of stationary time effects is violated.

Our second key assumption is that age, period and cohort effects are orthogonal and additive. Most notably, this assumption implies that reform effects are captured by cohort effects that are constant in the years after application. Hence, we do not allow the elapsed duration profiles

[20] In light of the long time period that is observed after the WIA reform, we allow for a more flexible specification that distinguishes short-term from long-term effects. Obviously, the common trends assumption is more stringent for the long-term effects.

to change due to reforms.[21] To test for the sensitivity of our findings, we will therefore consider more flexible model specifications with distinct age profiles for time periods with different policy regimes. In doing so, we compare the accumulated cohort effects for this model with the model that assumes constant age profiles over the full time period under investigation.

Third, the identification of incentive effects in Equation (3.2) requires an additional assumption. Specifically, we assume common compositional changes for awarded and rejected applicants that lead to common employment trends stemming from compositional changes. We argue this assumption is plausible, since the medical assessment and the derivation of degrees of disability did not change fundamentally in the time period under consideration. Given the common compositional trends assumption, the increase in employment rates that is specific to the awarded applicants can be interpreted as the effect of changes in benefit conditions. This 'DiD' increase is the equivalent of the change in the Bound estimate.

In light of the eyeball tests in Section 3.3, the assumption of common changes in compositional effects is more plausible if rejected applicants are compared to awarded applicants which are (also) deemed to have substantial residual earnings capacity. This calls for the estimation of models where we compare partially awarded applicants to rejected applicants or compare samples that are stratified by the degree of disability. As far as there are secular trends or reforms with compositional effects, applicants with similar earnings capacities are likely to be affected equally by this. Another way to test for common changes in compositional effects will be to compare model outcomes with and without the inclusion of control variables that were discussed earlier. If compositional changes affect awarded and rejected applicants equally, estimates should not be affected.

In what follows, we start by presenting OLS estimation results of Equation (3.1) and Equation (3.2) without (time constant) control variables that may or may not embody application cohort effects. As a result, the cohort estimates show the composite impact of *all* time-invariant variables that affect employment. Later on we also estimate model versions that

[21] Likewise, we model common time effects for DI applicants that are observed in the first and in later years after the application moment.

include dummies for five-year age groups, gender, ethnicity, impairment types and the pre-disability employment status as controls, so as to obtain insight in the sources of compositional changes that drive cohort effects.[22]

3.5 Estimation Results

3.5.1 The Age-Period-Cohort model

Figure 3.6 graphically presents the elapsed time (or: 'age'), period and cohort profiles of the employment for our full sample of DI applicants for four model variants. Following from Equation (3.1), our primary focus is on the APC-DP model as a benchmark; the respective results are indicated by the black, solid lines. To start with, our benchmark model shows elapsed time profiles since the DI decision – i.e., the 'age'-effect – that display a kinked pattern (see panel A). Since individual controls are not included, the estimates reflect the joint long-term effect of application over time and the actual aging of applicants. The drop in employment is largest in the first and second year after the DI decision, amounting to a decrease of nearly 20 percentage points. In this period applicants awarded benefits may leave the labor market and a large fraction of rejected applicants is laid off by their employer.[23] Subsequently, the employment rate of applicants declines with approximately 2 percentage points per year, such that the total decrease after 17 years equals roughly 45 percentage points.

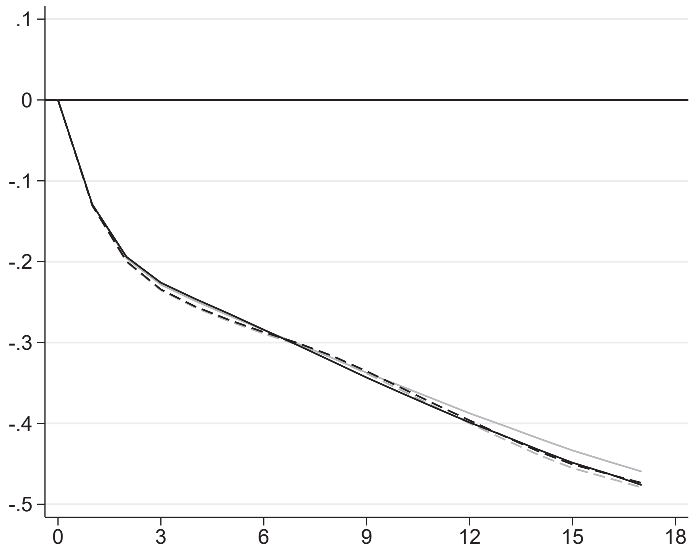
Panel B of Figure 3.6 points at small period effects. The spread of period effects is less than 5 percentage points, whereas the time and cohort effects add up to to about 45 and 30 percentage points, respectively. When comparing these findings to those of the AC model (without period effects), transitory period effects explain a negligible part of the variation

[22]We use the employment status in the year before application. We also estimated models using the employment status two years before application for cohorts after the WIA reform, taking into account that these applicant cohorts face a longer waiting period. This yields similar results.

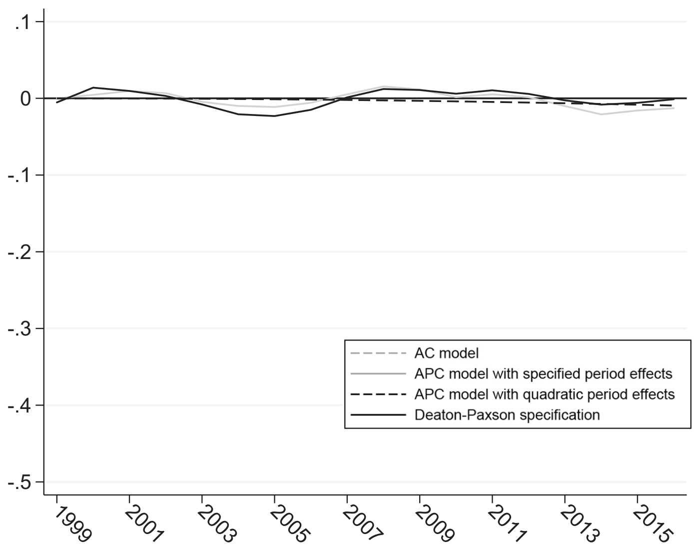
[23]Note that this contrasts to the SSDI system, where applicants typically have no (substantial) earnings from employment to begin with.

Figure 3.6: Elapsed time ('age'), period and application cohort effects on the employment of DI applicants

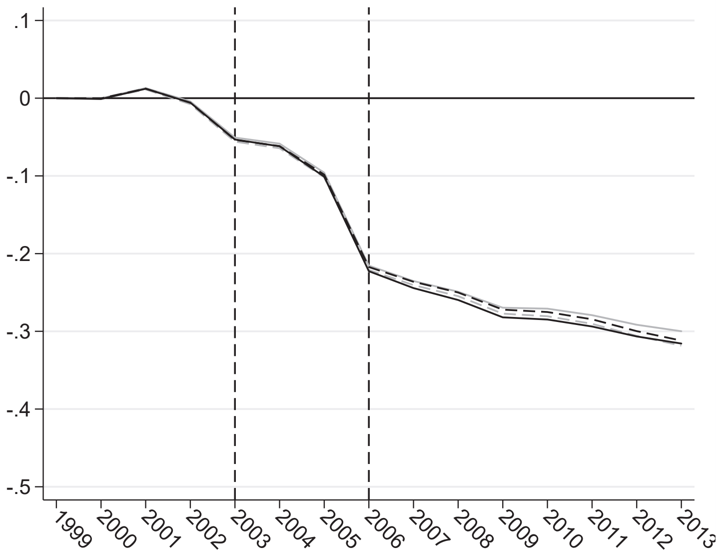
(a) Age effects



(b) Period effects



(c) Cohort effects



Note: The figure displays the results of the following four models: (i) Age-Cohort (AC) model, (ii) APC model with the period effects depend on the ratio of vacancies to unemployment and the employment rate of low-educated individuals, (iii) APC model with the period effects specified as a quadratic function, and (iv) Deaton-Paxson specification. The sample consists of all workers who applied for DI benefits in the Netherlands between 1999-2013.

in the employment.[24] Still, the small period effects of the DP model mimic business cycle patterns seemingly well, with peaks in 2001 and 2008.

Panel C indicates sizable application cohort effects, particularly when the GKP and WIA came into force. Changes in application cohort effects add up to a 30 percentage points difference between 1999 and 2013. This difference largely stems from a 4 percentage points drop in 2003 and another drop of about 13 percentages points in 2006. Again, these discrete changes suggest the impact of the Gatekeeper and the WIA reform. The cohort effects also show a continued decline in the years after the start of the WIA in 2006. In total, more than half of the change in cohort effects is confined to the reform years 2003 and 2006.

[24]The R-squared of the APC-DP model is 0.0683 and for the AC model 0.0680, respectively.

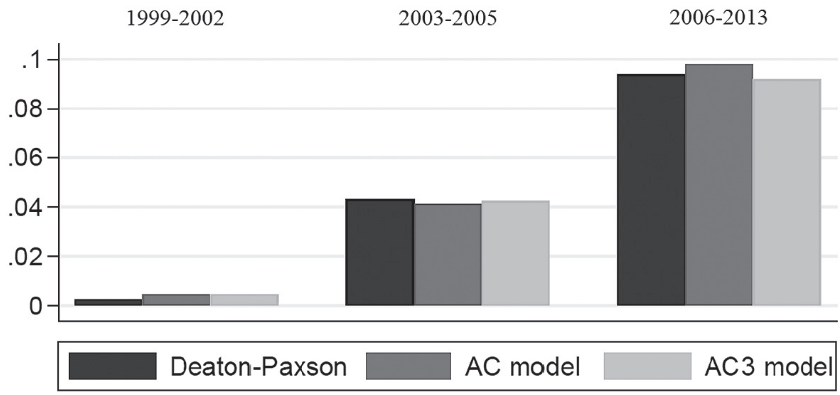
Robustness 3.5.2

In line with the orthogonality assumptions that were discussed earlier, we investigate the robustness of our findings. Since the DP-model does not allow for non-transitory period effects, our first concern is that structural trends are absorbed by the age and cohort effect estimates. We assess the stringency of the orthogonality assumptions of the DP model in two ways – the results are (also) shown in Figure 3.6. First, we estimate an APC model where period effects are specified as a quadratic function and age and cohort effects as (non-parametric) step functions. Albeit that this specification does not result in the (full) identification of APC effects, it enables us to estimate a part of identifiable non-linear period effects that may be non-transitory. Accordingly, the coefficient of the quadratic period effect, together with changes in age and cohort effect estimates, provides us with conservative tests on the existence of non-transitory period trends. The concerning coefficient is statistically significant, but the dashed black lines in Figure 3.6 show its magnitude is negligible and the accumulated application cohort and age effects are very similar to those for the DP model. Second, we consider parametric specifications where period effects depend on the ratio of vacancies to unemployment and employment rates of low-educated individuals. Arguing that low-educated individuals are over-represented among DI applicants, this auxiliary information can be used to proxy period effects that may also may show more structural trends e.g. arising from Skill-Biased Technological Change (SBTC).[25] We then find one percentage point increase in the employment rate of low-educated workers to be associated with a 0.6 percentage point increase in the period effect. While this estimate is statistically significant, the resulting range of period effects is comparable to those for the APC-DP model.

As stated earlier, our benchmark model also imposes orthogonality assumptions on the interrelation between age, period and cohort effects. As a result, changes in elapsed duration profiles induced by the reforms

[25] From 2005 onward, we observe employment rates of disabled individuals in the public scheme for disabled individuals that have no eligibility into the DI scheme (i.e., the 'Wajong'). For this limited time period, this variable did not yield a significant coefficient estimate.

Figure 3.7: Comparing implied absolute declines in cohort effects of three models, measured for 1999-2002, 2003-2005, and 2006-2013



Note: The three models: (i) Deaton-Paxson specification, (ii) AC model, and (iii) AC model with specific age and control effects for each of the three time periods.

are absorbed by applicant cohort effects. We therefore re-estimate the APC model with specifications with distinct age profiles and application cohort effects across three time periods: 1999-2002, 2003-2005, and 2006-2013. The implied changes in accumulated application cohort effects are shown in Figure 3.7. The bars in the figure indicate (i) the implied total change in cohort effects for the baseline DP model; (ii) for the AC model (i.e., without period effects); and (iii) for the AC model with distinct age and cohort effects for the three time periods.[26] The figure shows that the AC model yields application cohort effects for the three time periods that are virtually equal to those of the DP model. The negative application cohort effects after 2006 represent either learning or adaptation effects of the WIA reform or point at a secular trend in health and labor market conditions that are specific to new applicant cohorts.

[26]Note that the estimation of APC models with distinct age effects would give rise to identification problems of period effects. Since period effects we find are generally small for the total period, setting these equal to zero is not a strong restriction to make.

Table 3.2: DiD incentive effects of the Gatekeeper Protocol (GKP) and short-term and long-term incentive effect of the WIA reform

	Rejected vs partially allowed		degree of disability $< 35\%$ vs. $35-80\%$		degree of disability $< 35\%$ vs. $35-55\%$	
	(1)	(2)	(3)	(4)	(5)	(6)
$\tilde{\gamma}_{gkp}$	-0.005*	-0.005*	-0.001	0.000	-0.001	-0.001
	(0.00)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
$\tilde{\gamma}_{wia,shortterm}$	0.009*	0.026***	-0.009*	0.013***	-0.005	0.013***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.006)	(0.005)
$\tilde{\gamma}_{wia,longterm}$	0.029***	0.032***	0.018***	0.025***	0.018***	0.022***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Separate age, period and common cohort effects	✓	✓	✓	✓	✓	✓
Controls	—	✓	—	✓	—	✓
Observations	6,730,460	5,561,737	6,736,052	5,567,329	6,193,528	5,095,192
R^2	0.0642	0.2026	0.0650	0.2030	0.0622	0.2001

Note: Control variables include individual characteristics (age, gender, ethnicity), impairment types and employment history (employment status, UI benefit receipt and sector of employment). Individual clustered standard errors in the parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Decomposing incentive effects 3.5.3

Table 3.2 presents the estimation results for the incentive effect of the GKP and WIA reforms, $\tilde{\gamma}_{\tau}$, using the restricted (DiD) model of Equation (3.2) and Equation (3.3). Recall that the incentive effect measures changes in the Bound estimate, with positive changes pointing at a worsening of the employment probability of awarded applicants (and reverse). The findings for the restricted model are complemented with the unrestricted Bound estimates for all annual application cohorts – as shown in Figure 3.8. For both the restricted and the unrestricted model, we compare (differenced) application cohort effects of the following groups: (i) rejected applicants versus awarded applicants with partial benefits in columns (1-2); (ii) applicants with a degree of disability below 35% versus applicants with a degree of disability between 35 and 80% in columns (3-4); (iii) applicants with degree of disability below 35% versus applicants with a degree of disability between 35 and 55% in columns (5-6).

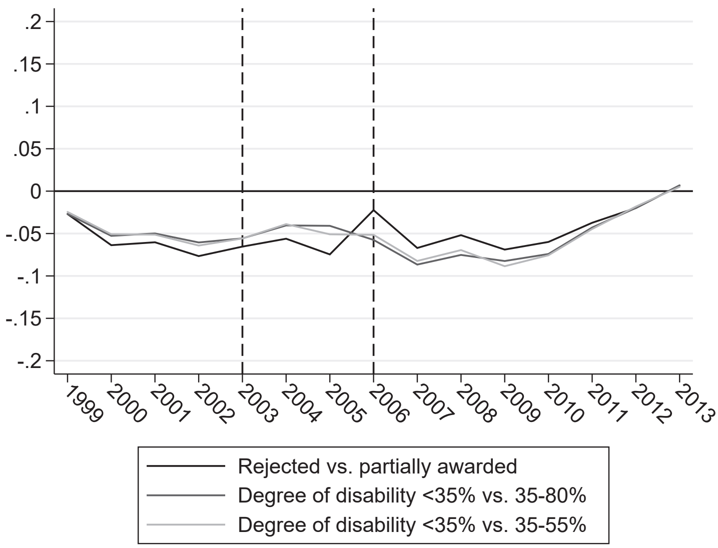
The DiD estimates in Table 3.2 suggest no incentive changes at the start of the GKP reform for all group comparisons. As the GKP aimed at changing the screening process before application, these results are in line with expectations and can be considered as placebo-outcomes. The evidence for the incentive effects of the WIA reform, however, is less clear-cut. As to the effects in the first four years since the reform (i.e., 2006-2009), all model specifications without controls show negligible and only weakly statistically significant estimates of the incentive effects.[27] The estimates increase somewhat after the inclusion of controls, suggesting that the common compositional cohorts assumption may be violated. For the long-term incentive effects (i.e., 2010-2013), our results indicate decreases of work incentives ranging between 2 and 3 percentage points – i.e., an increase in the Bound estimate – for partially awarded applicants. While these findings may appear more robust than the short-term effects at first sight, the common compositional trends assumption is more stringent for long-term effects.

We next move to the unrestricted Bound estimates as shown in Figure 3.8.[28] Similar to the graphical inference that was discussed earlier, the initial difference in application cohort effects is negative and fairly constant up till 2005 for all group comparisons. This again underlines the notion that the GKP increased the reintegration responsibilities during the waiting period for all DI applicants. For the WIA reform, again, there is no clear pattern that emerges. Depending on the stratification of groups, the Bound estimate can either stay more or less constant or decrease in 2006 (which implies a positive incentive impact). If any, Figure 3.8 suggests that the incentive effects of the WIA reform are small. Moreover, it appears unlikely that the increases in the Bound estimate after 2010 can be interpreted as the effect of the WIA reform. Taking a broader perspective, we are safe to say that the accumulated changes in application cohort effects by far cannot be explained by changes in DI benefit incentives.

[27] Recall that both Koning and van Sonsbeek (2017) and Kantarci et al. (2023) also find only small causal employment effects of the WIA reform.

[28] All parameter estimates of $\tilde{\gamma}_\tau$, both without and with controls, can be found in Table 3.3 in the Appendix, together with additional F-statistics which follow from multiple testing.

Figure 3.8: Annual Bound-estimates for the unrestricted APC-DP models



Note: The vertical axis displays the parameter estimates of $\tilde{\gamma}_\tau$ from Equation (3.2). The estimation allows for different age en period effects for rejected and awarded applicants.

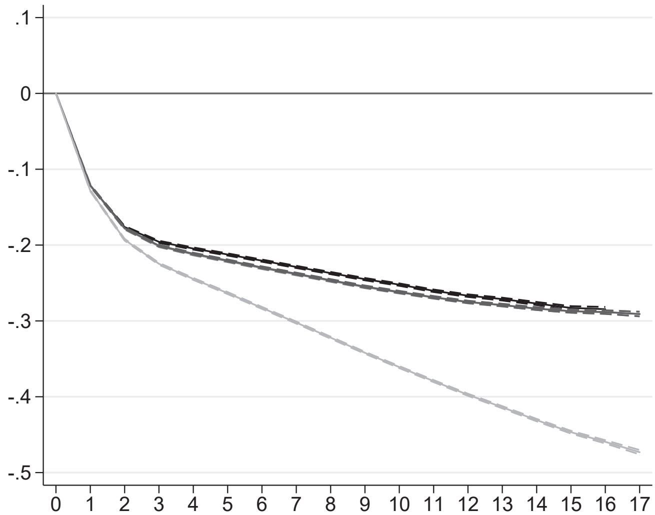
Application cohort effects in more detail 3.5.4

Our results so far point at sizable compositional cohort effects as the main driving force of employment trends, together with steep elapsed duration effects. We therefore investigate the origins of these cohort effects in three ways: we re-estimate the APC-DP model using various sets of controls and for sub-samples, as well as on outcome variables other than employment that are indicative for the labor market performance and health of applicants.

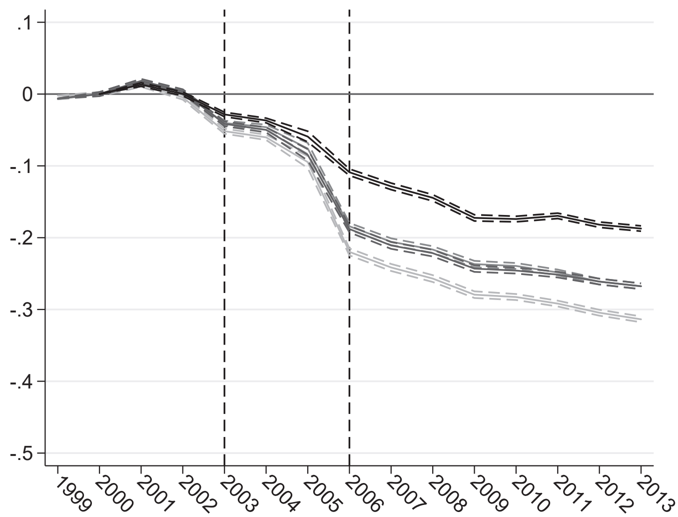
First, Figure 3.9 shows the results for the DP model with various sets of control variables that are added sequentially: (i) individual characteristics that include dummies for five-year age groups, gender and ethnicity; (ii) impairment types; and (iii) the employment history in the year before application (employment status, UI benefit receipt and the sector of previous employment). The inclusion of control variables causes the elapsed time

Figure 3.9: Deaton-Paxson estimation results of elapsed time ('age') and cohort effects with step-wise inclusion of sets of control variables

(a) Elapsed time effects



(b) Cohort effects



Note: The base specification (light grey line) is the model without control variables. We subsequently add: (i) dummies for age groups of five years, gender and ethnicity; (ii) impairment types; and (iii) employment status in the year prior application (employment status, UI benefit receipt and sector of employment). The dashed lines outline the 95-percent confidence intervals. The sample consists of all workers who applied for DI benefits in the Netherlands between 1999-2013.

effect estimates to level out after the first two years. More specifically, the observed change in elapsed time effect estimates in the first two years after application is almost entirely induced by the inclusion of age dummies.[29] We also see substantial reductions in application cohort effects with the control variables, suggesting self-screening effects on the average employment that occur before the DI decision. Specifically, the inclusion of age dummies and the employment status in the year before application reduces the application cohort effects substantially, whereas there are no changes in application cohort effects when including impairment types. Roughly speaking, about 40 percent of the 31 percentage points decline in employment rates of subsequent application cohorts is explained by self-screening on observed variables. As we have a limited set of controls, this estimate should be interpreted as a lower bound for the total effect of self-screening.

Interestingly, the inclusion of controls does not change application cohort effects substantially until 2006. The GKP probably discouraged workers with less-severe impairments from applying, rather than those with better labor market prospects. By contrast, the instantaneous drop in employment rates at the time of the WIA reform can largely be explained by the screening out of workers with better labor prospects, causing the remaining applicant pool to have less permanent contracts and a higher fraction being unemployed one year before application.[30] These compositional changes support the idea that the changes in application cohort effects are a representation of increased self-screening during the waiting period. In addition, we find that the gradual further decline in employment after the onset of the WIA reform can partially be explained by gradual compositional changes in observed controls.

Second, we estimated age and application cohort effects of samples that are stratified according to award decisions (rejected, partially awarded, fully awarded), gender, age groups (18-44 vs. 45-64) and impairment types

[29] The results are similar when we use 10-year age groups. The employment rates drop after the applicants reach their retirement age; this effect amounts to more than 20 percentage points.

[30] The newer application cohorts are also older (the last cohort is on average 5 years older than the first cohort), more often male (10%-points) and for a larger share non-native (8%-points).

(mental, musculoskeletal, cardiovascular and all other types) – these are all shown in Figure 3.13 in the Appendix. In line with expectations, we see larger and initially steeper age profiles for groups with a higher degree of disability, older ages and those diagnosed with cardiovascular disorders. This contrasts with rejected and partially awarded applicants and younger applicants who show more persistent employment profiles after the award decision. As to the application cohort effects, the initial decline since the start of the WIA is more substantial among those awarded full benefits, but next the partially awarded applicants catch up and experience a similar aggregate decline.[31] Changes in application cohort effects are most substantial for workers with mental impairments and already materialize in the year the GKP reform took place. This suggests moral hazard was present among workers with mental impairments, as the GKP implied stronger screening before application.[32]

Finally, Figure 3.14 in the Appendix presents age and cohort effects on wage earnings, contract types, Unemployment Insurance (UI) receipt and mortality rates of DI applicants. Panel A shows that earnings cohort effect have a similar pattern as the incidence of employment. Application cohort effects accumulate to 10,000 Euro per year, corresponding to 40 percent of the average income at the time of application. This effect is roughly equal to the extensive margin effect on employment. Panel B shows that the decline in application cohort effects of the probability on a permanent contract is roughly equal to that for all contracts, suggesting the decline is confined to permanent contracts. This finding underlines the importance of changes in the composition of application cohorts, with recent applicant cohorts having more vulnerable labor market positions. The evidence for UI benefit receipt in panel C suggests that application cohorts are more likely to loose (partial) employment. This may point at substitution effects into UI – see e.g. Koning and Van Vuuren (2010), Borghans et al. (2014) and Benitez-Silva et al. (2010). Finally, panel D

[31] Koning and van Sonsbeek (2017) argue that the stronger work incentives induced by the WIA has increased the relevance of a ‘cash-cliff’ to the fully and temporary disabled beneficiaries.

[32] Moral hazard may have been more important among workers with mental problems as it is a more heterogeneous group, with a high share of conflicts at work that prevent rehabilitation of sick-listed workers.

shows that mortality rates among applicant cohorts have increased after the reforms.[33] All these results highlight the importance of compositional effects among DI applicant cohorts as a driving factor of the changes in employment rates.

Conclusions 3.6

In this paper we expand on Age-Period-Cohort (APC) models to explain changes in the employment rates of Disability Insurance (DI) applicants. We use administrative data of DI application cohorts for the Netherlands, a country that experienced major disability reforms that intensified the screening process before application, tightened eligibility conditions and increased work incentives for benefit recipients. Using a Deaton-Paxson specification, we first decompose application cohort effects from period and age effects. The resulting application cohort effects represent the joint effect of (i) compositional changes induced by disability reforms; (ii) compositional changes induced by labor market trends; and (iii) behavioral changes in the employment rate of awarded applicants ('incentive effects') induced by reforms. To separate incentive from compositional effects, we develop a further decomposition that compares the employment rates of awarded applicants to those of rejected applicants. Assuming that compositional cohort effects for employment – both induced by reforms and changes in the labor market – affected both groups equally, the Difference-in-Differences (DiD) estimate of the reforms equals the change in the individual employment probability. This effect can be characterized as incentive effects of the reforms on benefit recipients.

We find that application cohort effects are the key driver of the observed decline in employment rates of DI applicants in the Netherlands. Both secular application cohort trends in the labor market and large instantaneous self-screening effects induced by reforms affected new applicant cohorts, rather than period effects or changes in work incentives for awarded applicants. Even though the period effects mimic the business cycle, its

[33]To calculate mortality rates, we follow the approach by Johansson et al. (2014) who use post-application mortality as proxy for ex-ante health.

relevance in explaining employment changes is negligible. Likewise, our further decomposition of application cohort effects into compositional and incentive effects suggests that changes in incentive effects are dwarfed by compositional effects. This highlights the importance of self-screening effects that were inherent with the reforms, with sick-listed workers that were discouraged to apply for DI benefits. Self-screening has dramatically changed the composition DI applicants, with less room for workers with residual earnings capacities who complement their labor income with benefits. Stated differently, the reforms have changed the targeting of the DI benefit scheme, rather than the work incentives.

Our results provide a novel perspective on evaluations that generally point at large inflow and enrollment effects of disability reforms in the Netherlands (De Jong et al. 2011, Godard et al. 2022, Koning and Lindeboom 2015). While these reforms have drastically changed the targeting of the DI scheme, the behavioral work impact of changes in the design of DI benefits has only been limited. This resembles recent findings from Haller et al. (2020), who also argue that reforms that change the eligibility of DI benefits have much stronger consequences for targeting than changes in benefits. Our findings also add to international analyses that suggest a trend of more vulnerable, low-skilled labor market groups becoming applicants for disability benefits (Autor and Duggan 2003, Maestas 2019, Von Wachter et al. 2011). Specifically, we find changes in the initial labor market position and sector of employment of applicants as important drivers of the observed decline in employment. This change applies to new applicant cohorts, rather than affecting all individuals that have applied for benefits at some point in time. To some extent, the dominant role of application cohort effects may stem from the relatively strict Employment Protection Legislation (EPL) that prevails in the Netherlands. In light of the constant inflow rates since that time and the absence of further reforms, it appears that gradual changes in the composition of applicant cohorts explain the employment decline since 2006. The higher share of vulnerable groups among applicants may point at a gradual sorting of low-skilled workers with health conditions into temporary and flexible jobs without employer obligations. This then points at changes in the underlying insured population of workers. We leave this topic for future research.

3.A Additional tables and figures

Figure 3.10: Fractions of awarded and rejected DI applicants by application cohort

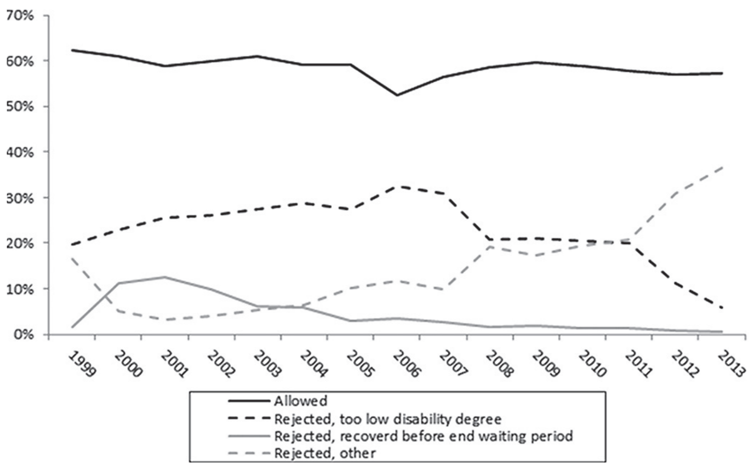


Figure 3.11: Cumulative distribution of the most important impairment groups of all applications for disability insurance between 1999-2013 by application cohort

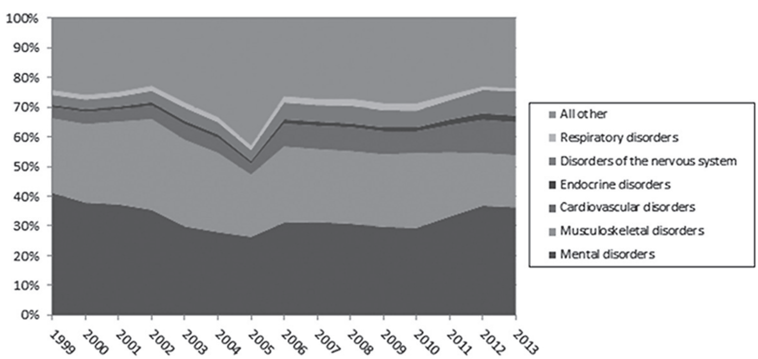
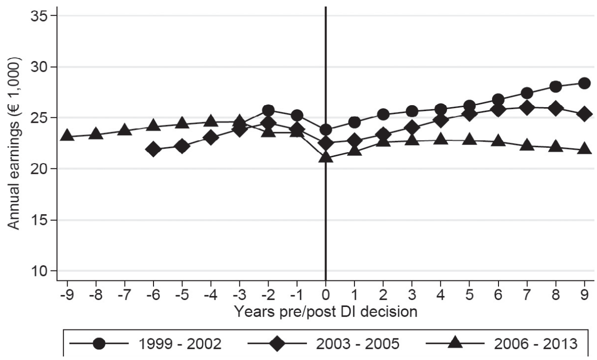
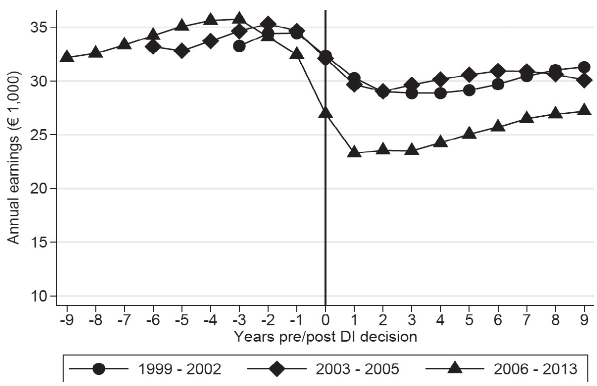


Figure 3.12: Annual average earnings of rejected, and partially and fully awarded applicant cohorts for three time regimes, before and after application for DI benefits

Panel A. Positive annual earnings of rejected applicants



Panel B. Positive annual earnings of applicants awarded partial benefits



Panel C. Positive annual earnings of applicants awarded full benefits

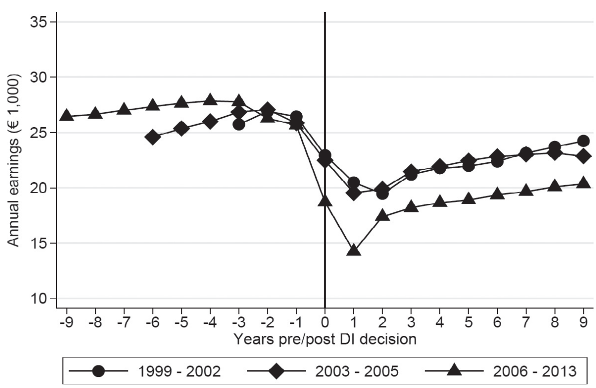


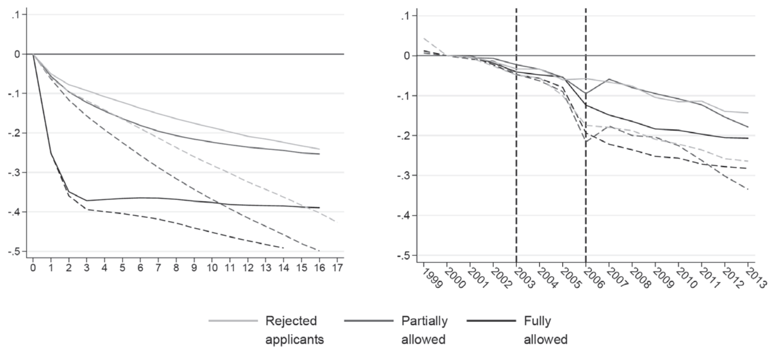
Table 3.3: Estimated cohort differentials of rejected vs. awarded DI applicants

	Rejected vs partially awarded		degree of disability $< 35\%$ vs. $35-80\%$		degree of disability $< 35\%$ vs. $35-55\%$	
	(1)	(2)	(3)	(4)	(5)	(6)
$\tilde{\gamma}_{1999}$	-0.027***	—	-0.027***	—	-0.025***	—
	(0.003)		(0.003)		(0.004)	
$\tilde{\gamma}_{2000}$	-0.064***	-0.035***	-0.053***	-0.028***	-0.051***	-0.023***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)
$\tilde{\gamma}_{2001}$	-0.060***	-0.032***	-0.050***	-0.028***	-0.051***	-0.024***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
$\tilde{\gamma}_{2002}$	-0.077***	-0.047***	-0.060***	-0.036***	-0.064***	-0.038***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
$\tilde{\gamma}_{2003}$ (GKP reform)	-0.065***	-0.046***	-0.056***	-0.036***	-0.056***	-0.038***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
$\tilde{\gamma}_{2004}$	-0.056***	-0.037***	-0.040***	-0.023***	-0.039***	-0.019***
	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)
$\tilde{\gamma}_{2005}$	-0.075***	-0.042***	-0.040***	-0.020*	-0.051***	-0.025**
	(0.010)	(0.009)	(0.012)	(0.010)	(0.014)	(0.012)
$\tilde{\gamma}_{2006}$ (WIA reform)	-0.023***	0.011	-0.057***	-0.014**	-0.051***	-0.009
	(0.008)	(0.007)	(0.008)	(0.007)	(0.010)	(0.009)
$\tilde{\gamma}_{2007}$	-0.067***	-0.026***	-0.087***	-0.035***	-0.082***	-0.034***
	(0.008)	(0.007)	(0.008)	(0.007)	(0.009)	(0.008)
$\tilde{\gamma}_{2008}$	-0.052***	-0.018***	-0.075***	-0.031***	-0.069***	-0.033***
	(0.008)	(0.007)	(0.007)	(0.006)	(0.009)	(0.008)
$\tilde{\gamma}_{2009}$	-0.069***	-0.034***	-0.082***	-0.040***	-0.088***	-0.046***
	(0.008)	(0.007)	(0.007)	(0.006)	(0.009)	(0.008)
$\tilde{\gamma}_{2010}$	-0.060***	-0.029***	-0.074***	-0.033***	-0.075***	-0.037***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.008)	(0.007)
$\tilde{\gamma}_{2011}$	-0.037***	-0.013**	-0.043***	-0.013**	-0.044***	-0.016**
	(0.006)	(0.005)	(0.006)	(0.005)	(0.007)	(0.006)
$\tilde{\gamma}_{2012}$	-0.020***	-0.009*	-0.020***	-0.004	-0.018***	-0.004
	(0.006)	(0.005)	(0.006)	(0.005)	(0.007)	(0.006)
$\tilde{\gamma}_{2013}$	0.007	0.015***	0.007	0.019***	0.005	0.017***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.007)	(0.006)
F-statistic differenced cohort effects						
All cohorts	23.40	17.14	18.68	11.51	13.70	7.92
Regime 1	55.16	9.86	18.61	3.40	18.72	6.50
Regime 2	2.51	1.93	3.47	3.45	2.83	3.46
Regime 3	16.98	9.37	29.37	13.87	19.25	10.07
F-stat differenced age effects	197.85	49.05	223.79	65.56	127.03	26.22
F-stat differenced period effects	16.28	20.36	14.03	16.11	9.11	10.84
Age, period and common cohort effects	✓	✓	✓	✓	✓	✓
Controls	—	✓	—	✓	—	✓
Observations	6,730,460	5,561,737	6,736,052	5,567,329	6,193,528	5,095,192
R^2	0.0645	0.2027	0.0650	0.2030	0.0623	0.2002

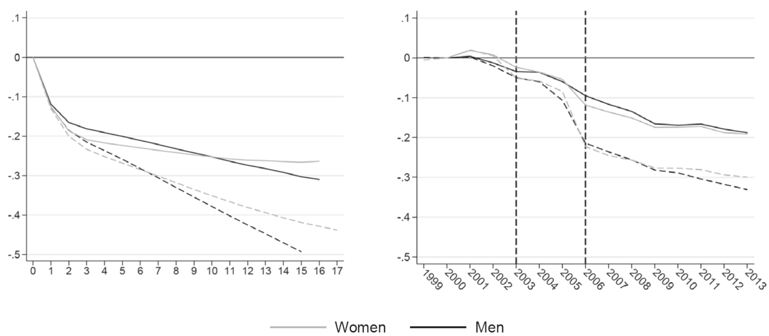
Note: Control variables include individual characteristics, impairment types and employment history. Reported F-statistics for multiple testing are Holm-adjusted. Individual clustered standard errors in the parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 3.13: Heterogeneous Deaton-Paxson estimates for age and cohort effects for employment

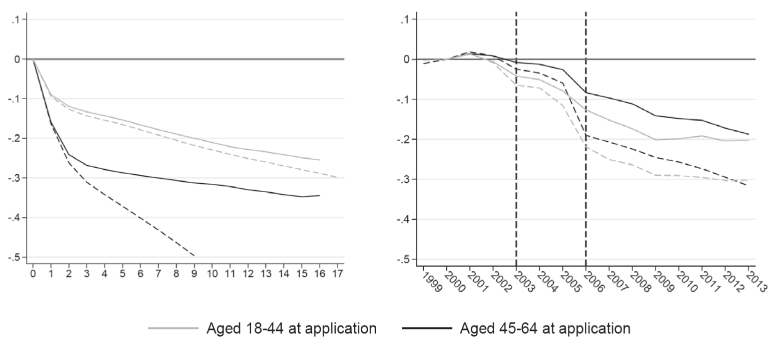
Panel A. Estimation results for rejected, and partially and fully awarded applicants



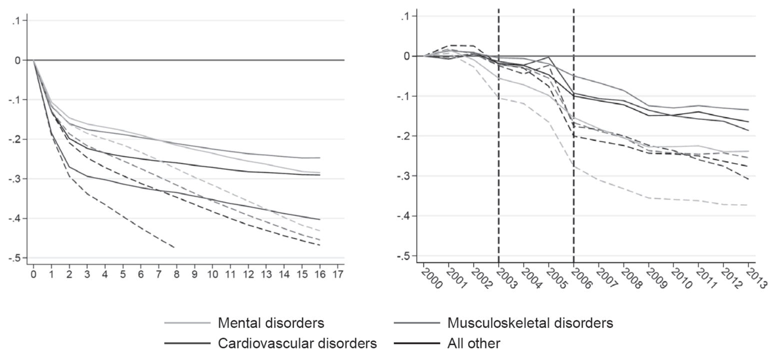
Panel B. Estimation results stratified by gender



Panel C. Estimation results stratified by age at application (18-44 vs. 45-64)



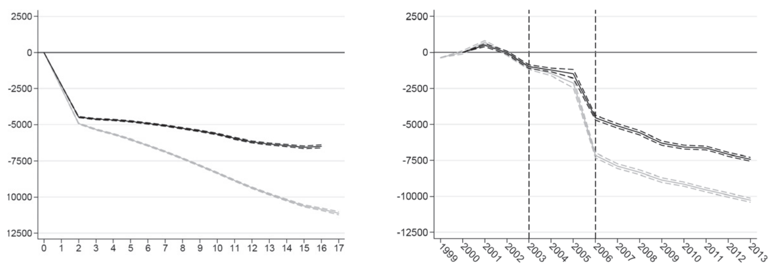
Panel D. Estimation results stratified by impairment types



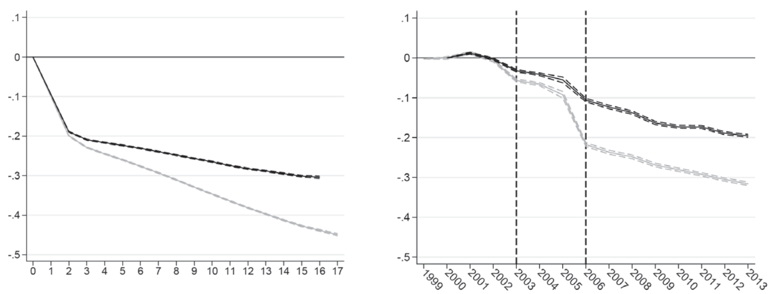
Note: Estimates without (dashed line) and with (solid line) control variables. Control variables include individual characteristics, impairment types and employment history. The sample consists of all workers who applied for DI benefits in the Netherlands between 1999-2013.

Figure 3.14: Deaton-Paxson estimation results of age and cohort effects for earnings, probability of a permanent contract, UI benefit receipt, social assistance benefit receipt and mortality

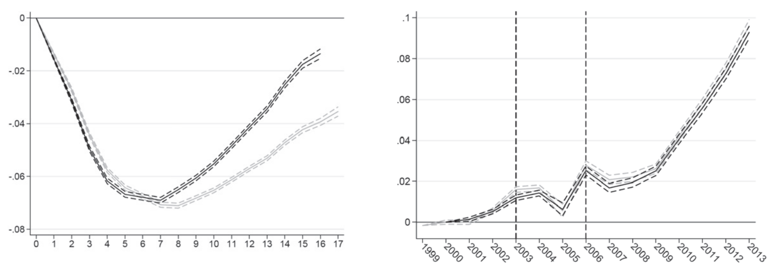
Panel A. Annual gross earnings (in 2015 Euros)



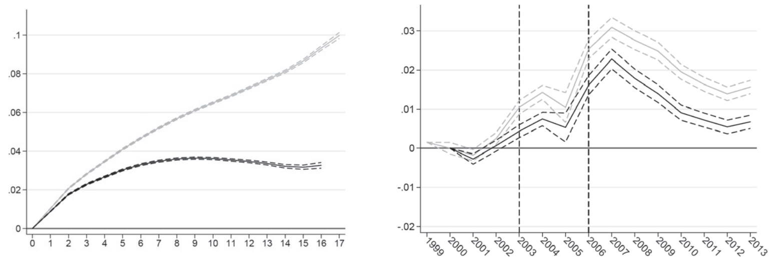
Panel B. Having a permanent contract



Panel C. Unemployment insurance benefit receipt



Panel D. Deceased



Note: Estimates without (grey) and with (black) control variables. Control variables include individual characteristics, impairment types and employment history. Dashed lines outline the 95-percent confidence intervals. The sample consists of all workers who applied for DI benefits in the Netherlands between 1999-2013.

4 Empirical evaluation of broader job search requirements for unemployed workers

Abstract

This paper exploits a large-scale field experiment where unemployed workers were randomly assigned to an additional caseworker meeting focussing on broader job search. The meeting significantly increases job finding and is cost effective. However, caseworkers differ in the rate at which they impose broader job search. We exploit this heterogeneity in caseworker stringency and the random assignment of unemployed workers to caseworkers to evaluate the broader search requirement. Our results show that imposing the broader search requirements reduces job finding. We argue that restricting the job search opportunities forces unemployed workers to search sub-optimally which negatively affects labor market outcome.

A working paper version of this paper is published as van der Klaauw and Vethaak (2022). This chapter is co-authored by Bas van der Klaauw. Bas van der Klaauw acknowledges financial support from a Vici-grant from the Dutch Science Foundation (NWO) and Heike Vethaak acknowledges financial support by Instituut Gak. The field experiment is registered under AEARCTR-0010370. We are grateful to Peter Berkhout for his indispensable help with the data and valuable comments. We also thank Bart Cockx, Marloes de Graaf-Zijl, Han van der Heul, Pierre Koning and Hans Terpstra, and seminar participants in Maastricht, at CREST-Paris, at the KVS New Paper Sessions 2020, the IZA/University of Sheffield Workshop: Evaluation of Labor Market Policies 2020, EEA-ESEM 2021, EALE 2021 and CEPR-Bank of Italy Workshop: Labour Market Policies and Institutions 2022 for useful comments to presentations and earlier drafts of the paper.

4.1 Introduction

In a recent study Belot et al. (2019) advocate that unemployed workers are searching for work too narrowly. This has drawn attention from both policy makers and researchers, who are considering policies to stimulate broader job search. Indeed, an increasing number of OECD countries require unemployed workers to search and accept jobs beyond the occupation of their previous employment. The underlying idea is that unemployed workers have biased beliefs about their labor market prospects. In particular, they anchor their reservation wage on their previous wage and search too often for work that resembles their previous job (Krueger and Mueller 2016, Mueller et al. 2021). Stimulating unemployed workers to search more broadly may then positively affect labor market outcomes and this would yield low costs to benefits administrations. However, Moscarini (2001) argues that only workers without comparative advantages should apply broadly for jobs while specialized workers should search narrowly.

In this paper, we empirically evaluate a program that enforces the requirement that unemployed workers search broadly for work. Individuals who have been collecting unemployment insurance (UI) benefits for six months are invited for a caseworker meeting to discuss job search strategies. When the caseworker concludes that the unemployed worker applies mainly for a narrow set of vacancies, she can give the unemployed worker a task to broaden the job search.[1] The unemployed worker is obliged to complete this task and this is monitored by the caseworker. In practice, it means that the unemployed worker should actively apply for jobs that are in different sectors, may have a longer commuting distance, offer a lower wage and may require a lower level of education.

For the empirical evaluation we use data from a large-scale field experiment conducted at the Dutch UI administration. A random subsample of about 130,000 unemployed workers has been invited to the caseworker meeting on job search strategies. We use this random assignment to estimate the causal effects of having the additional caseworker meeting.

[1]The Dutch law allows benefits recipients to only apply for jobs that meet their qualifications during the first six months of UI. After these initial six months benefits recipients are obliged to broaden their search to jobs that have lower requirements than their qualifications.

During this meeting the caseworker has the discretion to impose the broader search task on the unemployed worker. To estimate the causal effect of this broader search task we exploit that within local UI offices unemployed workers are randomly assigned to caseworkers and that there is substantial variation between caseworkers in the rate of imposing the broader search task. The identifying assumption is that if caseworkers differ in other dimensions that are important for supporting unemployed workers, these dimensions are orthogonal to the rate at which they impose the broader search task.

Our identification of the broader search task relates to the literature using judge stringency as instrumental variable. Kling (2006), Aizer and Doyle Jr (2015), Doyle Jr (2007, 2008) and Bhuller et al. (2020) use the random assignment to judges to estimate the effects of judge decisions on various socioeconomic and crime outcomes. Maestas et al. (2013) and French and Song (2014) use the assignment to an examiner to show that receipt of disability insurance benefits reduces labor supply. Most closely related to our approach is Arni and Schiprowski (2019), who use caseworker assignment to evaluate the relevance of job search requirements for unemployed workers. They consider a setting where caseworkers meetings occur more frequently (monthly) and search requirements can change between meetings. We study a setting with much less interaction between the caseworker and unemployed worker, which increases the plausibility of the validity of the empirical design.

Our paper contributes to the recent literature on broader job search requirements and to the literature on search requirements and active labor market programs. Within an online environment Belot et al. (2019) have randomly provided job seekers with additional vacancies to stimulate them to search more broadly. They find that this broader search encouragement increases the incidence of job interviews particularly for job seekers who initially searched narrowly. Altmann et al. (2018) randomly distributed an information brochure – with information about job search strategies and consequences of unemployment – among unemployed workers who are at risk of long-term unemployment and find that recipients of the brochure are more likely to find work. Skandalis (2019) shows that when the media announces intended hiring by plants, the composition of job applicants changes to individuals living further away. These studies show that job seekers benefit from the broader search induced by the information provision. The program we study in this paper has the same goal of stimulating broader job search, but as a formal policy it makes broader job search compulsory. This may imply that unemployed workers are restricted in their job search behavior and, therefore, are forced to search sub-optimally. We show within a simple job search model the potential effects of imposing the broader search task. The model shows that if unemployed workers do not have biased belief, the broader search task may stimulate job finding if narrow and broad search are close complements or do not differ substantially in their effectiveness. This coincides with Moscarini (2001) who argues that broader search is mainly useful for workers without comparative advantage.

Our paper further relates to a relatively extensive literature on caseworker meetings, and imposing and monitoring job search requirements. Recent studies by Maibom et al. (2017) and Schiprowski (2020) show non-negligible effects of caseworker meetings. They consider regular caseworker meetings, while we study a single meeting focussing on broader job search. The literature shows that additional job search requirements shorten the period of unemployment (Arni and Schiprowski 2019, Johnson and Klepinger 1994, Klepinger et al. 2002, Lammers et al. 2013). In our case the number of required job applications remains unaffected, but unemployed workers should also apply to jobs that are less closely related to their previous job. Finally, the caseworker meeting evaluates if the unemployed worker makes enough job applications and if these are already sufficiently broad. The caseworker meeting thus also contains an element of monitoring. The evidence on the effectiveness of job search monitoring is mixed (Van den Berg and Van der Klaauw 2006, Petrongolo 2009).

In the empirical analysis we use administrative data provided by the Dutch UI administration on all participants in the randomized experiment. Our evaluation of the experiment shows that the broader search program shortens the unemployment duration. We next exploit that unemployed workers are randomly assigned to caseworkers and that the rate at which caseworkers impose the broader search task is unrelated to other types of assistance. We find that imposing the broader job search task reduces the effect of the program, i.e. job finding is reduced after the broader search task. Even though being imprecisely estimated, marginal treatment effects suggest that broader search task are most often imposed on unemployed workers for whom the adverse effects are largest. Finally, we provide a decomposition of the effect of the broader search program in an effect of the broader search task and an effect of the meeting. This decomposition takes into account that groups of compliers differ when evaluating the program and the task. Our results differ from previous studies that often found positive effects of stimulating broader search. This shows the limitations of incorporating a broader search requirement in a formal (low-cost) policy.

The remainder of the paper is organized as follows. In the next section, we describe the Dutch UI system, the broader search policy, and the design and implementation of the experiment. Section 4.3 contains a description of the data and shows an evaluation of the broader search program. In section 4.4 we provide our empirical framework to estimate the effects of imposing the broader search task and we justify the use of caseworker stringency as instrumental variable. Section 4.5 presents some theoretical predictions of imposing the broader search task and shows the estimated effects as well as a decomposition of the program effects in an effect of caseworker meeting and the broader search task. Finally, section 4.6 concludes.

Background of the experiment 4.2

In this section we first provide a brief description of the Dutch UI system. Next, we discuss the content of the broader search program and finally we give some details on the experiment.

The Dutch UI system 4.2.1

In the Netherlands, the UI system insures workers against loss of working hours. If an individual worked 26 of the previous 36 weeks and loses at least five working hours, the individual becomes entitled to UI benefits.

During the first two months of UI the benefits level is 75% of the previous wage (capped at a maximum) and after that it becomes 70%. All eligible individuals are entitled to at least three months of UI benefits. The entitlement period to UI benefits depends on the work history.[2]

While collecting UI benefits, workers are obliged to (i) attend meetings with caseworkers when being invited, (ii) make at least one job application each week, and (iii) accept suitable job offers. During the first six months of UI, a job is considered suitable when it is in line with the worker's educational level, experience and previous wage. After these six months all jobs are considered suitable. During the first year of UI workers have three meetings with caseworkers, in the fourth, the seventh and the tenth month.

The meeting in the seventh month is affected by the experiment described in this paper. The meeting is eliminated for untreated individuals, while for treated individuals this meeting focuses on broader job search. The untreated individuals have the same (broader) search obligation, but since they do not receive the invitation letter and do not have the meeting this is less actively communicated.

4.2.2 The treatment

In 2015 the UI administration introduced a program to stimulate broader job search of workers who were collecting benefits for six months and for whom thus all jobs are considered suitable. Towards the end of the sixth month of UI, individuals receive a letter inviting them for the meeting with a caseworker in the seventh month of UI. This letter explains that the UI spell is approaching six months and that, therefore, the worker should apply for a broad set of jobs, including jobs requiring lower levels of education, in other sectors, with longer commuting times and lower wages than the previous job. The letter states that the purpose of the mandatory caseworker meeting is to discuss future job search strategies. The unemployed worker should bring two suitable vacancies, a curriculum vitae, past applications and the reactions of employers on these applications to

[2] De Groot and Van der Klaauw (2019) provide a more extensive discussion on the Dutch UI system.

the meeting. The untreated individuals do not receive the invitation letter for the caseworker meeting.

During the meeting the caseworker reviews the recent job applications. If the caseworker assesses the recent job search as narrow, the caseworker should give the unemployed worker a task to search more broadly. As a start of the task, the caseworker often provides two vacancies that are considered broader to which the unemployed worker must apply. When the broader search task is imposed, this is registered. Fulfilling the task is then an obligation and compliance can be evaluated in the subsequent months.

To summarize, the broader search program involves an invitation letter emphasizing the broader search obligation after six months of unemployment, a meeting with the caseworker evaluating the past job search and possibly a task for the unemployed worker to apply for jobs more broadly. Unemployed workers who are not subject to the program do not receive the letter, do not have the meeting and, therefore, cannot get the broader search task.

The experiment 4.2.3

The UI administration organized a randomized experiment with the intention to evaluate the broader search program. Excluded from the experiment are individuals who were older than 50 years and individuals who were entitled to less than ten months of UI benefits. A random subsample of the eligible workers who are approaching six months of benefits receipt between April 2015 and March 2017 were invited for the caseworker meeting discussing the broader search requirement. The randomization was organised such that individuals with one specific final digit of their social security number were not receiving the invitation letter for the caseworker meeting. Therefore, 10% of the eligible unemployed workers are assigned to the control group and the other 90% to the treatment group. Individuals who attend the meeting are assigned to a caseworker in their local office of the UI administration. The assignment is based on the current caseload of the caseworker, i.e. each unemployed worker is assigned to a caseworker with a caseload below the maximum caseload. In practice this often means that unemployed workers are assigned to the caseworker in their local office with the lowest caseload.

Table 4.1: Caseworker services received by the treatment and control group

	Treatment group	Control group	p-value
	(1)	(2)	(3)
Panel A: Services between 1-23 weeks			
Caseworker meeting	78.7%	79.1%	0.31
Contact by phone	0.9%	0.8%	0.80
Online contact	1.4%	1.7%	0.02
Panel B: Services between 24-36 weeks			
Caseworker meeting	62.4%	25.9%	0.00
Contact by phone	3.0%	9.0%	0.00
Online contact	12.0%	36.4%	0.00
Broader search task	43.1%	4.5%	0.00
Panel C: Services 37 weeks and later			
Caseworker meeting	17.6%	16.5%	0.00
Contact by phone	3.8%	4.0%	0.17
Online contact	24.7%	27.3%	0.00
Number of workers	118,697	13,420	

Note: The p-values in column (3) apply to t-tests of different means for the treatment and control group.

Table 4.1 shows for the treatment and control group how often they meet their caseworker in the period before the experimental intervention (1-23 weeks), during the experimental intervention (24-36 weeks) and after the experimental intervention (37 weeks and later). The population describes unemployed workers who have been collecting benefits for at least six months and thus entered the experiment. During the first 23 weeks of unemployment both in the treatment and control group about 80% of the individuals met their caseworker. Contact by phone or online contact is very rare in this period. After the randomization about 62% of the individuals in the treatment group and 26% of the individuals in the control group had a meeting with their caseworker. The individuals in the control group have much more often online contact or contact by phone. About 43% of the individuals in the treatment group and less than

5% of the individuals in the control group get a broader search task from their caseworker. This shows that the randomization actually affected the services provided to individuals, but compliance to the randomization is not perfect. The noncompliance is mainly caused by caseworkers not inviting individuals for a meeting, individuals that receive an invitation letter generally attend the caseworker meeting.[3] Panel C shows that after the period affected by the experiment, differences in services provided to the treatment and control group are modest.

Data and experimental evaluation 4.3

In this section we first provide a description of the data. Next, we consider the randomized experiment and show that participation in the broader search program significantly increases exit from UI.

Data description 4.3.1

For the empirical analysis we use administrative data available at the Dutch UI administration. Our sample contains all 132,177 individuals who participated in the randomized experiment. This means that they entered UI between October 2014 and September 2016 with a benefits entitlement period of at least ten months. In addition, they collected UI benefits for at least six consecutive months and were at that moment younger than 50 years. Individuals that previously worked as teacher or for the government are excluded as well as individuals participating in an entrepreneurship program.[4]

For each individual in the experiment sample, we observe if the individual was assigned to the treatment or control group, whether the individual attended the caseworker meeting during the seventh month of UI, the identity of the caseworker and whether a broader search task was imposed. In addition we observe for all individuals information on the UI

[3] Letters are imprecisely registered in our data, e.g. often there is no identifier of the content of the letter. So we cannot always determine exactly when the invitation letter for the caseworker meeting was send.

[4] Teachers and civil servants are covered by separate UI schemes.

spells (start and end date, monthly benefits payments and re-integration activities), employment contracts (start and end date, monthly earnings and working hours, type of contract and sector) and personal characteristics such as the date of birth, gender, nationality and level of education. We use the data to construct for all individuals a labor market history for the 32 months after starting collecting UI benefits.

Table 4.2: Descriptive statistics, balancing and compliance to the experiment

	Explanatory variables		Dependent Variables			
			Treatment group		**Meeting**	
	Mean	Standard Deviation	Coefficient Estimate	Standard Error	Coefficient Estimate	Standard Error
	(1)	(2)	(3)	(4)	(5)	(6)
Demographics						
Age	39.41	(6.39)	0.0000	(0.0002)	0.003***	(0.000)
Female	0.571	(0.495)	-0.0001	(0.0020)	0.032***	(0.003)
Dutch nationality	0.954	(0.210)	0.0085**	(0.0041)	0.016**	(0.007)
Low educated	0.176	(0.381)	–		–	
Middle educated	0.518	(0.500)	0.0025	(0.0023)	0.018***	(0.004)
High educated	0.307	(0.461)	-0.0014	(0.0028)	-0.001	(0.005)
Previous employment and benefit eligibility						
Monthly earnings (€)	2,325	(1,098)	0.0018	(0.0012)	0.008***	(0.002)
Hours per week	31.55	(9.10)	0.0000	(0.0001)	-0.001***	(0.000)
Max. entitlement (weeks)	87.63	(28.36)	0.0000	(0.0001)	-0.001***	(0.000)
Employed at 6 months UI	0.296	(0.457)	-0.0015	(0.0021)	-0.159***	(0.003)
Sector last job						
Financial	0.234	(0.424)	0.0037	(0.0034)	-0.002	(0.005)
Retail and trade	0.195	(0.396)	-0.0010	(0.0035)	0.020***	(0.006)
Health care	0.191	(0.393)	-0.0010	(0.0036)	-0.006	(0.006)
Temporary employment	0.088	(0.283)	0.0058	(0.0040)	-0.032***	(0.006)
Industrial	0.086	(0.281)	–		–	
Transport	0.057	(0.232)	-0.0022	(0.0045)	0.009	(0.007)
Other	0.149	(0.356)	0.0034	(0.0036)	-0.017***	(0.006)
F-statistic for joint significance			1.40		93.48	
[p-value]			[0.138]		[0.000]	
Number of workers = 132,177						

Note: OLS estimates of regressing assignment to the treatment group (column (3)) and attending the caseworker meeting (column (5)) on worker characteristics. Robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Columns (1) and (2) of Table 4.2 show for the different characteristics the sample means and the standard deviations, respectively. Individuals are, on average, slightly younger than 40 years, 57% are female and almost 95% have the Dutch nationality. The mean level of education is relatively high, 31% have a high education (university or college), 52% have a middle education (higher vocational or high school) and about 18% low education (lower vocational or primary education). Before entering UI, the average monthly earnings was 2,325 euros and individuals worked, on average, almost 32 hours per week. At the start of UI the average benefits entitlement period is almost 88 weeks. After six months of UI benefits almost 30% of the individuals have some employment and are thus collecting UI part-time.[5] Most individuals in the experiment entered UI after having worked in the financial sector, retail and trade, or the health care sector.

Table 4.2 shows that the treatment and control group are balanced. Column (3) presents the results of regressing assignment to the treatment group on the individual characteristics, and the standard errors are in column (4). There is only a significant effect of having the Dutch nationality on being assigned to the treatment group. Having another nationality is very rare in our sample and the size of the difference is small. All results are robust against using a sample of only individuals with the Dutch nationality. For all other characteristics we do not find any significant difference between the treatment and control group. The F-test at the bottom of the table shows that jointly all characteristics do not have a significant effect on the assignment to the treatment or control group.

Column (5) of Table 4.2 shows which characteristics predict the incidence of attending the caseworker meeting. This is informative on how compliance to the treatment differs between individuals.[6] The strongest effect is that individuals who have some part-time employment after six months of UI, so at the moment of the invitation, are less likely to attend the caseworker meeting. There are also some other characteristics that affect the likelihood that a worker will attend the caseworker meeting and

[5]The Dutch UI system compensates loss of weekly working hours. A worker can enter UI when losing part of the working hours and remaining working for the other part. Furthermore, when a UI recipients finds a part-time job with fewer working hours than the UI entitlement, the workers remains collecting UI benefits for the remaining hours.

[6]The regression uses the full sample. The estimation results are unaffected when only considering the treatment group.

all estimated covariate effects are jointly significant. This indicates that there is selection in which individuals meet the caseworker. The (limited) information available from the invitation letters seems to suggest that the selection is mainly induced by local offices not scheduling a meeting with all unemployed workers rather than the behavior of the worker, who may succeed in canceling the meeting.

4.3.2 Evaluating the broader search program

The randomized experiment evaluates the broader search program that starts with the caseworker meeting. The program can also contain the broader search task (when imposed during the caseworker meeting) and the monitoring of compliance to this task. Above it was shown that there is partial compliance to the random assignment to the treatment and control group. To deal with the partial compliance we use instrumental variable estimation. We specify the following regression equation for outcome Y_i observed for worker i,

$$Y_i = \alpha + \delta M_i + X_i'\beta + \varepsilon_i \tag{4.1}$$

The variable M_i indicates attendance of the caseworker meeting, so our parameter of interest δ describes the effect of participating in the broader search program. The effect also includes that attending the caseworker meeting can result in a broader search task, which may change job search behavior. The vector X_i contains all characteristics described in Table 4.2.

The initial assignment to the treatment group T_i is used as instrumental variable for attending the caseworker meeting. This provides the first-stage equation

$$M_i = \kappa + \gamma T_i + X_i'\phi + \nu_i \tag{4.2}$$

The randomization ensures that initial assignment is orthogonal to (unobserved) individual characteristics. To use initial assignment as instrumental variable it is also required that there are no other pathways in which initial assignment can affect outcomes. In practice, this requires that the invitation to the caseworker meeting should not directly affect outcomes. Recall that this is not the first meeting with a caseworker during the UI spell and it is a single not very time-consuming meeting. Therefore, we argue that behavioral responses to the invitation letter are unlikely. We also do not find differences in outcomes between the treatment and control group in the month at which the invitation letters were sent. The first-stage regression shows that the estimate for γ is 0.366, and the F-test statistic equals 8,213. The instrumental variable is very strong and when assuming monotonicity about 36.6% of the population in the experiment are compliers.[7] The parameter δ should be interpreted as the causal effect for these compliers.

Figure 4.1 presents the effects of enrolling in the broader search program for four outcomes for each month since starting collecting UI. Recall that the program starts with a meeting in the seventh month of UI and none of the outcomes shows any significant effect for the earlier period. Enrolling in the broader search program significantly reduces UI benefits receipt, after 12 months the fraction of individuals collecting UI benefits is reduced by about four percentage points (graph (a)). The effect on employment mirrors the effect on receiving UI, but after 12 months the effect declines slightly faster (graph (b)). After 18 months of UI the program effect disappears, so the program stimulates individuals to find work faster but the nonparticipants catch up later.

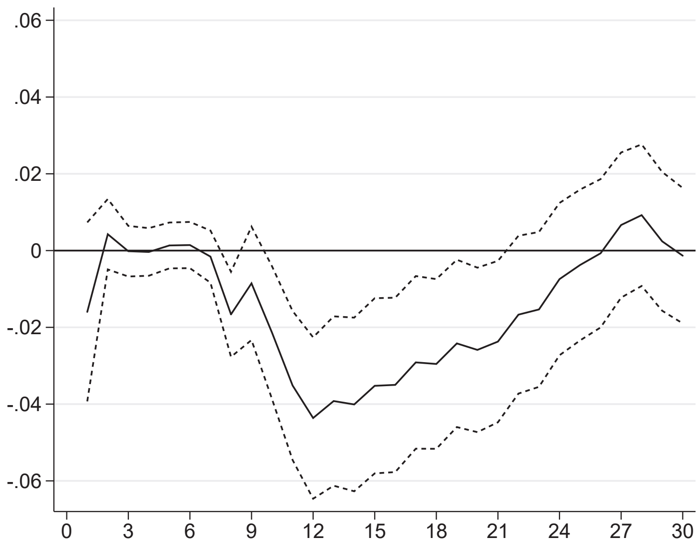
The bottom two graphs of Figure 4.1 show that the increased exit from UI also reduces the average monthly UI benefits payments. This effect is already significant in the eighth month and increases to about €100 less in benefits payments in the twelfth month. The effect on earnings is about half the effect on benefits payments and is never significant. This means that the reduced benefits receipt is, on average, not compensated by increased earnings from work.[8]

[7] In the control group 25.9% of the individuals attend a caseworker meeting ('always takers'), and in the treatment group 37.6% do not attend the caseworker meeting ('never takers').

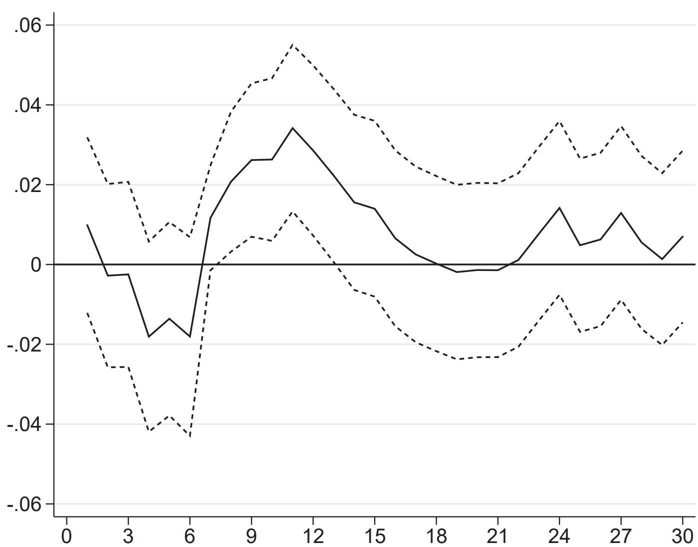
[8] Our data do not contain information on self-employment and other social insurance or welfare schemes. However, our target population is not eligible for other benefits schemes and exit to self-employment is minor and often not affected by labor market policies (De Groot and Van der Klaauw 2019).

Figure 4.1: Effects of participating in the broader search program - instrumental variable estimates

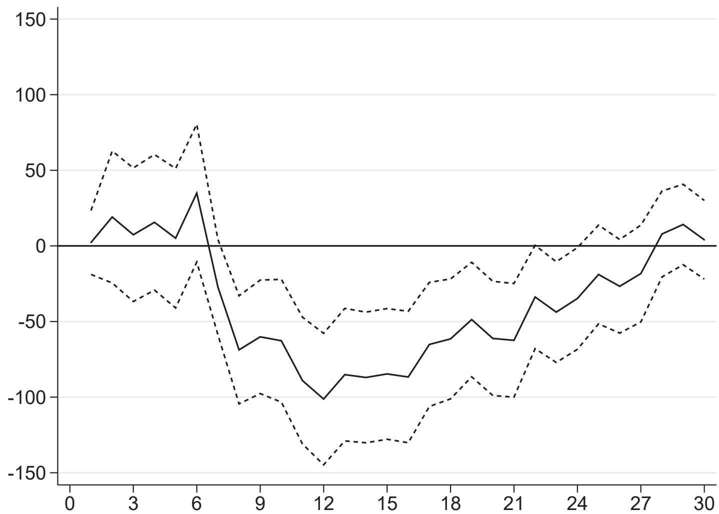
(a) UI benefits receipt



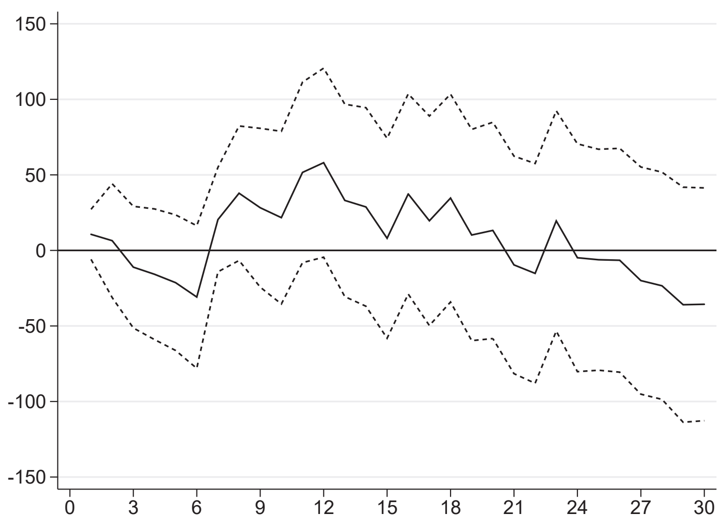
(b) Employment



(c) Amount of UI benefits



(d) Earnings



Note: The estimated effects are based on regressions including controls for age, gender, nationality, education, previous wage, sector and working hours, and UI benefits eligibility. $N = 132,117$ for all estimated effects in all panels. The 95% confidence interval are based on robust standard errors. $t = 0$ is the start of collecting UI, the broader search program starts with a caseworker meeting in the seventh month of UI.

Table 4.3: Effects of participating in the broader search program on cumulative outcomes - instrumental variable estimates

Dependent variable:	Weeks of collecting UI	UI Benefits	Weeks of employment	Earnings
	(1)	(2)	(3)	(4)
18 months after start UI	-1.41***	-879***	0.90**	379
	(0.34)	(174)	(0.43)	(291)
Dependent mean	36.76	10,357	28.24	11,635
30 months after start UI	-1.84***	-1,202***	1.15	265
	(0.60)	(264)	(0.85)	(630)
Dependent mean	51.11	13,893	64.06	29,050
Number of workers		132,117		

Note: All regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, and UI benefits eligibility. Robust standard errors in parentheses. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

Table 4.3 shows the estimated effects on cumulative outcomes 18 and 30 months after starting collecting UI, so about one and two years after the start of the program. These estimates confirm the earlier findings. The largest part of the program effects are in the year after the start of the program. In that period the UI period is reduced by, on average, 1.4 weeks and total benefits payments are 880 euros lower. The program is very cost effective for the UI administration, who estimate that the costs of offering the program are about €169 per invited individual.[9] The additional earnings for the participants are almost half of the reduction in UI benefits payments and insignificant. The increase in weeks of employment is also less than the reduction in weeks of UI.

[9]Figure 4.5 in Appendix 4.A provides a back-of-the-envelope calculation showing how the program costs compare to intention-to-treat effects of the program.

Broader search task 4.4

In the previous section we showed that participating in the broader search program stimulates the exit from UI and increases job finding. The two key elements of the program are a meeting with a caseworker and the broader search task. Caseworker meetings not always result in a broader search task. There is ample empirical evidence that attending a caseworker meeting positively affects job finding (Card et al. 2010, Maibom et al. 2017, Schiprowski 2020). There is much less evidence on mandating unemployed workers to search more broadly for work. In this section we present our empirical approach to estimate the effects of the broader search task.

Empirical approach and data 4.4.1

During the meeting the caseworker assesses the job search behavior of the unemployed worker. If the caseworker considers the search behavior as too narrow, the caseworker can give the unemployed worker a task to search more broadly. We exploit that within local offices of the UI administration unemployed workers are randomly assigned to caseworkers and that caseworkers differ in the rate at which they impose the broader search task. Our instrument variable approach is inspired by Bhuller et al. (2020), who use judge stringency as instrumental variable for incarceration and Arni and Schiprowski (2019) who use caseworker stringency as instrumental variable for required job search effort. For the empirical analysis, we use unemployed workers that attended the caseworker meeting, because only these individuals have a caseworker and only for them it is observed whether or not they received a broader search task. Furthermore, we restrict the sample to individuals who were assigned to the treatment group. This implies that we do not rely on variation induced by the randomized experiment. Individuals in the control group who attended a caseworker meeting are always takers in the randomized experiment. As will be discussed below excluding these individuals allows for a more straightforward interpretation and is necessary for the decomposition in Subsection 4.5.4.

We use the following regression equation to model how the outcome Y_{ic} of unemployed worker i at local office c of the UI administration depends on whether or not a broader search task B_{ic} has been imposed,

$$Y_{ic} = \alpha_c + \delta B_{ic} + X_i'\beta + \varepsilon_{ic} \tag{4.3}$$

The parameters α_c are the fixed effects for the local office at which the unemployed worker attends the caseworker meeting.[10] The vector X_i includes the characteristics discussed in Table 4.2. The parameter of interest δ describes the effect of the broader search task in addition to the caseworker meeting.

Caseworkers impose the broader search task when they believe that the unemployed worker focuses her job search activities too narrow. The broader search task is thus imposed on a selective subsample of unemployed workers who attend a caseworker meeting and the selection depends on unobserved job search behavior. However, caseworkers may assess job search behavior differently, which introduces exogenous variation in imposing the broader search task. An unemployed worker who receives a broader search task during the caseworker meeting, might not have received this task if she would have been assigned to another caseworker in the same local office (or vise versa). In accordance with earlier studies we refer to the rate at which a caseworker imposes the broader search task as the stringency of the caseworker (e.g. Aizer and Doyle Jr 2015, Bhuller et al. 2020, Kling 2006, Maestas et al. 2013).

Because within a local office unemployed workers are randomly assigned to a caseworker, whether or not an unemployed worker receives a broader search task depends on the stringency of her caseworker. This provides the first-stage regression equation

$$B_{ic} = \gamma_c + \lambda Z_{j(i)c} + X_i'\theta + \nu_{ic} \tag{4.4}$$

The instrumental variable $Z_{j(i)c}$ describes the stringency of caseworker $j(i)$ who is assigned to unemployed worker i. To compute the caseworker stringency faced by the unemployed worker, we use the leave-out mean.

[10] In the estimation we interact fixed effects for the local offices with calendar month to take account that the pool of caseworkers within a local office may change over time.

Table 4.4: Descriptive statistics, assignment of caseworker stringency and the broader search task

	Explanatory variables		Dependent variables			
			Caseworker stringency		**Broader search task**	
	Mean	Standard Deviation	Coefficient Estimate	Standard Error	Coefficient Estimate	Standard Error
	(1)	(2)	(3)	(4)	(5)	(6)
Demographics						
Age	39.47	(6.33)	0.0002	(0.0002)	0.002***	(0.001)
Female	0.592	(0.492)	0.0020	(0.0016)	0.051***	(0.005)
Native	0.960	(0.196)	-0.0010	(0.0031)	0.038***	(0.012)
Low educated	0.166	(0.372)	–		–	
Middle educated	0.520	(0.500)	0.0023	(0.0016)	0.044***	(0.006)
High educated	0.314	(0.464)	0.0040*	(0.0022)	0.064***	(0.007)
Previous employment and benefit eligibility						
Wage (€)	2,337	(1,094)	-0.0005	(0.0007)	0.001	(0.003)
Hours per week	31.28	(9.09)	0.0000	(0.0001)	0.000	(0.000)
Maximum entitlement	87.83	(28.03)	0.0000	(0.0000)	0.001***	(0.000)
Employed at 6 months	0.237	(0.425)	-0.0026*	(0.0016)	-0.160***	(0.006)
Sector last job						
Financial	0.244	(0.430)	0.0021	(0.0028)	0.012	(0.008)
Retail and trade	0.202	(0.402)	0.0011	(0.0027)	0.026***	(0.009)
Health care	0.194	(0.396)	0.0016	(0.0027)	-0.005	(0.009)
Industrial	0.087	(0.282)	–		–	
Temporary employment	0.079	(0.266)	0.0056*	(0.0031)	-0.040***	(0.011)
Transport	0.054	(0.227)	0.0007	(0.0035)	-0.023**	(0.012)
Other	0.139	(0.346)	0.0026	(0.0026)	-0.006	(0.009)
F-statistic for joint significance			1.32		89.35	
[p-value]			[.181]		[.000]	
	Number of workers = 42,605			Number of caseworkers = 461		

Note: OLS estimates of caseworker stringency (column (3)) and imposing the broader search task (column (5)) on individual characteristics. All regressions include controls for local office fixed effects interacted with month fixed effect. Standard errors are robust and clustered at the caseworker level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

This implies that we consider all other unemployed workers assigned to caseworker $j(i)$ (excluding unemployed worker i) and take the fraction that received a broader search task in this group. The leave-out mean is also used by Aizer and Doyle Jr (2015), Bhuller et al. (2020), Maestas et al. (2013).[11]

[11] Bhuller et al. (2020) argue that both the leave-out mean and the split-sample estimator perform well when the number of cases per judge is large enough. However, the split-sample estimator substantially reduces the sample. We perform the split-sample estimator as robustness check.

To estimate the model, we require information on the identity of the caseworker and whether or not a broader search task has been imposed. We can thus only use the sample of unemployed workers that actually attended the caseworker meeting. To keep the sample representative for the usual unemployed workers that would attend the caseworker meeting, we exclude workers in the control group of the experiment that attended the caseworker meeting.[12] In addition, we restrict the sample further to workers who attended a meeting with a caseworker who met with at least 50 and at most 400 unemployed workers during the experiment period.[13] After applying the sample selection criteria, the data include 42,605 workers and 461 caseworkers. Each of the 36 local UI offices has, on average, 13 caseworkers over the experimental period and each caseworker in our sample met about 92 unemployed workers participating in the randomized experiment. Column (1) of Table 5.2 shows the summary statistics of the subsample that we use to evaluate the broader search task. In terms of observed characteristics this sample is very similar to the full sample of participants in the experiment (see for comparison column (1) of Table 4.2).

4.4.2 Justification of the IV assumptions

The instrumental variable approach relies on three key assumptions, i.e. independence, exclusion restriction and relevance. Furthermore, for the interpretation of the estimates often a monotonicity assumption is made. Below we discuss the validity of these assumptions in our setting.

The validity of the instrumental variable approach relies on both an independence assumption and an exclusion restriction. The (conditional) independence of the instrumental variable is guaranteed by the random assignment of unemployed workers to caseworkers within local offices

[12] In the control group only the always takers in the randomized experiment attend the caseworker meeting, while in the treatment group both the always takers and the compliers attend the caseworker meeting. Including the attendants in the control group would bias the sample towards always takers.

[13] The minimum of 50 is imposed to obtain a reliable estimate for caseworker stringency. The maximum of 400 is used to exclude a few managers who register workers to themselves before assigning them to caseworkers.

of the UI administration. An unemployed worker thus gets randomly assigned a risk of receiving a broader search task. Since the unemployed worker can only be matched to a caseworker within the local office, it is essential to include fixed effects for local offices in the regression equations. To justify the random assignment of unemployed workers to caseworkers, we regress the caseworker stringency on the worker characteristics and fixed effects for the local offices (interacted with calendar time). The parameter estimates for this regression are shown in column (3) of Table 5.2. A joint test shows that worker characteristics do not predict stringency of the caseworker (p-value equals 0.181).

The independence assumption allows to give a causal interpretation to the estimate for λ in the first-stage regression. However, for the validity of caseworker stringency as instrumental variable also an exclusion restriction is required. The exclusion restriction imposes that caseworker stringency only affects the (labor market) outcomes of the unemployed worker via the broader search task. This rules out that caseworker stringency is correlated to assistance provided by the caseworker who may help the unemployed workers in finding work. Recall that the caseworker meeting is part of a new program that is evaluated using a randomized experiment. The program focuses solely on broader search and, therefore, the caseworker meeting has a clear agenda. This reduces the discretion for caseworkers to consider other interventions. Furthermore, the program only contains a single caseworker meeting, which limits the scope for caseworkers to provide additional support.

To provide some justification for the exclusion restriction we follow Arni and Schiprowski (2019) and consider other policy choices made by the caseworker. Column (1) of Table 4.5 shows summary statistics on the most frequent other interventions in the period after the caseworker meeting. Almost 19% of the unemployed workers are at some point in time after the meeting exempted from job search, while 7.4% get a punitive benefits reduction for not complying to the guidelines on the UI administration and 3.3% are assigned to participating in a (job search training) workshop. Column (3) shows that the use of these interventions is uncorrelated to caseworker stringency. Even without correcting for the local office fixed effects all estimated coefficients are already insignificant. So, caseworkers

Table 4.5: Use of other policy tools related to caseworker stringency and the broader search task

	Explanatory variables		Dependent variables			
	Mean	Standard Deviation	**Caseworker stringency**		**Broader search task**	
	(1)	(2)	(3)	(4)	(5)	(6)
Workshop participation	0.033	(0.178)	-0.006	0.003	0.082***	0.105***
			(0.006)	(0.005)	(0.012)	(0.012)
Benefits sanction	0.074	(0.263)	0.000	0.001	0.042***	0.046***
			(0.003)	(0.002)	(0.008)	(0.008)
Job search exemption	0.189	(0.392)	0.003	0.000	0.067***	0.061***
			(0.003)	(0.001)	(0.006)	(0.005)
F-statistic for joint test			2.34	0.34	83.33	84.10
[p-value]			[0.071]	[0.795]	[0.000]	[0.000]
Office x month FEs			—	✓	—	✓
Number of workers = 42,605			Number of caseworkers = 461			

Note: OLS estimates of caseworker stringency (columns (3) and (4)) and the broader search task (columns (5) and (6)) on other caseworker behavior. Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours and UI benefits eligibility. Standard errors in parenthesis are robust and clustered at the caseworker level. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

who are less likely to impose the broader search task do not compensate this with more frequent use of other interventions.[14] To provide further support for the exclusion restriction we include caseworker behavior as additional regressors in our empirical model. This validity check is discussed in Subsection 4.5.2 and shows that the estimated effects of the broader search task are robust to including these additional regressors.

The instrumental variable approach requires that caseworker stringency is relevant, which means that it has sufficient explanatory power on assigning the broader search task. The explanatory power is expressed in the first-stage regression, and in particular in the parameter λ. Table 5.5 shows the estimate for λ in the full sample and also for individuals with

[14]Caseworker added-value may be heterogeneous and related to caseworker stringency. To test this we consider unemployed workers that were to old to participate in the experiment (above 50) and who entered UI in the six months before the experimental sample. OLS estimates for this sample do not show any significant relation between the stringency of the caseworkers they met and job finding or exit from UI.

different characteristics. In the full sample the estimate is 0.826, which implies that when an unemployed workers is assigned to a caseworker who imposed the broader search task during 90% of the meetings instead of 50%, the probability that this worker receives is broader search task is about 0.33 higher. The estimate for λ is highly significant and does not differ much between individuals with different characteristics. From this we conclude that caseworker stringency is a relevant and a very strong instrumental variable (F-statistic equals 1,778).[15]

To interpret the instrumental variable estimates the monotonicity assumption is helpful. This assumption states that when the stringency of the caseworker increases, the treatment status of the unemployed worker can not switch from receiving broader search task to not receiving this task. Since caseworker stringency is a continuous instrumental variable, our estimate for the broader search task is a weighted average of marginal treatment effects (see Subsection 4.5.3 for an analysis of marginal treatment effects). Figure 5.3 displays the distribution of caseworker stringency unconditional and conditional on local office-month interactions. In the unconditional distribution (left panel), the caseworker at the 5th percentile imposes the broader search task in 46% of the meetings, while this is 91% for the caseworker in the 95th percentile. Roughly speaking under the monotonicity assumption about 46% of the unemployed workers are always takers (receive the broader search task from all caseworkers), 9% are never takers (none of the caseworkers assigns them a broader search task) and the remaining 45% are compliers (depends on which caseworker they are assigned to if they receive the broader search task).[16] Our estimates for the effect of the broader search task are informative on the compliers. When we condition on the local office-month fixed effects (right panel),

[15]In Appendix 4.B we show the robustness by including additional controls (Table 5.10), the split-sample approach (Table 4.11) and the reverse-sample approach (Table 4.12). In the split-sample approach the sample is randomly split in two. The first sample is used to compute the caseworker stringency, while the second sample is used for the regression. In the reverse-sampling approach, the caseworker stringency is computed using opposite types, e.g. for women we use in the regression the caseworker stringency computed on men.

[16]Maestas et al. (2013) consider the caseworker with the lowest (0.06) and highest (0.98) stringency and multiply the difference with the first stage coefficient to obtain the share of compliers $0.826 \times 0.92 = 76\%$. This computation is sensitive to outliers, which may be caseworkers that only meet few unemployed workers.

Table 4.6: First-stage estimates by demographics

	Coefficient	S.e.	F-stat	N	Dependent Mean
	(1)	(2)	(3)	(4)	(5)
Full sample					
Full sample	0.826***	(0.020)	1,778	42,605	0.730
Gender					
Female	0.827***	(0.027)	921	25,207	0.752
Male	0.820***	(0.025)	1,036	17,398	0.698
Nationality					
Native	0.824***	(0.020)	1,644	40,899	0.733
Non-native	0.973***	(0.123)	63	1,706	0.663
Educational level					
Low educated	0.862***	(0.047)	336	7,063	0.674
Middle educated	0.816***	(0.028)	877	22,156	0.764
High Educated	0.823***	(0.034)	571	13,386	0.709
Age					
Younger than 40	0.845***	(0.027)	949	20,323	0.709
Older than 40	0.813***	(0.028)	857	22,282	0.750
Employment status					
Not employed at 6 months	0.818***	(0.024)	1,207	32,504	0.769
Employed at 6 months	0.850***	(0.044)	369	10,101	0.605

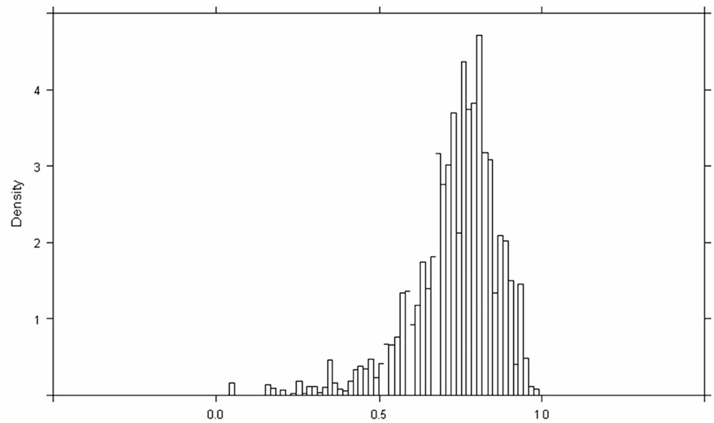
Note: Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. Standard errors are robust and clustered at the caseworker level. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

the standard deviation of the distribution increases, 0.151 compared to 0.135 unconditional. This implies that variation in caseworker stringency is not driven by variation between local offices.

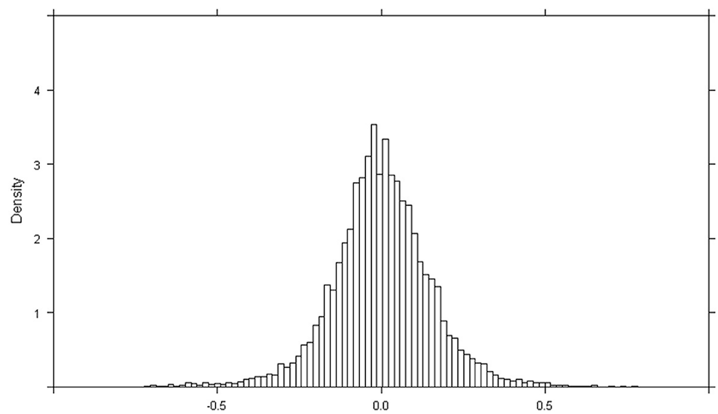
The monotonicity assumption is violated when different caseworkers are strict to different groups of unemployed workers (De Chaisemartin 2017). For example, younger caseworkers may impose the broader search task more often to younger unemployed workers and less often to older unemployed workers, while an older caseworker does the opposite. If there are more younger than older unemployed workers, the younger caseworker is stricter but monotonicity is violated. Although this is not concluding evidence, recall from Table 5.5 that there are hardly any differences in how likely individuals in different groups receive the broader

Figure 4.2: Distribution of caseworker stringency and demeaned by local UI office and month

(a) Caseworker stringency



(b) Caseworker stringency (demeaned)



search task. Imbens and Angrist (1994) consider this as support in favor of the monotonicity assumption.

4.5 Effects of the broader job search task

In this section we first show using a theoretical job search model how imposing the broader search task can affect the job search behavior of unemployed workers and their labor market outcomes. Next, in Subsection 4.5.2 we present the estimated effects, which show that imposing the broader search task reduced exit from UI. In this subsection we also discuss the robustness of the estimated treatment effects and heterogeneous treatment effects. In Subsection 4.5.3 we explore that there is substantial variation in the stringency of caseworkers and show marginal treatment effects. Finally, in Subsection 4.5.4 we decompose the estimated effect of the program in an effect of the broader search task and an effect of the caseworker meeting.

4.5.1 Theoretical predictions

To get an idea about the expected effects from the broader search requirement, we consider a job search model with two search channels following Van den Berg and Van der Klaauw (2006). We refer to the two search channels as narrow search n and broader search b. A worker has to decide how much effort she devotes to narrow search s_n and to broader search s_b. The rate at which narrow and broader search effort result in job offer is given by $\lambda_n s_n$ and $\lambda_b s_b$, respectively. Search is costly to the unemployed worker and the costs are described by $c(s_n, s_b)$. Since broader search implies in practice searching for jobs that pay a lower wage, we assume that the wage offer distribution from narrow search $F_n(\cdot)$ first-order stochastically dominates the wage offer distribution from broader search $F_b(\cdot)$ and that broader search effort is more likely to result in a job offer $\lambda_b > \lambda_n$. The unemployed worker accepts all wage offers that exceed the reservation wage ϕ. As shown in Van den Berg and Van der Klaauw (2006) the optimal reservation wage follows from solving the Bellman's equation

$$\phi = \max_{s_n, s_b \geq 0} \omega - c(s_n, s_b) + \sum_{j=\{n,b\}} \frac{\lambda_j s_j}{\rho} \int_{\phi}^{\infty} w dF_j(w)$$

where ρ is the discount rate and ω the level of UI benefits.

Van den Berg and Van der Klaauw (2019) discuss two specifications for the search costs function. First, substitution between both search channels $c(s_n, s_b) = (s_n + s_b)^2$ and second effort devoted to both channels are complements $c(s_n, s_b) = s_n^2 + s_b^2$.[17] We consider both specifications below and discuss the consequence of imposing a broader search requirement. The broader search requirement means $s_b \geq \bar{s}_b$, so broader search effort should exceed a minimum $\bar{s}_b$ set by the UI administration.

If both search channels are substitutes, then the optimal behavior of the unemployed worker is to devote only search effort to the channel with the highest return. If the broader search channel is the channel with the highest return, then the worker only searches broadly and should satisfy the broader search requirement already.[18] This is typically the case where caseworker does not impose the broader search task and the unemployed worker should be considered as a never taker. The more interesting case is when the narrow search channel yields the highest returns. In that case the unrestricted optimal search behavior is $s_b^* = 0$ and

$$s_n^* = \frac{\lambda_n}{2\rho} \int_{\phi^*}^{\infty} w dF_n(w)$$

and the job finding rate equals $\lambda_n s_n^* (1 - F_n(\phi^*))$ where ϕ^* is the (unrestricted) reservation wage. The broader search task sets $\tilde{s}_b = \bar{s}_b$, which changes the optimal narrow search effort to

$$\tilde{s}_n = \frac{\lambda_n}{2\rho} \int_{\tilde{\phi}}^{\infty} w dF_n(w) - \tilde{s}_b$$

Since unemployed workers are restricted in their job search behavior $\tilde{\phi} < \phi^*$. Furthermore, the total search effort increases, $\tilde{s}_b + \tilde{s}_n > s_n^*$. Because $\lambda_b > \lambda_n$, the job offer arrival rate increases, i.e. $\lambda_b \tilde{s}_b + \lambda_n \tilde{s}_n > \lambda_n s_n^*$. Accepted wages decline for two reasons, first the reservation wage declines

[17] Van den Berg and Van der Klaauw (2019) allow for different marginal costs of effort to both channels. However, for our purpose this is not necessary since we can always scale s_n, s_b, λ_n and λ_b, such that the marginal costs are similar.

[18] If the broader search requirement requires more effort than the optimal effort, the worker should increase the broader search effort and reduce the reservation wage. This increases job finding and reduces the average accepted wage.

and second a share of the job offers is now drawn from $F_b(\cdot)$ rather than $F_n(\cdot)$. Finally, the effect on the job finding rate is ambiguous. If there is a large difference between $F_n(\cdot)$ and $F_b(\cdot)$, then many job offers obtained via broader search will be declined by the unemployed worker, while fewer job offers are generated via narrow search.

Now consider the case that both channels are complements, i.e. $c(s_n, s_b) = s_n^2 + s_b^2$. In that case the optimal search is given by

$$s_j^* = \frac{\lambda_j}{2\rho} \int_{\phi^*}^{\infty} w dF_j(w) \qquad j = n, b$$

If the optimal effort to broader search s_b^* already exceed the minimum requirement $\bar{s}_b$, then the minimum broader search task does not affect job search behavior. The more interesting case is when the broader search requirement causes that unemployed workers have to devote more effort to broader search, so that $\tilde{s}_b = \bar{s}_b$. The optimal amount of narrow search effort becomes

$$\tilde{s}_n = \frac{\lambda_n}{2\rho} \int_{\tilde{\phi}}^{\infty} w dF_n(w)$$

Since the unemployed worker is restricted in her behavior, the reservation wage decline $\tilde{\phi} < \phi^*$ and thus narrow job search effort also increases $\tilde{s}_n > s_n^*$. Because the broader search task increases both broad and narrow job search effort and reduces the reservation wage, the job finding rate increases. Furthermore, expected accepted wages will decline since the reservation wage decreases and because a larger share of the accepted jobs will be found broadly.

Van den Berg and Van der Klaauw (2019) estimate the cost function of effort and find that different channels are almost perfect substitutes.[19] In that case, a broader search requirement increases total job search effort, reduced mean accepted wages, but the effect on job finding is ambiguous and depends on how acceptable job offers from the broader search channel are. These effects only apply to unemployed workers who without

[19] Van den Berg and Van der Klaauw (2019) distinguish between formal and informal job search and parameterize the cost function as $c(s_1, s_2) = (s_1^\gamma + s_2^\gamma)^{2/\gamma}$. Their structural analysis shows that γ is close to 1.

the broader search task devote their effort mainly to narrow job search. Moscarini (2001) argues that in equilibrium specialized workers search narrow because they have a comparative advantage in narrow job search.

Our job search model assumes that unemployed workers are fully informed about the job search environment. This is a strong assumption, unemployed workers may be too optimistic about job finding. When unemployed workers overestimate the returns to search, they set their reservation wages too high (Krueger and Mueller 2016, Mueller et al. 2021). Belot et al. (2019) argue that initially too optimistic unemployed workers target their job search towards better jobs and, therefore, search narrowly, but they get more pessimistic when search is unsuccessful. When they have updated their beliefs sufficiently, Information on broader search may induce unemployed workers to update their beliefs faster and consequently change their job search behavior. The theoretical prediction of Altmann et al. (2018) is that when the additional information makes beliefs of the unemployed workers more realistic, then job search behavior becomes more efficient. Both Belot et al. (2019) and Altmann et al. (2018) state that providing information is particularly useful for individuals at risk of staying unemployed long-term. Recall that our broader search program targets individuals who have been collecting UI benefits for six consecutive months, which makes it likely that the target population is responsive.

Estimated effects 4.5.2

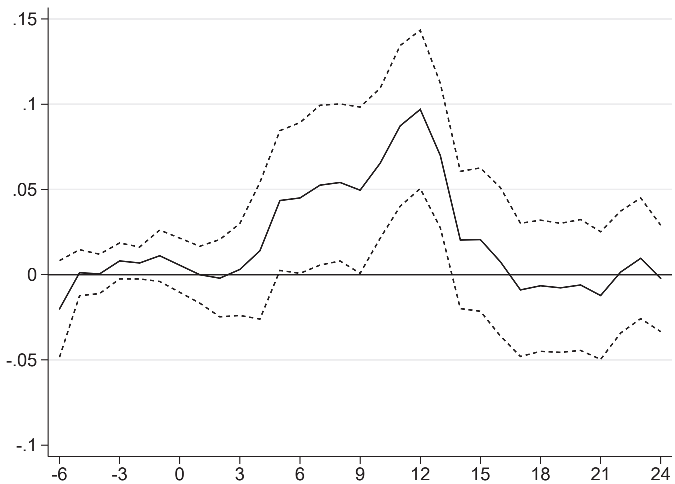
We use the empirical model discussed in the previous section to estimate the effects of imposing the broader search task for each month since the caseworker meeting. Figure 4.3 shows the estimated effects on the four main labor market outcomes, where time $t = 0$ is the moment of the caseworker meeting. Six months after the broader search task is imposed, the exit rate from UI significantly decreases (graph (a)). The broader search task causes that one year after the meeting the dependency on UI benefits is almost 10 percentage points higher. At that moment employment is about five percentage points lower and this is just significant (graph (b)). So the increased dependency on UI benefits is only partly explained by reduced job finding. One year after the meeting the effects diminish relatively fast, which might be partly due to unemployed workers reaching the end of the UI entitlement period. Recall that the individuals in our sample, on average, are entitled to 18 months of UI benefits, which is one year after the caseworker meeting. The average amount of UI benefits payments is somewhat higher when a broader search task is imposed (graph (c)). Again this peaks one year after the caseworker meeting and then the effect is just significant. The effects on earnings (graph (d)) show the opposite pattern with the exception that beyond one year after the caseworker meeting the effects remain negative rather than that they diminish. However, the effects on earnings are never significant.

Table 5.6 presents the estimated effects on cumulative outcomes one and two years after the caseworker meeting. The broader search task increases the period of collecting UI benefits with about 2.3 weeks. The majority of this effect is already present one year after the caseworker meeting. This also holds for the cumulative amount of the UI benefits payments, which increases by, on average, €800. The increased benefits payments cause that imposing the broader search task is costly for the UI administration. The cumulative weeks of employment decreases less than the weeks of UI benefits increase. The broader search task causes that individuals have about 1.6 fewer weeks of employment. After one year the negative effect on earnings has about the same size as the positive effect on UI benefits payments. The negative effect on earnings increases further during the second year (the large standard error causes that the effect is not significant). We should, however, be careful in concluding that individuals financially suffer from receiving the broader search task. When UI benefits end during the second year, individuals may become eligible for welfare benefits and these benefits are not registered in our data.

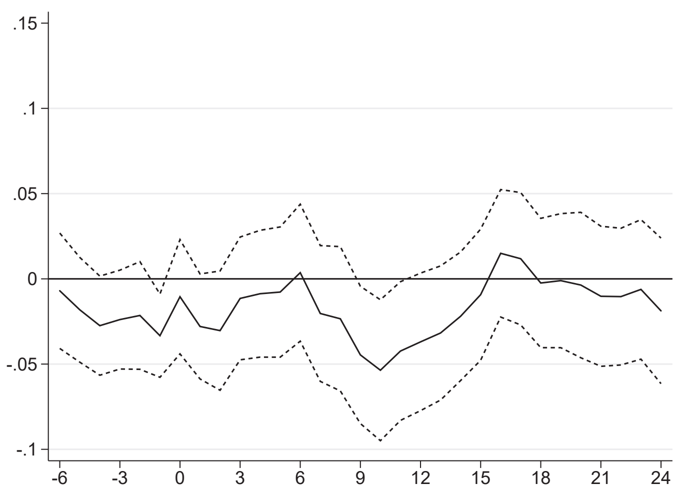
Our estimation results show that imposing the broader search task has adverse effects on job finding. This contradicts that the broader search task repairs systematic mistakes in the job search behavior, for example because unemployed workers are too optimistic about their labor market prospects. We use the theoretical predictions from the previous subsection to explain the reduced job finding. The broader search task forces unemployed

Figure 4.3: Effects of imposing the broader search task - instrumental variable estimates

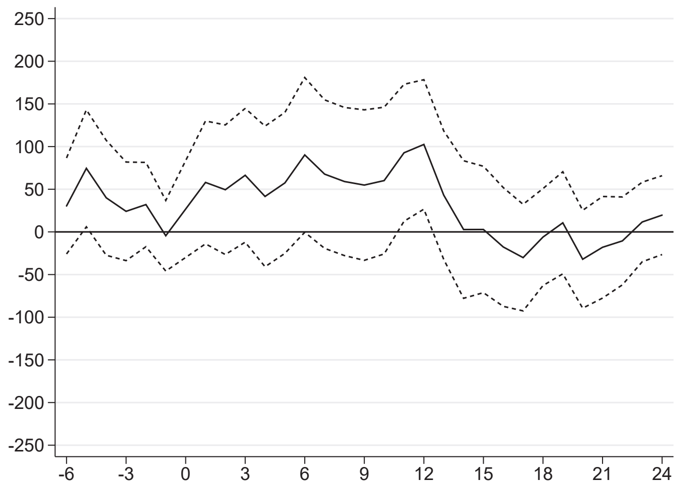
(a) UI benefits receipt



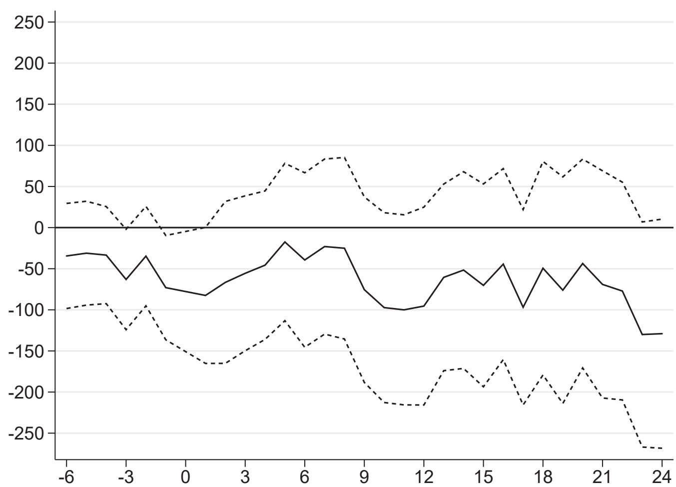
(b) Employment



(c) Amount of UI benefits



(d) Earnings



Note: Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. $N = 42,605$ for all estimated effects in all panels. Dashed lines display the 95% confidence interval based on standard errors clustered on caseworker level. $t = 0$ is the time of the caseworker meeting.

Table 4.7: Effects of imposing the broader search task on cumulative outcomes - instrumental variable estimates

Dependent variable:	UI duration	UI benefits	Employment duration	Earnings
	(1)	(2)	(3)	(4)
One year after meeting	2.27***	838**	-1.17	-671
	(0.86)	(389)	(0.79)	(515)
Dependent mean	36.96	10,728	27.93	10,958
Two years after meeting	2.70*	856	-1.65	-1,462
	(1.45)	(562)	(1.55)	(1,117)
Dependent mean	50.92	14,210	63.91	28,126
Number of workers		42,605		

Note: Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. Standard errors in parentheses are robust and clustered at the caseworker level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

workers to change their job search behavior and they can no longer search optimally. The negative effect on job finding is in line with the model specification where broader job search is a substitute for narrow job search. That different search channels are close substitutes concurs with Van den Berg and Van der Klaauw (2019). The theoretical model predicts that reservation wages decline and more often jobs will be found using broader job search. The latter is associated with lower wages. We investigate these predictions by considering the effects of imposing the broader search task on the job characteristics.[20]

Table 4.8 shows the estimated effects of the jobs that the individuals had one year after the caseworker meeting and two years after the caseworker meeting. Imposing the broader search task does not have a significant

[20]Ideally, we would also consider job application data. However, job applications in the online account are only observed for less than 40% of the individuals and often bunch at one application each week, which is the mandatory job search requirement. Also data from the online job search platform are incomplete and older applications are overwritten.

effect on the hourly wage, but two years after the meeting it significantly reduces weekly working hours and the likelihood of having a permanent contract. The effects are quite substantial, the reduction in weekly working hours due to the broader search task is more than 6% and the broader search task reduces the probability to have a permanent contract by about 21%. Permanent contracts and more working hours are indicators for better job quality. So, due to the broader search task unemployed workers may have lowered their job requirements. This concurs with the theoretical predictions. We do not find evidence that due to the broader search task unemployed workers are more likely to work in a different sector than before they became unemployed or that the commuting distance to their job is larger.

Table 4.8: Effect of broader search task on job characteristics - instrumental variable estimates

Dependent variable:	Hourly wage	Weekly hours	Permanent contract	Different sector	Distance > 20km
	(1)	(2)	(3)	(4)	(5)
One year after meeting	-0.688	-0.687	0.012	-0.027	-0.021
	(3.963)	(0.732)	(0.023)	(0.022)	(0.036)
Dependent mean	17.45	26.93	0.23	0.80	0.51
Number of workers			27,358		26,088
Two years after meeting	-0.302	-1.800***	-0.082***	-0.002	0.010
	(0.843)	(0.688)	(0.029)	(0.024)	(0.034)
Dependent mean	17.37	28.74	0.39	0.81	0.49
Number of workers			30,249		29,235

Note: All estimations are conditional on employment. Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. Regressions of distance have fewer observations because the postal code of individuals and jobs are sometimes missing. Standard errors are robust and clustered at the caseworker level. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

The effects of the broader search task may differ between individuals. Table 4.16 in Appendix 4.C shows the estimated effects for different groups of benefits recipients on the UI benefits duration and the employment duration. The estimated effect on the UI benefits duration is slightly higher for females than for males, for lower and middle educated than for higher educated and for older individuals than for younger individuals. But differences between groups are often not significant.

In the estimations above we used the leave-out mean to compute stringency of the caseworker. The robustness of the estimated effects of the broader search task is assessed by applying the sample-split approach and the reverse-sampling approach. The estimates in Table 4.13 in Appendix 4.B and Table 4.17 in Appendix 4.C show that both alternative approaches give very similar estimated effects of the broader search task.[21] In Subsection 4.4.2 we stated that the exclusion restriction requires that caseworker stringency does not affect outcomes other than via the broader search task. We earlier showed that other possible support offered by the caseworker is orthogonal to caseworker stringency. We now explore this further by including caseworker policy use as additional regressors to the empirical model. Table 4.15 in Appendix 4.B shows that all estimated effects are very robust against including these additional regressors. These results do not provide any indication that the exclusion restriction may be violated.

Our empirical findings do not concur with Belot et al. (2019), who stress that unemployed workers may benefit from broader job search. An important difference is that in their experiment unemployed workers can decide themselves whether they want to apply on the broader vacancies which are randomly provided. In our setting the broader search is a formal policy and unemployed workers are obliged to comply. The mandatory nature of our policy may also explain why our results differ from Altmann et al. (2018) and Skandalis (2019), who consider information provision on alternative job search strategies. In our setting, also unemployed workers without biased beliefs on their labor market prospects have to change their search behavior to comply to the broader search task. The theoretical predictions of Belot et al. (2019) and Moscarini (2001) suggest that broadening

[21] Table 4.14 in Appendix 4.B shows the robustness of our results to the period effects. If we include quarters instead of month fixed effects, the results remain unaffected.

the search behavior is only beneficial for more disadvantaged workers (e.g. non-specialized and long-term unemployed workers).

4.5.3 Marginal treatment effects

The rate at which caseworkers impose the broader search task differs substantially between caseworkers. The caseworker stringency states which share of a random subsample of unemployed workers would receive the broader search task from the caseworker. If monotonicity holds, then unemployed workers could be ranked by their propensity to receive the broader search task. Unemployed workers with a high propensity would receive a broader search task from all caseworkers, while unemployed workers with a low propensity would only receive a broader search task from the most strict caseworkers. In that case, our empirical analysis provides a mixture of marginal treatment effects (Carneiro et al. 2010, Heckman and Vytlacil 2001). Below we provide an analysis of the marginal treatment effects (MTEs) to study if treatment effects differ between unemployed workers who are very likely and very unlikely to receive a broader search task.

We define z as the inverse propensity that an unemployed workers receives the broader search task. The variable z is uniformly distributed within the sample of unemployed workers. When an unemployed worker is assigned to a caseworker with stringency Z, a broader search task is imposed if $z < Z$. We use a polynomial $\delta_0 + \delta_1 z + \delta_2 z^2$ for the treatment effect for an unemployed worker with characteristic z. The key problem is that z is unobserved. However, we observed the caseworker stringency $Z_{j(i)c}$ faced by unemployed worker i in local office c. If this caseworker receives a broader search task, then z is uniformly drawn from 0 to $Z_{j(i)c}$, so the expected treatment effect δ_{ic} equals

$$\delta_{ic} = \frac{1}{Z_{j(i)c}} \int_0^{Z_{j(i)c}} \delta_0 + \delta_1 z + \delta_2 z^2 dz = \delta_0 + \delta_1 \frac{Z_{j(i)c}}{2} + \delta_2 \frac{Z_{j(i)c}^2}{3}$$

Therefore, we extend the second-stage regression to

$$Y_{ic} = \alpha_c + \delta_0 B_{ic} + \delta_1 \frac{Z_{j(i)c} B_{ic}}{2} + \delta_2 \frac{Z^2_{j(i)c} B_{ic}}{3} + X_i' \beta + \varepsilon_{ic} \tag{4.5}$$

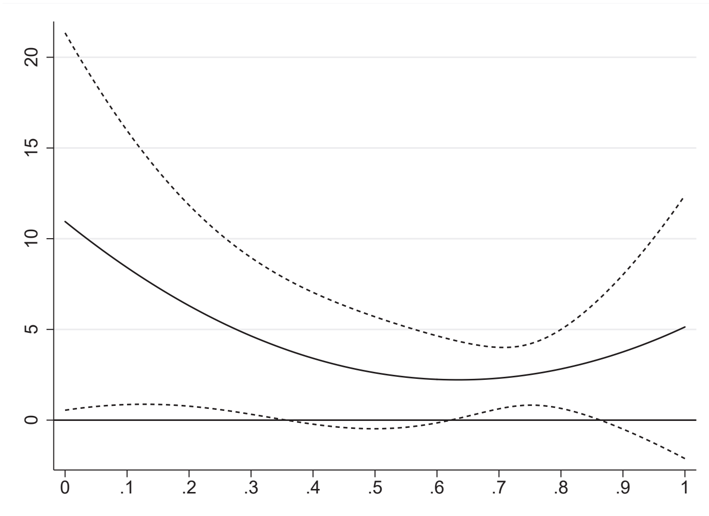
Figure 4.4 shows the estimated MTEs for the cumulative labor market outcomes one year after the caseworker meeting.[22] The effects of the broader search task are most adverse for individuals with a low value for z. Recall that a low level of z implies that the broader search task will be imposed by all caseworkers. The adverse effects of the broader search task are smallest for unemployed workers with z is about 0.65. But labor market outcomes of these individuals still do not improve after receiving the broader search task. The conclusion is thus that no unemployed worker benefits from receiving a broader search task, but the targeting of caseworkers makes the average effect worse. However, only for cumulative earnings the marginal treatment effects are significantly different from a homogeneous treatment effect (see the coefficient estimates in Table 4.18 in Appendix 4.C).

A possible explanation is that caseworkers target the broader search task towards unemployed workers who devote the least job search effort toward broader search. The reason why these individuals mainly search narrowly is that this yields the highest returns for them. These are what Moscarini (2001) refers to as the specialized workers. If they substitute broader job search for narrow job search, their labor market outcomes become substantially worse.

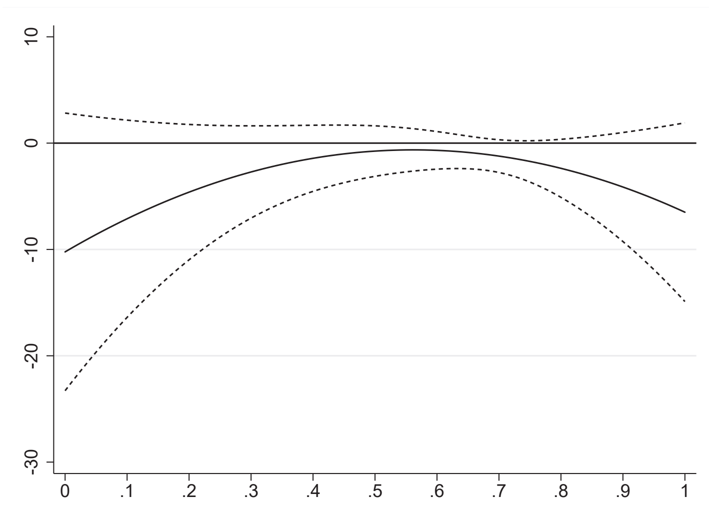
[22]The same figures with the estimated MTEs for the cumulative outcomes after two years are in Figure 4.6 in Appendix 4.C. These results yield similar conclusions.

Figure 4.4: Marginal treatment effects of imposing the broader search task on cumulative outcomes after one year

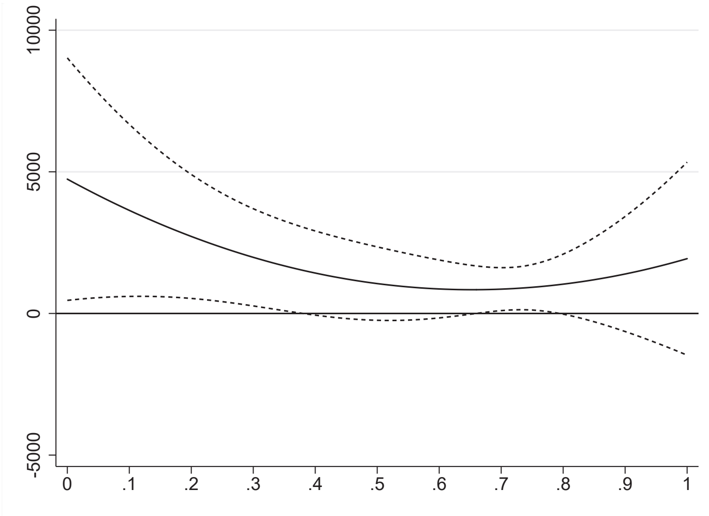
(a) UI benefits receipt



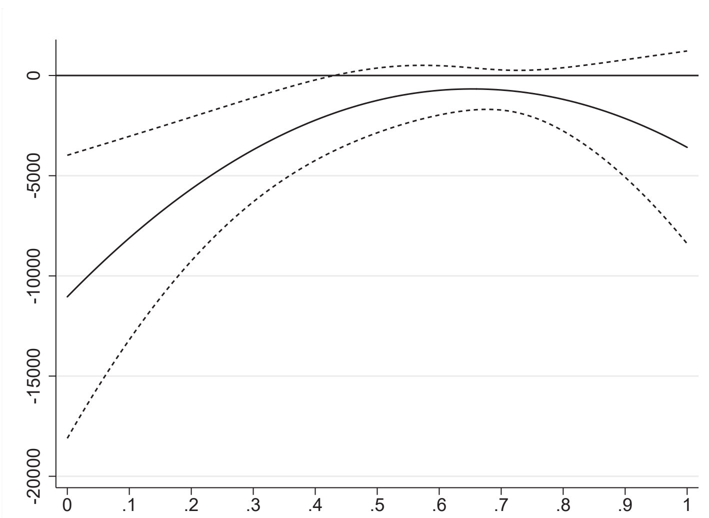
(b) Employment



(c) Amount of UI benefits



(d) Earnings



Note: The horizontal axis displays the unobserved resistance to the broader search task. Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. Dashed lines display the 95% confidence interval based on standard errors clustered on caseworker level.

4.5.4 Decomposing the effects of caseworker meetings

The broader search program has positive effects on labor market outcomes. The program consists of a caseworker meeting and possibly a broader search task. Imposing the broader search task has negative effects on labor market outcomes. The program effect is estimated for the compliers to the experiment, while the effect for the broader search task is estimated for unemployed workers in the treatment group who attended a caseworker meeting. The latter includes both compliers and always takers, which complicates a direct comparison of effects. Below, we decompose the program effects in an effect of the caseworker meeting and an effect of the broader search task.

To compare the estimated effects of participating in the broader search program to the estimated effect of imposing the broader search task, it is necessary to make the populations comparable. Let ϕ_C and ϕ_A be the sizes of populations of respectively the compliers and always takers in the randomized experiment. We define Y^{11} as the potential outcome after attending a caseworker meeting in which a broader search task has been imposed and Y^{10} as the potential outcome after attending a caseworker meeting in which the caseworker meeting has not been imposed. Following Imbens and Rubin (1997), we write

$$\mathrm{E}[Y^{11} - Y^{10}|C] = \frac{1}{\phi_C}\Big((\phi_C + \phi_A)\mathrm{E}[Y^{11} - Y^{10}|C \vee A] - \phi_A \mathrm{E}[Y^{11} - Y^{10}|A]\Big)$$

The population shares ϕ_C and ϕ_A are observed in the randomized experiment and equal 0.366 and 0.259, respectively. The treatment effect $\mathrm{E}[Y^{11} - Y^{10}|C \vee A]$ is estimated using the sample of program participants in the treatment group, which coincides with the results in Subsection 4.5.2. We can estimate $\mathrm{E}[Y^{11} - Y^{10}|A]$ using the program participants in the control group.[23] The estimates on cumulative outcomes for the

[23]The caseworker stringency is estimated using the treatment group in the randomized experiment. Next, we pool the data on the unemployed workers with a caseworker meeting in the treatment and control group to get a precise estimate for the treatment effects for the always takers in the control group. This is necessary because there are too few unemployed working with a caseworker meeting in the control group to estimate all local office and time fixed effects. Pooling the treatment and control group allows us to estimate the average treatment effect for the compliers directly.

subsequent estimates for $\mathrm{E}[Y^{11} - Y^{10}|C]$ and $\mathrm{E}[Y^{11} - Y^{10}|A]$ are shown in panels A and B of Table 4.19 in Appendix 4.D. The effects of the broader search task are more adverse for the always takers than for the compliers to the randomized experiment.

We define Y^1 as the potential outcome of participating in the broader search program and Y^0 as the potential outcome of not participating. The evaluation of the randomized experiment provides an average treatment effect for the compliers to the random assignment, so $\mathrm{E}[Y^1 - Y^0|C]$. The potential treated outcome can be decomposed in a potential outcome Y^{11} with the broader search task and a potential outcome Y^{10} without the broader search task. For ease of simplicity we assume that the effects of the broader search task do not vary by the propensity to receive the task. Accordingly, we can write $Y^1 = pY^{11} + (1-p)Y^{10}$ where p is the propensity to be assigned the broader search task. Recall that the marginal treatment effects did not provide strong evidence that treatment effects vary with the propensity to receive the task.[24] Then the effect of only the caseworker meeting without the broader search task becomes

$$\mathrm{E}[Y^{10} - Y^0|C] = \mathrm{E}[Y^1 - Y^0|C] - p\mathrm{E}[Y^{11} - Y^{10}|C]$$

The probability p to be assigned to the broader search task should apply to the compliers in the experiment. So this means that $p = \frac{1}{\phi_C}((\phi_C + \phi_A)\bar{Z}_1 - \phi_A\bar{Z}_0)$, where $\bar{Z}_1$ and $\bar{Z}_0$ are the rates of applying the broader search task in the treatment and control group of the randomized experiment. The rates of applying the broader search task $\bar{Z}_1$ and $\bar{Z}_0$ follow from Table 4.1 and equal 0.691 en 0.174, respectively, which implies $p = 1$.[25]

The results of the decomposition are shown in Table 4.9. The program and the broader search task have opposite effects. Therefore, the caseworker meeting is very effective. The caseworker meeting shortens the period of collecting UI with more than three weeks, which reduces UI benefits payments with about €1,500. These effects are in agreement with

[24] Allowing the effects of the broader search task to depend on the propensity to receive the task, would require to estimate the distribution of the propensity to receive the task among the compliers from the distributions in the control and treatment group and to also compare marginal treatment effects of both treatment groups.

[25] Filling in all fractions actually gives $p = 1.06$.

Table 4.9: Effects of the program, the broader search task and the caseworker meeting

Dependent variable:	Weeks of collecting UI	UI Benefits	Weeks of employment	Earnings
	(1)	(2)	(3)	(4)
One year after meeting				
Program: $E[Y^1 - Y^0 \vert C]$	-1.41	-879	0.90	379
Broader search task: $E[Y^{11} - Y^{10} \vert C]$	1.77	622	-0.57	-444
Caseworker meeting: $E[Y^{10} - Y^0 \vert C]$	-3.18	-1,501	1.48	823
Two years after meeting				
Program: $E[Y^1 - Y^0 \vert C]$	-1.84	-1,202	1.15	265
Broader search task: $E[Y^{11} - Y^{10} \vert C]$	1.67	351	0.61	-413
Caseworker meeting: $E[Y^{10} - Y^0 \vert C]$	-3.51	-1,553	0.54	678

Note: The standard errors of the program effects and the broader search task effects for the compliers are in Table 5.6 and Table 4.19 in Appendix 4.D, respectively.

(e.g. Card et al. 2010, Maibom et al. 2017, Schiprowski 2020), who also find strong effects of caseworker meetings. The effect on weeks of employment and earnings is less strong. So workers do not fully compensate the reduced UI benefits with additional earnings.

4.6 Conclusion

The paper evaluates a broader search program for unemployed workers in the Netherlands. Results from a field experiment show that participation in the program stimulates the exit from unemployment. However, the reduced benefits payments are not fully compensated with additional earnings from work. The broader search program starts with a caseworker meeting, which takes place at the moment that unemployed workers should expand their job search to jobs that not necessarily match their skills and requirements. If during the meeting the caseworker assesses the past job search as too narrow, the caseworker can give the unemployed worker a mandatory task to apply for a broader set of vacancies. We exploit that the rate of giving this task differs between caseworkers and that unemployed workers are randomly assigned to a caseworker. Our estimation results show that the broader search task has a negative effect on labor market outcomes. Our decomposition analysis, therefore, shows a positive effect of the caseworker meeting, which is in agreement with e.g. Card et al. (2010), Schiprowski (2020) and Maibom et al. (2017).

The adverse effects of the broader search task seem in contradiction with Altmann et al. (2018) and Belot et al. (2019), who show positive effects of encouraging broader search. While our study considers recipients of unemployment insurance benefits, Belot et al. (2019) focus on more disadvantaged unemployed workers. Moscarini (2001) argues that broader search is more beneficial for workers with weak comparable advantages, such as the more disadvantaged workers studied by Belot et al. (2019). Furthermore, Altmann et al. (2018), Belot et al. (2019) consider an encouragement to search more broadly, while our broader search task is mandatory. An encouragement may mainly affect unemployed workers with biased belief about the returns to job search, while a mandatory program also affects unemployed who already optimize their search effort. Our job search model shows adverse effects of mandatory broader search for specialized workers who benefit more from narrow search than broader search. Since caseworkers direct broader search tasks to unemployed workers who are mainly searching narrowly, specialized workers are most likely to receive the broader search task. Our marginal treatment effect estimates suggest that the adverse effects of the broader search task are highest for unemployed workers that are most likely to receive the task.

Our findings show that evaluations from an encouragement or information treatment are not easily translated in a (low-cost) active labor market program. Active labor market programs are often more mandatory, which implies that the treated population is larger than the compliers to an encouragement or information treatment. In particular, caseworkers will target the program to a different group than the respondents to, for example, a brochure on the advantages of broader search as is studied by Altmann et al. (2018). To some extent we could interpret the caseworker meetings without the mandatory task as more similar to the information treatments. If no task is given to the unemployed worker, the meeting still focuses on broader search but is advisory and informative. The results from our decomposition show strong positive effects of the caseworker meeting without task that are in line with encouragement and information treatments on broader search.

4.A Back-of-the-envelope costs-benefits analysis of the broader search program

For the UI administration the average costs of the broader search program are €169 per invited unemployed worker.[26] To compare these costs to the benefits for the UI administration, we consider the reduction in cumulative UI benefits payments for each invited unemployment workers. Therefore, we consider the intention-to-treat effect, which is estimated using the following regression equation

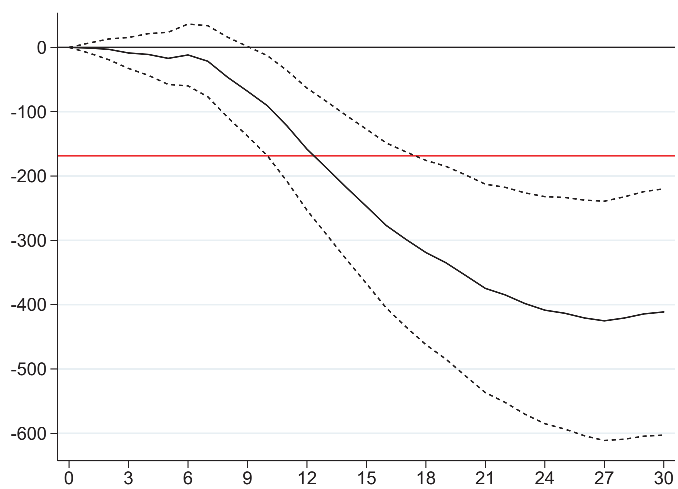
$$Y_i = \alpha + \delta T_i + X_i' \beta + \varepsilon_i$$

Since this specification regresses the outcome on the random assignment, it can be estimated using OLS. Further, we assume an annual discount rate of 5 percent.

The results are summarized in the Figure 4.5. The break-even moment is after 13 months of collecting UI benefits (roughly 6 months after the meeting with the caseworker). In the long run the program is cost effective for the UI administration since the reduction in UI benefits equals about €400 per invited worker.

[26]The average costs per treated worker is calculated by dividing the total lump-sum costs of the experiment (€20 million) by the number of workers in the treatment group ($N = 118,697$).

Figure 4.5: Intention to treat effects of the broader search program on cumulative UI benefits - OLS estimates



Note: The black line displays the estimated intention to treat effects. The estimated effects are based on regressions controlling for age, gender, nationality, education, previous wage, sector, working hours and UI benefits eligibility. Dotted lines display the 95% confidence interval, based on robust standard errors. The red line displays the average costs per invited worker. $t = 0$ is the start of collecting UI, the broader search program starts with a caseworker meeting in the seventh month of UI.

Robustness of the effects of the broader search task 4.B

Table 4.10: First-stage estimates using different sample selections on caseworkers and different controls

	Worker per caseworker				
Sample selection	40-400	50-400†	60-400	50-300	50-500
	(1)	(2)	(3)	(4)	(5)
Panel A. No controls					
Caseworker stringency	0.820***	0.834***	0.842***	0.836***	0.832***
	(0.018)	(0.020)	(0.023)	(0.020)	(0.020)
F-stat. (Instrument)	2,104	1,782	1,395	1,804	1,754
Panel B. Add demographic controls					
Caseworker stringency	0.820***	0.830***	0.836***	0.832***	0.827***
	(0.018)	(0.020)	(0.023)	(0.020)	(0.020)
F-stat. (Instrument)	2,082	1,763	1,377	1,784	1,735
Panel C. Add labor market history controls					
Caseworker stringency	0.820***	0.826***	0.834***	0.828***	0.824***
	(0.018)	(0.020)	(0.022)	(0.019)	(0.020)
F-stat. (Instrument)	2,076	1,778	1,388	1,811	1,739

Note: †The baseline analysis uses caseworkers meeting 50-400 benefits recipients. All regressions include local office fixed effects interacted with month fixed effect. The demographic controls are age, gender, nationality and education. The labor market history controls are previous wage, sector, working hours and UI benefits eligibility. Standard errors are robust and clustered at the caseworker level. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

Table 4.11: First-stage estimates using different sample selections on caseworkers and different controls - split-sample approach

Sample selection	40-400	50-400	60-400	50-300	50-500
	(1)	(2)	(3)	(4)	(5)
Panel A. No controls					
Caseworker stringency	0.718***	0.745***	0.774***	0.751***	0.743***
	(0.033)	(0.038)	(0.043)	(0.038)	(0.038)
F-stat. (Instrument)	468	386	328	393	392
Panel B. Add demographic controls					
Caseworker stringency	0.715***	0.742***	0.768***	0.748***	0.739***
	(0.033)	(0.038)	(0.043)	(0.038)	(0.038)
F-stat. (Instrument)	463	383	323	390	388
Panel C. Add labor market history controls					
Caseworker stringency	0.709***	0.733***	0.762***	0.740***	0.731***
	(0.034)	(0.039)	(0.044)	(0.038)	(0.038)
F-stat. (Instrument)	444	363	299	373	367

Note: All regressions include local office fixed effects interacted with month fixed effect. The demographic controls are age, gender, nationality and education. The labor market history controls are previous wage, sector, working hours and UI benefits eligibility. Standard errors are robust and clustered at the caseworker level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 4.12: First-stage estimates for different groups of benefits recipients - reverse-sample approach

	Coefficient	S.e.	F-stat	N	Dependent Mean
	(1)	(2)	(3)	(4)	(5)
Gender					
Female	0.874***	(0.086)	103	11,491	0.769
Male	0.750***	(0.051)	220	12,282	0.720
Nationality					
Native	–			476	0.742
Non-native	0.985***	(0.127)	61	1,608	0.669
Educational level					
Low educated	0.876***	(0.061)	209	6,167	0.682
Middle educated	0.777***	(0.077)	102	11,643	0.749
High Educated	0.771***	(0.057)	182	11,194	0.768
Age					
Younger than 40	0.732***	(0.066)	124	12,046	0.730
Older than 40	0.757***	(0.075)	103	12,454	0.771
Employment status					
Not employed at 6 months	0.343***	(0.076)	20	10,480	0.778
Employed at 6 months	0.836***	(0.068)	153	7,977	0.619

Note: Regressions include with the exception of the control for the relevant group controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. Standard errors are robust and clustered at the caseworker level. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

Table 4.13: Effects of imposing the broader search task on cumulative outcomes - split-sample approach

Dependent variable:	UI duration	UI benefits	Employment duration	Earnings
	(1)	(2)	(3)	(4)
One year after meeting	3.11***	879*	-1.62	-621
	(1.20)	(534)	(1.30)	(720)
Dependent mean	36.89	10,710	27.94	11,007
Two years after meeting	3.63*	846	-2.46	-1,312
	(2.12)	(787)	(2.48)	(1,608)
Dependent mean	50.81	14,210	63.87	28,173
Number of workers		21,342		

Note: Instrumental variable regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. Standard errors in parentheses are robust and clustered at the caseworker level. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

Table 4.14: Effects of imposing the broader search task on cumulative outcomes – instrumental variable estimates with quarter fixed effects

Dependent variable:	UI duration	UI benefits	Employment duration	Earnings
	(1)	(2)	(3)	(4)
One year after meeting	2.21**	905**	-1.25	-713
	(0.88)	(391)	(0.79)	(517)
Dependent mean	36.96	10,728	27.93	10,958
Two years after meeting	2.57*	940*	-1.76	-1,519
	(1.47)	(560)	(1.56)	(1,120)
Dependent mean	50.92	14,210	63.91	28,126
Number of workers		42,605		

Note: Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with quarter fixed effect. Standard errors in parentheses are robust and clustered at the caseworker level. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

Table 4.15: Effects of imposing the broader search task on cumulative outcomes – instrumental variable estimates with additional controls for caseworker policy choices

Dependent variable:	UI duration	UI benefits	Employment duration	Earnings
	(1)	(2)	(3)	(4)
One year after meeting	2.27***	849**	-1.22	-676
	(0.88)	(393)	(0.79)	(518)
Dependent mean	36.96	10,728	27.93	10,958
Two years after meeting	2.66*	876	-1.75	-1,487
	(1.48)	(573)	(1.56)	(1,127)
Dependent mean	50.92	14,210	63.91	28,126
Number of workers		42,605		

Note: Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. The additional controls for other policy tools used by caseworkers are participation in workshops, sanctions and search exemptions, which are computed using leave-out means. Standard errors in parentheses are robust and clustered at the caseworker level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Heterogeneous effects of the broader search task 4.C

Table 4.16: Effects of imposing the broader search task on cumulative outcomes after one year for different demographic groups - instrumental variable estimates

Dependent variable:	UI duration	Employment duration
	(1)	(2)
A. GENDER:		
1. Female		
Estimate	2.69***	-1.09
(s.e.)	(0.91)	(1.22)
Dependent mean	37.62	28.23
Number of workers	25,207	25,207
2. Male		
Estimate	1.73	-1.11
(s.e.)	(1.16)	(1.52)
Dependent mean	36.00	27.49
Number of workers	17,398	17,398
B. EDUCATIONAL LEVEL:		
1. Low educated		
Estimate	2.40	-0.62
(s.e.)	(1.72)	(2.26)
Dependent mean	36.90	27.51
Number of workers	7,063	7,063
2. Middle educated		
Estimate	3.04***	-3.13**
(s.e.)	(1.00)	(1.32)
Dependent mean	36.97	28.97
Number of workers	22,156	22,156
3. High Educated		
Estimate	1.16	1.75
(s.e.)	(1.33)	(1.77)
Dependent mean	36.97	26.42
Number of workers	13,386	13,386
C. AGE:		
1. Younger than 40		
Estimate	1.55*	-1.87
(s.e.)	(0.91)	(1.31)
Dependent mean	32.67	29.70
Number of workers	20,323	20,323
2. Older than 40		
Estimate	2.35**	-0.16
(s.e.)	(1.07)	(1.37)
Dependent mean	40.87	26.31
Number of workers	22,282	22,282

Note: Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. Standard errors in parentheses are robust and clustered at the caseworker level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

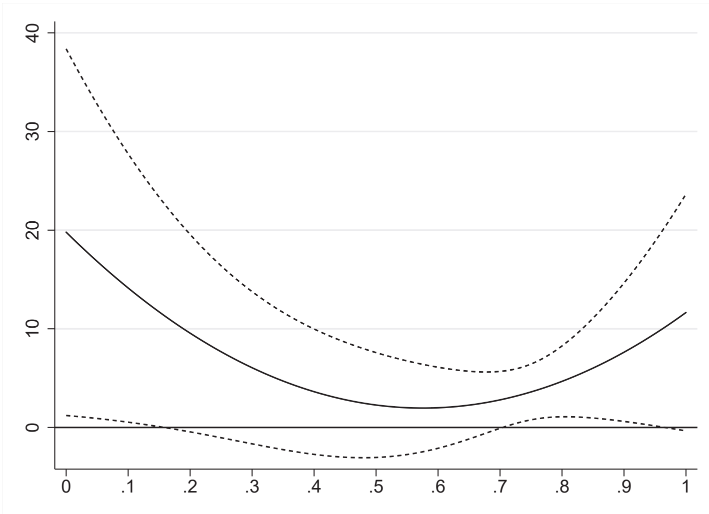
Table 4.17: Effects of imposing the broader search task on cumulative outcomes after one year for different demographic groups - reverse-sample instrumental variable estimates

Dependent variable:	UI duration	Employment duration
	(1)	(2)
A. GENDER:		
1. Female		
Estimate	4.49***	-2.14
(s.e.)	(1.52)	(1.70)
Dependent mean	37.37	27.38
Number of workers	11,678	11,678
2. Male		
Estimate	-0.30	-0.01
(s.e.)	(1.46)	(1.70)
Dependent mean	36.17	27.10
Number of workers	12,607	12,607
B. EDUCATIONAL LEVEL:		
1. Low educated		
Estimate	2.24*	1.80
(s.e.)	(1.90)	(3.29)
Dependent mean	36.96	27.40
Number of workers	6,374	6,374
2. Middle educated		
Estimate	2.40*	0.88
(s.e.)	(1.44)	(2.26)
Dependent mean	36.91	28.47
Number of workers	11,875	11,875
3. High Educated		
Estimate	0.98	-0.70
(s.e.)	(1.54)	(1.84)
Dependent mean	36.79	26.16
Number of workers	11,441	11,441
C. AGE:		
1. Younger than 40		
Estimate	1.17	0.41
(s.e.)	(1.85)	(2.63)
Dependent mean	32.51	29.25
Number of workers	12,283	12,283
2. Older than 40		
Estimate	3.61**	-3.21
(s.e.)	(1.68)	(2.23)
Dependent mean	41.00	25.40
Number of workers	12,706	12,706

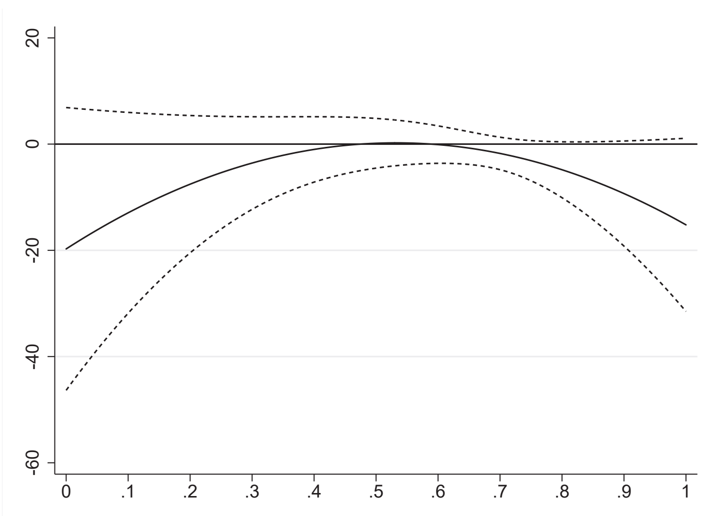
Note: Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. Standard errors in parentheses are robust and clustered at the caseworker level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Figure 4.6: Marginal treatment effects of imposing the broader search task on cumulative outcomes after 2 years

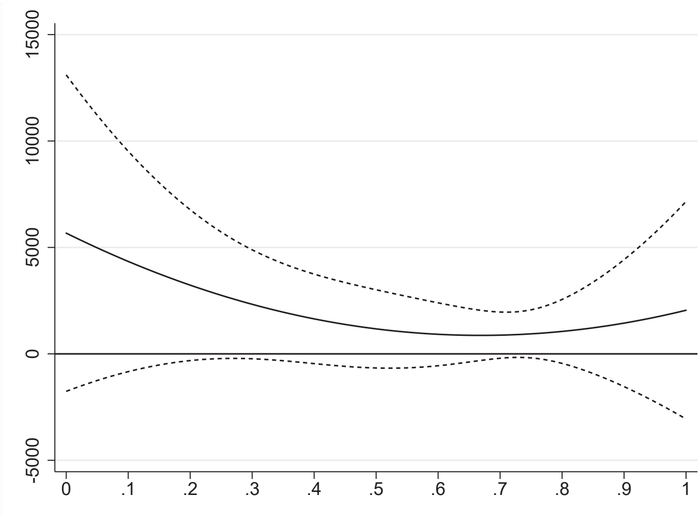
(a) UI benefits receipt



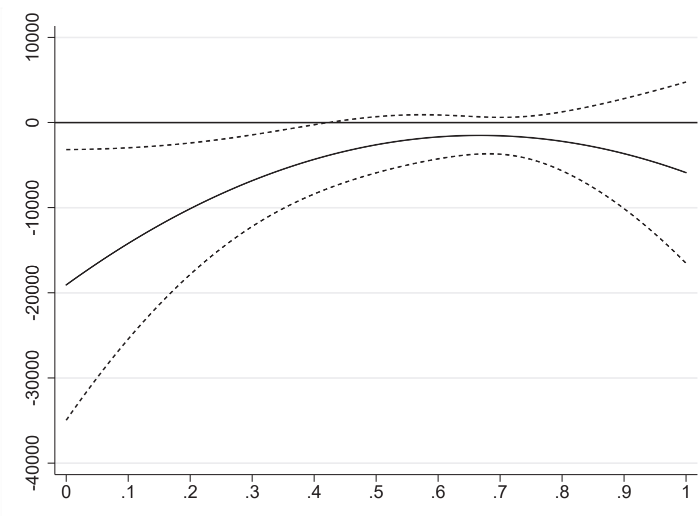
(b) Employment



(c) Amount of UI benefits



(d) Earnings



Note: Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. Dashed lines display the 95% confidence interval based on standard errors clustered on caseworker level.

Table 4.18: Marginal treatment effects coefficients of imposing the broader search task on cumulative outcomes - instrumental variable estimates

Dependent variable:	UI duration	UI benefits	Employment duration	Earnings
	(1)	(2)	(3)	(4)
Panel A: Linear specification				
One year after meeting				
δ_0	4.45	2,006	-1.11	-3,774*
	(3.69)	(1,583)	(3.75)	(2,208)
δ_1	-3.09	-1,661	-0.08	4,418
	(4.85)	(2,154)	(5.53)	(3,146)
Two years after meeting				
δ_0	3.69	2,453	1.37	-7.253
	(6.43)	(2,297)	(7.39)	(4,601)
δ_1	-1.41	-2,274	-4.30	8,245
	(8.50)	(3,167)	(10.82)	(6,681)
Panel B: Quadratic specification				
One year after meeting				
δ_0	10.95**	4,742**	-10.23	-11,042***
	(5.31)	(2,184)	(6.66)	(3,603)
δ_1	-27.53	-11,943	34.19	31,741**
	(18.84)	(8,255)	(22.82)	(12,403)
δ_2	21.72	9,136	-30.45	-24,283**
	(17.69)	(7,961)	(20.40)	(11,343)
P-value joint test	0.26	0.19	0.32	0.01
Two years after meeting				
δ_0	19.80**	5,673	-19.73	-19,064**
	(9.48)	(3,792)	(13.58)	(8,110)
δ_1	-61.95*	-14,370	75.04	52,647*
	(32.56)	(14,109)	(46.17)	(28,516)
δ_2	53.80*	10,747	-70.51*	-39,461
	(29.96)	(13,034)	(40.77)	(25,873)
P-value joint test	0.16	0.43	0.22	0.09
Number of workers		42,605		

Note: All outcome variables are measured 1 year after the caseworker meeting. Regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. Standard errors in parentheses are robust and clustered at the caseworker level. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

4.D Decomposition of the effects for the broader search program

Table 4.19: Effects of imposing the broader search task on cumulative outcomes for compliers and always takers - instrumental variable estimates

Dependent variable:	UI duration	UI benefits	Employment duration	Earnings
	(1)	(2)	(3)	(4)
Panel A: E$[Y^{11} - Y^{10} \mid C]$				
One year after meeting	1.77*	622	-0.57	-444
	(0.91)	(417)	(0.94)	(580)
Dependent mean	36.96	10,728	27.67	10,958
Two years after meeting	1.67	351	0.61	-413
	(1.54)	(605)	(1.80)	(1,278)
Dependent mean	50.92	14,210	62.61	28,126
Number of workers		42,605		
Panel B: E$[Y^{11} - Y^{10} \mid A]$				
One year after meeting	2.83***	1,059**	-2.39**	-1,150*
	(0.99)	(460)	(1.07)	(662)
Dependent mean	37.40	10,500	27.67	10,688
Two years after meeting	3.85**	1,375*	-4.92**	-3,348**
	(1.72)	(702)	(2.10)	(1,450)
Dependent mean	51.47	13,972	62.61	27,032
Number of workers		1,477		

Note: Pooled regressions include controls for age, gender, nationality, education, previous wage, sector and working hours, UI benefits eligibility, and local office fixed effects interacted with month fixed effect. Standard errors in parentheses are robust and clustered at the caseworker level. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

5 The effects of application processing times and prepayments on welfare receipt and employment

Abstract

This paper investigates the benefit and employment effects of application processing times and benefit prepayments for welfare applicants. For causal inference we exploit exogenous variation in application processing times and the use of provisional benefit prepayments among randomly assigned caseworkers. We find that processing times mostly affect the timing of benefit receipt and employment of the applicants, but to a lesser extent their cumulative outcomes. The absence of larger effects can be explained by two opposing effects of longer processing times: (i) some applicants may withdraw their welfare application, while (ii) other applicants who receive welfare benefits after longer processing times may result in a longer-term reliance on welfare. Our results also suggest that providing benefit prepayments promotes the future employment outcomes of approved applicants who experience longer processing times.

Introduction 5.1

One of the key trade-offs in the optimal design of social security programs is that of ensuring income security and preserving work incentives for

This chapter is co-authored by Ernst-Jan de Bruijn, Marike Knoef and Pierre Koning. W&I Rotterdam is gratefully acknowledged for providing the data that are used in this paper. The authors are grateful to Lieke Kools for her help with the code. The authors thank Jim Been, Guido Imbens, Bas van der Klaauw and seminar participants in Leiden for useful comments to presentations and earlier drafts of the paper. This research is sponsored by Instituut Gak.

those able to work. In this setting, application costs for benefits can play an important role. Formal eligibility conditions, together with informal and administrative barriers – such as long processing times – raise application costs, which may lead to increased self-screening and improved targeting of programs (Kleven and Kopczuk 2011, Nichols and Zeckhauser 1982). However, there is also evidence pointing at increases in non-take-up of benefits among individuals in greatest need (Currie 2006, Deshpande and Li 2019, Finkelstein and Notowidigdo 2019, Ko and Moffitt 2022). Whether reductions in take-up are welfare improving depends crucially on the design of the program, its consequences for self-screening and the quality of the award decision.

To address these unintended consequences while still ensuring timely provision of income, social insurance programs often use provisional benefit prepayments. While such prepayments can facilitate consumption smoothing, they may also result in substantial repayments for those ultimately found ineligible. Depending on the length of the application process, this can cause significant financial stress for applicants.[1] The combined use of long application times – with more adequate eligibility assessments – and benefit prepayments may thus contribute to targeting and income smoothing, but also inhibits the risk of later financial stress for denied applicants.

This paper investigates the impact of welfare application processing times and benefit prepayments on the welfare receipt and employment of applicants. We focus on welfare in the Netherlands, where benefits serve as a social safety net for all households with insufficient means of subsistence. Welfare applicants need to provide caseworkers with detailed information on their living situation, income and wealth to establish their eligibility. Although there are formal standards for the maximum application processing time, the actual length and process of collecting and assessing information can vary substantially across individuals and caseworkers. During the application process, caseworkers may opt to

[1]The Covid-19 pandemic has led to substantial expansions of income support programs for which overpayments and subsequent repayments are relevant. Examples include the Universal Credit system in the United Kingdom and the Unemployment Compensation program and Child Tax Credit in the United States (DWP 2022, GAO 2022, NCSL 2023).

provide benefit prepayments. These prepayments are reconciled upon determination of the applicant's eligibility for welfare benefits, but must be returned if the applicant is found to be ineligible.

Application processing times can have opposing effects on the benefit and labor market outcomes of applicants. On the one hand, longer processing times may decrease the take-up of welfare benefits and increase employment. Due to the uncertainty of the award decision, applicants may opt to take up employment and withdraw their application while it is still being processed. At the same time, the complexity of the application process can be a significant and unforeseen burden for applicants, which may discourage them from completing their application. These effects may accumulate over time, with applicants becoming more likely to suffer from liquidity constraints (if benefit prepayments are absent). Applicants may therefore decide to increase job search efforts or accept lower wages. On the other hand, longer processing times may lead to higher levels of inactivity pending the award decision, which reduces the applicants' employment or earnings potential. As long as no benefit decision is made, the job search of applicants is often not monitored and welfare to work services are not started. Applicants may also assume that work resumption implies the potential loss of accumulated welfare benefits during the elapsed application time so far. Additionally, lower levels of job search activities may stem from increased financial stress and individuals may lose human capital. Overall, the effect of application processing times is therefore ambiguous.

As prepayments alleviate the impact of a temporary loss of income, applications with prepayments may be perceived as less burdensome. They provide more income security, and therefore reduce the risk of liquidity constraints. On the one hand, this may lower the need to return to work. On the other hand, less financial stress while awaiting the benefit decision may increase the quality of job search.

For the empirical analysis we use welfare application data from Rotterdam, combined with administrative records from Statistics Netherlands. Rotterdam is the second-largest city in the Netherlands and has relatively the highest number of inhabitants on welfare. We observe substantial variation in the application processing times between applicants, partly due to differences in processing speed among the caseworkers who act as adjudicators. To measure the causal effect of the length of the application processing times, we exploit variation among caseworkers and the random assignment of applications to caseworkers. Our key identifying assumption is that the processing speed of caseworkers is orthogonal to other relevant dimensions on which the caseworkers might differ, such as the award rate. Given the use of unambiguous eligibility rules for welfare benefits with little room for caseworker discretion, this assumption is likely to hold. In a similar fashion, we exploit variation between caseworkers in the propensity to grant benefit prepayments prior to the award decision to investigate the effects of benefit prepayments.

Our approach is inspired by a growing body of literature that exploits exogenous variation in judge or caseworker stringency as an instrumental variable. Kling (2006), Aizer and Doyle Jr (2015), Doyle Jr (2007, 2008), and Bhuller et al. (2020) use judges stringency to estimate the effects of judge decisions on various socioeconomic and crime outcomes. Arni and Schiprowski (2019) and van der Klaauw and Vethaak (2022) use caseworker stringency to estimate the effects of job search requirements on unemployed workers. Maestas et al. (2013) and French and Song (2014) use the assignment to a disability examiner to demonstrate the negative effects of disability insurance benefits on employment rates. Our study is most closely aligned with the work of Autor et al. (2015), who show that longer Social Security Disability Insurance (SSDI) application processing times reduce subsequent employment and earnings. They consider a setting in which applicants face substantial processing times with strong non-work incentives, resulting in prolonged periods out of the labor force without any form of income. In contrast, our study examines a setting with on average shorter processing times in which the applicants are subject to job search requirements and frequently receive benefit prepayments to cushion the income effects of longer application times.

Our paper contributes to the existing literature on the take-up and targeting of social security programs by examining the effects of processing times on program take-up. Previous research has shown that costs associated with the take-up of these programs are effective in deterring unemployed workers from applying.[2] Formal and mandatory waiting or job search periods are commonly used to increase the costs of application and are often justified as they aim at lowering moral hazard (Nichols et al. 1971, Parsons 1991). For example, Autor et al. (2014) and Storer and Van Audenrode (1995) find reduced benefits take-up of disability and unemployment insurance, respectively. Bolhaar et al. (2019) show that mandatory job search periods substantially and permanently reduce the take-up of Dutch welfare benefits. They consider a setting in which the applicants are strongly encouraged to actively search for work during the application. This job search period delays the first benefit payment by one month for all applicants. This contrasts with our setting in which the delay stems from informal processing times, which are ex ante unknown to the applicants. Informal processing times are found to increase the application costs and lower the take-up of disability programs (Autor et al. 2015). As these costs are unintended and work incentives are often not (well) formalized, the longer processing times might not be welfare improving.[3]

Our paper also contributes to the literature on the existence of liquidity constraints, by examining the relationship between benefit prepayments and job search. In economic theory, wealth is typically assumed to have a negative impact on job search (Blundell et al. 1997, Mortensen 1986). Empirical evidence supports this hypothesis, with unemployed workers finding jobs faster if they are liquidity-constrained (Basten et al. 2014, Card et al. 2007, Chetty 2008). However, there may also be negative employment effects induced by financial stress associated with severe liquidity constraints. A growing body of research highlights the detrimental impact of financial stress on (mental) health and labor market outcomes (Dobbie and Song 2015, 2020, Gathergood 2012, Ridley et al. 2020). To the best of our knowledge, we are the first to empirically investigate the effects of benefit prepayments, which have the potential to alleviate financial stress for those awaiting a benefit award decision.

[2]Currie (2006) provides an extensive overview of the literature on the take-up of social benefits.

[3]Since longer processing times in our setting partially originate from administrative burdens and program complexity, our paper also relates to the literature on the complexity of social programs (e.g. Currie 2006, Kleven and Kopczuk 2011, Ko and Moffitt 2022).

Our main research findings can be summarized as follows. First, processing times have a substantial mechanical effect on the timing of welfare receipt, but limited accumulated effects on welfare receipt or employment. Second, the absence of long-term effects can be attributed to two opposing effects of longer processing times on the applicants. Some applicants withdraw their welfare application due to longer processing times and evidently spend less time in welfare. These applicants tend to make up for lost benefits income through increased earnings. But accepted applicants with longer processing times tend to receive welfare for longer periods than those with fast applications. Finally, our findings indicate that the use of welfare prepayments increases employment of applicants who are ultimately awarded benefits after longer processing times. This suggests that lower financial stress facilitates successful job search.

The remainder of the paper is organized as follows. In the next section we describe the Dutch welfare system, the application process, the availability of benefit prepayments and the expected effects of both processing times and benefit prepayments. Section 5.3 contains a description of the data. In Section 5.4 we provide our empirical framework to estimate the effects of short processing times and benefit prepayments. Additionally, we validate the use of caseworker speed as an instrumental variable. Section 5.5 presents the results of the analysis before Section 5.6 concludes.

5.2 Institutional background

In this section, we first describe the Dutch welfare system. In doing so, a particular interest lies in the application process and the use of prepayments to mitigate the income effects of longer application times. Next, we hypothesize about the potential effects of both the length of the application process and the receipt of prepayments on welfare and labor market outcomes.

Welfare benefits in the Netherlands 5.2.1

In the Netherlands, welfare benefits function as a safety net for all unemployed individuals with insufficient means of subsistence. To be eligible for welfare benefits one should have insufficient earnings, possess no substantial assets, have no substantial partner income, and not or no longer be entitled to other social security benefits (such as unemployment insurance (UI) or disability insurance benefits).[4] With 22 percent of welfare applicants having exhausted their UI benefits, the majority of the inflow consists of individuals with insufficient work history for UI entitlement. In addition to meeting income and asset conditions, individuals should make sufficient job search efforts.[5] Non-compliance could be sanctioned with temporary benefit reductions or benefit suspensions.

Welfare benefits in 2019 were 1,026 euros per month for single individuals (both with and without children) and 1,465 euros for couples. Additionally, households may receive housing, child, and health insurance subsidies, which are not observed in our study. With benefits amounting to about 60 percent of the median disposable income, the Dutch welfare benefits system can be considered generous compared to most other OECD countries (OECD 2018). Individuals continue to receive welfare benefits until they find employment or reach the legal retirement age. Welfare recipients are not subject to benefit reductions equivalent to 25 percent of their monthly earnings, up to a threshold of approximately 200 euros. Any additional income earned above this threshold leads to a reduction in welfare benefits of an equal amount.

[4]For applicants younger than 27 there are different rules for both the application and while receiving welfare benefits. So are younger applicants subject to a so-called 'job-search-period'. For more information on the different rules for welfare recipients aged 18 to 26 see e.g., Cammeraat et al. (2022) and Stam et al. (2020). Similarly, for older applicants there are several differences in eligibility rules and search requirements. The first relevant age threshold for older applicants is at the age of 50 when partially disabled workers can supplement their (labor) income with welfare benefits up till the level of welfare benefits.

[5]An exemption from those search requirements can be requested by parents with full custody over children younger than five and those who are deemed fully and permanently disabled.

5.2.2 The application procedure

In our analysis, we focus on Rotterdam, which is the second-largest city in the Netherlands and the city with the highest welfare benefits dependence. Welfare benefit applications are assigned to caseworkers who have 8 weeks to review the application in order to determine eligibility and, if relevant, the level of benefits. The reviewing process can be characterized as back-office work, as caseworker rarely have direct contact with the applicants. The formal application period can be extended (multiple times) by the caseworker if the applicant does not provide all necessary information, such as detailed information on income transactions (bank statements, pay slip, other benefits, tax return and alimony), residence (rent and fellow residents), assets (savings, valuables and debts), a job seeker statement, and income and/or a job seeker statement of their partner, if relevant. Based on interviews with caseworkers, especially the reported living situation and wealth of the applicant can be unclear. In these cases, additional information from the applicant is required before the application can be assessed.

Given the complete information provided by the applicant and the unambiguous eligibility rules, there is little room for discretion in the award decision. However, caseworkers may differ in processing speed and in the frequency of calls for additional information of applicants. In case the applicant is deemed eligible, the benefits will be paid retroactively from the moment of the initial application. Longer processing times, however, might result in less (experienced) monitoring of job search requirements and less welfare to work services.

In the city of Rotterdam, caseworkers who assess welfare applications are not involved in monitoring of ongoing benefit conditions and in providing welfare-to-work services. Caseworkers involved with claims assessment are so-called ‘income caseworkers’ (in Dutch: *inkomensconsulent*), whereas caseworkers involved with return to work are ‘reintegration caseworkers’ (in Dutch: *werkcoach*). The assignment of welfare applications to caseworkers is based on the current caseload of the caseworkers. As an exception to this, some (less experienced) caseworkers assess more applications from applicants who exhausted UI benefits before applying for welfare benefits. These applications require less application time and effort from the caseworker, since the source and level of previous income is known.[6]

Benefit prepayments 5.2.3

To mitigate the short-term income effect of longer application times, caseworkers may issue benefit prepayments. The aim of prepayments is to prevent applicants from facing liquidity constraints while awaiting the application decision. Three possible scenarios for the use of prepayments exist, depending on the type of applicant. First, applicants with pending applications that are likely to be eligible for benefits receive 90% of their expected benefits level in advance, starting from week five onward and ending when their application is completed. Second, applicants with liquidity constraints pending the application can request a loan to bridge the income gap, which can exceed the monthly benefits level and is paid within 8 weeks. Third, applicants with urgent liquidity constraints pending the application, such that they are unable to make ends meet, can request a modest loan specifically for groceries (not for e.g., utilities, rent or insurances). These prepayments are considerably smaller, as they amount to at most a few hundred euro and are paid within two days. The prepayments are deducted from the first benefit payments if allowed benefits and remain outstanding claims when applications are rejected.

Theoretical predictions 5.2.4

In our setting, the length of the welfare application processing time is not known ex ante by applicants. The effects of longer processing times therefore do not affect the (initial) application decision, but are relevant during the application process pending the award decision of the caseworker. The overall effects of processing times on the take-up of benefits and employment are ambiguous: longer waiting times may discourage

[6] We address this issue of selection in our empirical analysis by controlling for the exhaustion of UI benefits before the application in the first-stage regression. The subsequent results are consistent with random assignment of applications to caseworkers.

applicants to proceed with their applicants, but applicants may also show higher levels of inactivity when time proceeds.

When someone applies for welfare benefits and it takes a long time to process their application, they may resume work and retract their application. As time proceeds, applicants may become more inclined to secure income and avoid further application costs. With higher job search efforts and lower reservation wages, this may result in a decline in the take-up of welfare benefits and increased employment. When particularly applicants with better labor market prospects withdraw their applications, the targeting of welfare benefits improves. Albeit that the application process for welfare benefits is usually shorter than for disability insurance, such behavioral and targeting effects can be similar as those found by Autor et al. (2014) for SSDI benefits. Applicants may be more likely to withdraw their application during the application process when applicants become liquidity-constrained. Empirical evidence lends credence to this explanation, with unemployed workers securing jobs more quickly if they are liquidity-constrained (see e.g., Basten et al. 2014, Card et al. 2007).

Longer processing times could also increase the take-up of welfare benefits and decrease employment rates. Two channels might explain these effects. First, the welfare application process itself may directly impact the post-application outcomes, as periods of inactivity may diminish the employment potential of applicants (Parsons 1991). This effect has not only been found among disability insurance applicants (Autor et al. 2015), but also among asylum seekers who face temporal employment restrictions upon arrival in a country (see e.g., Fasani et al. 2021, Marbach et al. 2018). Obviously, this channel is less important in our study as the average length of the application is shorter. Still, we argue that the absence of monitoring of job search requirements and the absence of welfare to work services pending the application decision may lead to higher levels of inactivity during the application as compared to when they are receiving welfare benefits.

Second, longer processing times may increase the take-up of welfare benefits and decrease employment rates due to financial stress. The reason for this is that as the period during which applicants have to manage without income lengthens, their financial situation worsens. This can lead to financial stress, which, in turn, can reduce their ability to search for work (Dobbie and Song 2015, 2020, Gathergood 2012). This mechanism may be particularly relevant in the context of means-tested welfare benefits, which are granted to applicants who have already limited or negative wealth at the time of application.

While longer processing times for welfare benefits applications can cause a temporary drop in income, the use of prepayments can largely offset this effect. One may therefore expect that prepayments diminish both the positive and negative employment effects. With benefit prepayments, the incentive to resume work may be lower. At the same time, however, they also may alleviate any financial stress, which in turn may increase the odds of finding work. On the other hand, prepayments can help alleviate liquidity constraints and financial stress, which may increase the likelihood of successful job search. Therefore, the overall effect of prepayments on welfare dependency and employment is ambiguous.

Data 5.3

Data sources 5.3.1

We use administrative individual-level data on all welfare applications submitted in the city of Rotterdam between 2013 and 2019. These data contain information on the application date, the main reason for application, and – if awarded benefits – the starting and ending date of the welfare benefit spell. Additionally, we have information on the caseworker assigned to the application (personal and caseworker team identifier).

We combine the application data with rich administrative records of Statistics Netherlands covering the period between 2012 and 2020. This yields individual-month panel data enabling us to follow applicants for at least one year and up to 8 years both before and after the welfare application. The administrative records provide us with demographic characteristics (such as gender, age, and migration background), labor market outcomes and usage of several social security programs (welfare, UI and disability insurance). The latter dataset also includes information on prepayments of welfare benefits during the application review period.

In total we observe the first welfare applications between 2013 and 2019 of 47,596 individuals. For our empirical analysis, we exclude applications by individuals younger than 27 or older than 49 at the time of application, as their applications are reviewed by different caseworkers (20,981)[7], applications for which the assigned caseworker is not observed (7,123) and applications of which the caseworker reviews too little or too many applications in the according year (5,626).[8] This reduces our final sample to 13,866 observations.[9]

5.3.2 Descriptive statistics

Figure 5.1 shows the distribution of application processing times and the corresponding probabilities of receiving benefit prepayments for our final sample.[10] About half of the applicants start receiving welfare benefits within the formal term of 8 weeks, 44 percent start receiving welfare benefits after 8 weeks, and only 6 percent do not result in welfare benefits

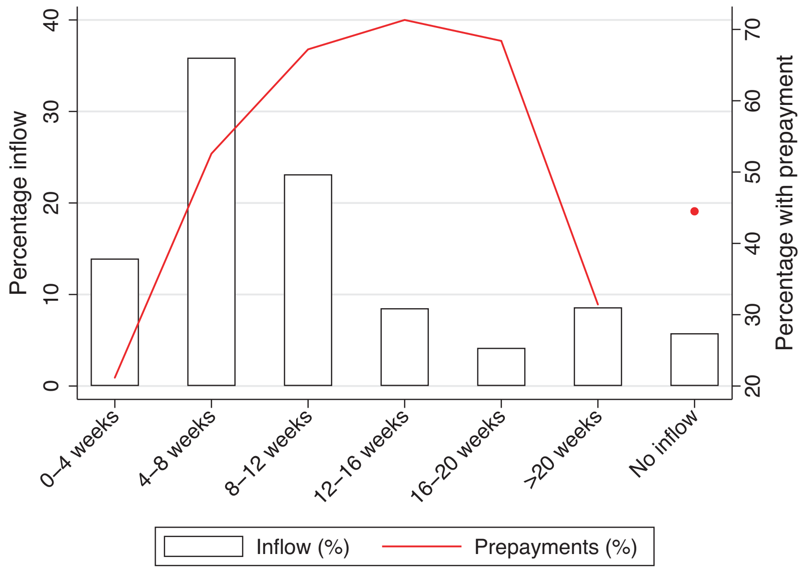
[7]In Figure 5.5 in Appendix 5.A, we see that the average age of the applicants assigned to a caseworker shows three spikes, namely at the ages 23, 38 and 49. This mirrors the fact that there are caseworkers mostly assessing applications for individuals below the age of 27, between the age of 27 and 49 and 50 years and older, respectively.

[8]In our analysis, we exploit information on the caseworker for causal inference. Consequently, we impose a lower bound on the number of applications per caseworker to decrease measurement error. The upper bound screens out administrative staff who assigned applications to themselves before assigning them to the caseworkers. We show the robustness of our results to different lower and upper bounds when discussing our main results in Subsection 5.5.1.

[9]In Table 5.8 in Appendix 5.A, we show that the selection rule based on age changes the sample substantially. The sample of applicants aged 27-49 is statistically significantly different from our main sample, but the differences are small. The remaining differences between the sample based on age and our main sample stem from the exclusion of applicants for which the caseworker is unobserved and not on the selection rule based on the number of applicants per caseworker. The applicants with no observed caseworker have less distance to the labor market. If these caseworkers are unobserved as a result of quick withdrawal of the application, than these applicants are not compliers. The selection rule based on the number of applicants per caseworker in a specific year is uncorrelated with individual characteristics.

[10]The application processing times are calculated as the difference between the application date and the date of entry into welfare. However, in the case of rejected or withdrawn applications, we cannot calculate the application processing times, as the date of rejection or withdrawal is unobserved.

Figure 5.1: Application processing times and received prepayments



receipt. Furthermore, the propensity of receiving prepayments increases with the application processing time. Not surprisingly, applicants who start receiving benefits faster are in less need of prepayments. Similarly, the prepayments for applicants who are expected to be eligible are being paid from week 5 onward. In effect, about 20 percent of the applicants with the shortest application processing times (up to 4 weeks) receive prepayments, while this share increases to between 50 and 70 percent for applicants with longer processing times. It is only for the subsample with processing times longer than 20 weeks that on average 30 percent receive prepayments before entering welfare. Of the sample of applicants that do not enter welfare, either due to rejection or withdrawal, about 45 percent receive prepayments. For them, the prepayments cannot be deducted from the benefit payments and remain outstanding claims.

Table 5.1 shows the individual characteristics and labor market histories of our sample. We are interested in the impact of welfare application processing times, as well as the role of prepayments pending the appli-

Table 5.1: Descriptive statistics of applicants by application outcome

	Full sample	Fast inflow	Slow inflow with prepayments	Slow inflow without prepayments	No inflow	p-value difference: Full vs. Fast inflow	Full vs. Slow inflow with prepayments	Full vs. Slow inflow without prepayments	Full vs. No inflow
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Demographics									
Age 27–31	0.32	0.33	0.33	0.28	0.37	0.02	0.17	0.00	0.00
Age 32–36	0.23	0.23	0.23	0.24	0.24	0.72	0.58	0.31	0.64
Age 37–41	0.17	0.17	0.17	0.18	0.15	0.26	0.73	0.02	0.04
Age 42–46	0.17	0.18	0.16	0.20	0.15	0.90	0.04	0.00	0.11
Age 47–49	0.10	0.10	0.10	0.11	0.09	0.13	0.38	0.13	0.11
Female	0.48	0.49	0.51	0.42	0.42	0.00	0.00	0.00	0.00
Native	0.22	0.23	0.22	0.18	0.29	0.15	0.76	0.00	0.00
First generation migrant	0.53	0.52	0.53	0.61	0.45	0.01	0.19	0.00	0.00
Second generation migrant	0.24	0.25	0.25	0.20	0.26	0.11	0.09	0.00	0.31
Labor market history and previous benefit eligibility									
Welfare benefit receipt	0.32	0.37	0.28	0.28	0.19	0.00	0.00	0.00	0.00
Employed	0.45	0.46	0.43	0.41	0.63	0.44	0.18	0.00	0.00
Exhaustion of UI benefits	0.22	0.28	0.16	0.13	0.25	0.00	0.00	0.00	0.05
Observations	13,866	6,907	3,768	2,391	800				
Share of sample	1.00	0.50	0.27	0.17	0.06				

Note: 'Fast inflow' is defined as inflow into welfare within 8 weeks, 'Slow inflow' as inflow after 8 weeks, and 'No inflow' as that the application does not result in welfare receipt. The p-values in columns (6-9) apply to t-tests of different means for the specific subs-sample and full sample. 'Welfare benefit receipt' and 'Employed' are dummies indicating the status in the year preceding the application. 'Exhaustion of UI benefits' is based on the variable indicating the application reason in the application data of the city.

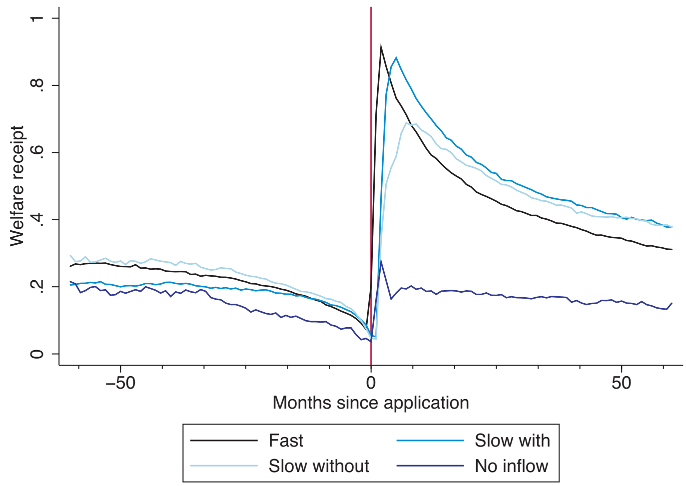
cations as a potential instrument to smooth consumption. Therefore, we split the sample in subsamples stratified by the outcome of the application process (inflow within 8 weeks, inflow after 8 weeks or no inflow) and whether prepayments were paid for applications lasting longer than 8 weeks. This results in the following four subsamples: (i) fast applications – i.e., application processing times below 8 weeks; (ii) slow applications – i.e., longer than 8 weeks – with the provision of prepayments; (iii) slow applications without the provision of prepayments; and (iv) all applications that did not result in welfare receipt (either rejected or withdrawn) and, hence, with unobserved processing times. Obviously, the absence of recorded processing times for either rejected or withdrawn applications raises the question whether these cases could be classified as fast or slow. Note that this concerns no more than 6% of the sample.

Table 5.1 shows that applicants with fast applications – i.e., subsample (i) – received relatively often welfare or UI benefits in the year preceding the application. This suggests that familiarity with the application system increases the probability of a fast application process and mirrors the fact that applicants who received UI benefits before applying are considered less complex to assess. Applicants with slow applications with prepayments – i.e., subsample (ii) – and applicants with slow applications without prepayments – i.e., subsample (iii) – differ the most on gender and migration background. Women receive more prepayments, which might be explained by the relatively large share of recently-divorced or single applicants with young children. The lower prepayment rate among first generation migrants could be the result of being less familiar with the system and/or language barriers that withhold them from requesting financial support. Applicants that do not enter welfare – i.e., subsample (iv) – are younger, more often male, native and were recently active on the labor market. These characteristics suggest that they have a smaller distance to the labor market than the applicants awarded welfare benefits.

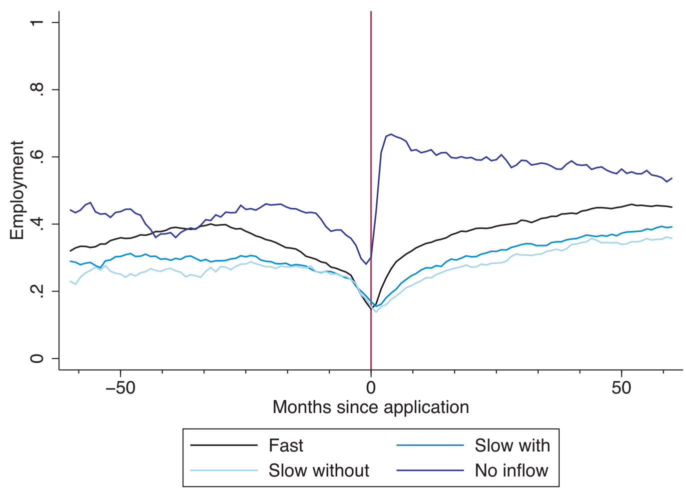
To shed more light on longitudinal patterns, Figure 5.2 shows welfare receipt, welfare benefits, employment and earnings for the four different subsamples before and after the date of application. Welfare benefits (panel (c)) and earnings (panel (d)) are unconditional on welfare receipt and employment. The figure provides three general insights. First, and by

Figure 5.2: Welfare receipt, welfare benefits, employment and earnings before and after month of application by application outcome

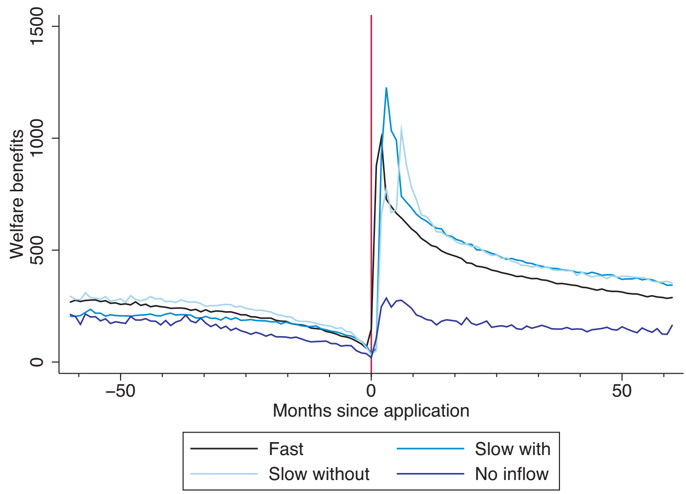
(a) Welfare receipt



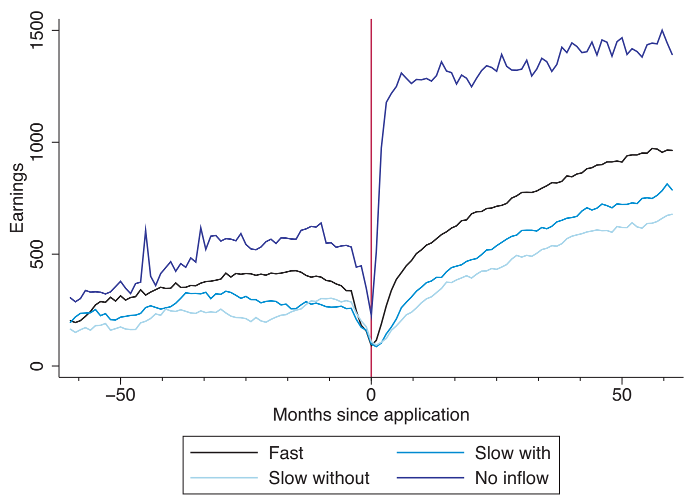
(b) Employment



(c) Welfare benefits



(d) Earnings



construction, the subsample with fast applications has a larger propensity to receive welfare benefits in the first months following the application than the applicants with slow applications, but a lower probability of receiving welfare benefits from four months after application onward. Similarly, those with fast applications have higher employment rates and earnings after the application than those with slow applications. Second, applicants with slow applications with and without prepayments show similar patterns before and after the application. There are only modest differences in the timing of entering welfare and post-application employment. The slow applications without prepayments enter welfare slightly later than those with prepayments and the applicants with prepayments have slightly higher employment rates and earnings after the moment of application. Third, of the applicants whose application does not result in welfare receipt about 20 percent still receive welfare benefits after the application. They also have relatively high employment rates and earnings before and after the application. This concurs with the idea that they have a smaller distance to the labor market than the other applicants and that they are more likely to withdraw their welfare applications.

5.4 Methodology

In the previous section we showed that applicants with fast and slow applications differ both in observed characteristics and pre- and post-application outcomes. This section explains our approach to investigate the causal effects of application processing times on the labor market outcomes and welfare receipt of applicants. Later on, we extend our analysis to benefit prepayments. Throughout our analysis, we rely on exogenous variation of the decision times across caseworkers.

5.4.1 Empirical approach

The applications for welfare benefits are processed by caseworkers to determine the eligibility and the level of benefits. These caseworkers may differ in processing speed and in the frequency of calls for additional information of applicants. As a consequence, application processing times vary across caseworkers. We exploit this variation, given that applications are quasi-randomly assigned to caseworkers; i.e., the assignment is random conditional on team, year and exhaustion of UI benefits before the application. Furthermore, it is important to note that the only way through which the caseworkers are influencing the outcomes of the welfare applicants is via the processing times. That is because the caseworkers involved with claims assessment are rarely in contact with the applicant and are not involved in the monitoring of job search requirements and/or the reintegration activities (other caseworkers are specifically involved with monitoring and reintegration activities). Also, the benefit award decisions themselves provide no discretionary room for the caseworkers. Our instrumental variable approach is inspired by an increasing number of studies exploiting judge or caseworker stringency as an instrumental variable (e.g. Aizer and Doyle Jr 2015, Arni and Schiprowski 2019, Bhuller et al. 2020, Doyle Jr 2007, 2008, Kling 2006, Maestas et al. 2013). Our approach is most similar to Bhuller et al. (2020) who use judge stringency as instrumental variable for incarceration and van der Klaauw and Vethaak (2022) who use caseworker stringency as instrumental variable for the imposition of broader job search requirements.

To estimate the effect of the application processing time T_{ict} on the outcome Y_{ict} of individual i assigned to a caseworker in caseworker team c at time t, we use the following regression equation:

$$Y_{ict} = \alpha_{ct} + \delta T_{ict} + \mathbf{X}'_{it}\beta + \varepsilon_{ict} \tag{5.1}$$

The parameters α_{ct} are the interacted caseworker team and year fixed effects of team c of the caseworker assigned to the application filed in year t. Vector $\mathbf{X}_{it}$ includes individual characteristics, namely gender, age groups, migration background, a dummy indicating previous welfare receipt, a dummy indicating previous employment and also whether the applicant exhausted UI benefits before subsequently applying for welfare benefits. The parameter of interest δ describes the effect of the application processing time T_{ict}. We define T_{ict} as a dummy that equals one if the application processing time for benefit awards is shorter than 8 weeks and is zero otherwise.

Column (5) of Table 5.2 shows that application processing times T_{ict} are related to individual characteristics of the applicant.[11] Hence, the probability of benefit receipt within 8 weeks is endogenous and concerns a selective subsample of the applicants. We therefore rely on exogenous variation in application speed that follows from the quasi-random assignment of applications to caseworkers who may differ in their processing speed. Stated differently, a welfare application that was approved within 8 weeks, might have been awarded after 8 weeks if assigned to a different caseworker in the same caseworker team (or vice versa). We refer to this variation as the processing speed of the caseworker. In our setting, a higher caseworker processing speed is defined as a higher percentage of welfare applications being awarded within 8 weeks after the start of the application. We use this exogenous variation in processing times across caseworkers as an instrument for the individuals' application processing time T_{ict}.

In doing so, we are able to incorporate the small group of applicants who do not enter welfare – either because of rejection, or because of retraction of the application – for which we cannot reconstruct application processing times (only 6% of the sample). With regard to rejections, applicants that do not qualify for welfare are randomly assigned to caseworkers, and caseworkers do not have discretionary room regarding the award decision, such that processing speed and rejections are unrelated (this will also be shown empirically in the next section). Longer processing times, however, can increase the probability of withdrawing the application. Therefore, excluding the either rejected or withdrawn applications will result in measurement bias. In the baseline specification we will thus compare the subsample with fast applications (subsample (i) in the data description) with all other applications (subsamples (ii)–(iv) in the data description). So, for the 'no inflow' group $T_{ict} = 0$. This provides us with the following first-stage regression equation:

[11] Fast application processes appear to be associated with, among other things, not being a first generation migrant and past welfare benefit receipt. This suggests that applicants can experience barriers in the application process that are related to language or familiarity with the system.

$$T_{ict} = \gamma_{ct} + \lambda Z_{j(i)t} + \mathbf{X}'_{it}\theta + \nu_{it} \tag{5.2}$$

The instrumental variable $Z_{j(i)t}$ describes the speed of caseworker j assigned to the application of individual i. We calculate the speed measure as the conventional leave-out mean (similar to e.g., Aizer and Doyle Jr 2015, Bhuller et al. 2020, Maestas et al. 2013). This corresponds to the average rate at which welfare benefits are rewarded within 8 weeks on all applications assigned to caseworker j within the same year, excluding the application of individual i. As a result, we can interpret the estimates of β in the second stage Equation (5.1) as local average treatment effects (LATEs), i.e., the average treatment effect on the group of applicants for whom the application processing time depends on whether they were assigned to a caseworker with a high or low processing speed.

Inherent to our empirical strategy, we can only use the sample of applications for which we can identify a caseworker with a reliable caseworker speed measure. Hence, as previously discussed in Section 5.3, we restrict the sample to applications assigned to caseworkers with at least 25 and at most 400 applications within the specific calendar year. The minimum of 25 is imposed to reduce measurement error in the calculated caseworker speed. The maximum of 400 is used to exclude a few teams which register applications to one staff member instead of individual caseworkers.[12] The remaining sample of 13,866 individual applications is assigned to 162 different caseworkers. Each caseworker team has on average 23 caseworkers over the whole period. Caseworkers are on average in the data for two years.

Extended IV model 5.4.2

As shown in Table 5.1, there are important differences in the use of benefit prepayments among applicants. These prepayments reduce the risk of liquidity constraints. To investigate the effects of prepayments during the application period on welfare receipt and employment, we will further

[12]When testing the robustness of our first stage, we will also choose different thresholds.

divide the subsample of applications with slow application processing times into two groups: slow applications with prepayments and slow applications without prepayments. Our analysis specifically focuses on the effect of prepayments among applicants with slow applications, as prepayments are most relevant in the case of longer application processing times.

We split our sample in the four subsamples listed in Table 5.1 in Section 5.3, namely: (i) fast inflow – i.e., application times below 8 weeks; (ii) slow inflow – i.e., longer than 8 weeks – with prepayments in the period pending the award decision; (iii) slow inflow without prepayments; and finally we also distinguish (iv) those applications that did not result in welfare receipt (we refer to these applications as 'no inflow'). To estimate the effects, we rewrite Equation (5.1) by replacing the dummy T_{ict} with vector $\mathbf{p}_{ict}$:

$$Y_{ict} = \alpha_{ct} + \mathbf{p}_{ict}\mu + \mathbf{X}'_{it}\beta + \varepsilon_{ict}, \tag{5.3}$$

where $\mathbf{p}_{ict}$ includes dummies for three of the four subsamples, with fast inflow as the reference group.[13] We instrument the elements of $\mathbf{p}_{ict}$ with the corresponding caseworker propensities, which again are estimated as leave-out means and where we explicitly allow for correlation between the error terms of the four equations.[14] In other words, for all caseworkers we estimate the propensity that an application results in slow inflow with prepayments, in slow inflow without prepayments or in no inflow, respectively.[15] Consequently, we can interpret the results as local average treatment effects for the subpopulation whose propensity of that particular treatment is affected by the caseworker instruments. We refer to the model

[13]We select the subsample with fast inflow as the baseline since this is the largest subsample.

[14]We instrument each element of $\mathbf{p}_{ict}$ exclusively with the corresponding caseworker propensity instrument, since the treatments and thus the caseworker propensities add up to one and are both mutually exclusive and correlated. Including multiple correlated instruments in each first-stage regression might violate the monotonicity assumption. For example, the effect of the caseworker propensity for slow inflow with prepayments on the probability of having a slow inflow without prepayments is ambiguous. The effect could be positive because of the similarity of slow inflow, but it could also be negative because of the mutual exclusiveness.

[15]Some caseworkers are more inclined to grant prepayments than others.

described in Equation (5.3) as the extended IV model. To improve efficiency of our estimates, we estimate this model using limited-information maximum likelihood (LIML), assuming homoskedastic normal standard errors (Greene 2003).[16]

Justification of the IV assumptions 5.4.3

Before we turn to the estimation results, we will discuss the validity of caseworker speed as an instrumental variable for application processing times. This concerns all four assumptions of the instrumental variables approach: independence, exclusiveness, relevance and monotonicity.

Independence. For causal interpretation of λ in Equation (5.2), the caseworker speed as an instrumental variable should be (quasi-)randomly assigned to individual applicants. Given that applications are randomly assigned to a caseworker *within* teams at the time of application, we control for fully interacted team and year fixed effects in the regression model. As explained in Section 5.2, the only exception to the quasi-randomized assignment is for applications that were filed after exhaustion of UI benefits.[17] Hence, we also control for UI exhaustion before applying for welfare benefits.[18] We test the conditional independence of the instrument by regressing the caseworker speed on individual characteristics of the applicants, while controlling for team and year fixed effects and exhaustion of UI benefits. The results for this test are shown in column (3) in Table 5.2. A joint F-test shows that individual characteristics do not predict the instrument (p-value equals 0.164).[19]

[16] We also estimated our main model – i.e., in which we use caseworker speed as an instrumental variable for processing times – using LIML. This provided us with virtually the same coefficients and standard errors.

[17] This can be seen Figure 5.6 in Appendix 5.A. The vast majority of caseworkers review samples of applications with a 10-30% share of the applicants exhausted UI benefits before applying for welfare benefits. However, there is some bunching at zero and some outliers with a distinctly larger share of applicants who exhausted UI benefits.

[18] This is similar to Maestas et al. (2013) who use examiner stringency as instrumental variable for SSDI receipt. They expect that body system and terminal illness indicators were taken into account in the assignment of cases to examiners. Therefore, they control for these variables in their first-stage regression.

[19] We perform a similar test for the three instruments used in the extended model. The results of this test in Table 5.9 in Appendix 5.B show that two instruments are

Table 5.2: Descriptive statistics, assignment of caseworker speed and the observed application processing time

	Explanatory variables		Dependent variables			
			Fast caseworker speed		**Fast processing time**	
	Mean	Standard Deviation	Coefficient Estimate	Standard Error	Coefficient Estimate	Standard Error
	(1)	(2)	(3)	(4)	(5)	(6)
Demographics						
Age 27–31	0.324	(0.468)	—		—	
Age 32–36	0.232	(0.422)	0.0017	(0.0035)	-0.0196*	(0.0111)
Age 37–41	0.171	(0.377)	0.0016	(0.0042)	-0.0277**	(0.0121)
Age 42–46	0.174	(0.379)	0.0011	(0.0040)	-0.0119	(0.0109)
Age 47–49	0.099	(0.299)	-0.0088*	(0.0050)	-0.0329**	(0.0149)
Female	0.479	(0.500)	0.0058*	(0.0030)	0.0275***	(0.0081)
Native	0.223	(0.416)	—		—	
First generation migrant	0.533	(0.499)	0.0005	(0.0037)	-0.0256**	(0.0100)
Second generation migrant	0.245	(0.430)	0.0010	(0.0039)	-0.0063	(0.0115)
Labor market history and previous benefit eligibility						
Welfare benefit receipt	0.319	(0.466)	0.0077*	(0.0040)	0.1188***	(0.0094)
Employed	0.451	(0.498)	-0.0040	(0.0031)	-0.0087	(0.0095)
F-statistic for joint significance			1.46		20.61	
[p-value]			[.164]		[.000]	
	Number of applications = 13,866			Number of caseworkers = 162		

Note: Column (3) shows OLS estimates of caseworker speed (=percentage of welfare applications awarded within 8 weeks) on individual characteristics of welfare applicants. Column (5) shows a linear probability model of fast processing time (1 = benefit receipt within 8 weeks) on individual characteristics of welfare applicants. All regressions include controls for exhaustion of UI benefits and team fixed effects interacted with year fixed effect. Standard errors are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Exclusion restriction. In our context, the exclusion restriction states that the instrumental variable affects the outcomes of the applicants through the application processing time channel, and not in any other way. If caseworkers differ in any other dimensions than their processing speed, these should be orthogonal to the caseworker speed instrument. Recall that the caseworkers to whom the applications are assigned are not involved in the monitoring of job search requirements and/or in the provision of job search assistance for the welfare recipients (if awarded benefits). Also, recall that the caseworkers have no discretionary room in the benefit award decision. The benefit award decision and the level of welfare benefits are purely based on the provided information. (We will show that there are no systematic differences between caseworkers in award rates and the level of benefits awarded.)

Caseworkers, however, also make decisions on benefit prepayments in case of (expected) longer applications.[20] And faster caseworkers might differ in the rate that they grant prepayments. In what follows, we will provide empirical evidence that the exclusion restriction holds.

First, Panel A of Table 5.3 shows the relationship between our instrumental variable (caseworker processing speed, which is the percentage of applications processed within 8 week), caseworkers award rate, and whether or not benefits are awarded.[21] Column (3) shows that caseworkers do not significantly differ in award rates and that the award decision is only affected by the processing speed of the caseworker and not by the stringency regarding the award decision. This is in line with the fact that the award decision follows directly from the information provided by the applicant and that there is little or no discretion for caseworkers to deviate from this outcome. Higher award rates then only follow from less withdrawn applications as a consequence of shorter waiting times.

uncorrelated with individual characteristics and that one instrument is only marginally significant at the 10%-level.

[20]This resembles the situation of Bhuller et al. (2020), where judges not only decided on incarceration. Instead, trial decisions were multidimensional, as judges could also decide on fines, community service, probation and guilt.

[21]The caseworker award rate is computed using leave-out means. This rate is not significantly correlated with individual characteristics.

Table 5.3: Effect of caseworker processing speed on award rates

	(1)	(2)	(3)
Panel A: Benefits awarded			
Caseworker processing speed	0.080***	—	0.085***
	(0.012)		(0.013)
Caseworker award rate	—	0.134	-0.040
		(0.086)	(0.096)
Panel B: Benefits awarded within 8 weeks			
Caseworker processing speed	0.805***	—	0.803***
	(0.021)		(0.022)
Caseworker award rate	—	1.669***	-0.027
		(0.227)	(0.064)
Panel C: Use of prepayments			
Caseworker processing speed	-0.292***	—	-0.189***
	(0.040)		(0.028)
Caseworker prepayment rate	—	0.487***	0.321***
		(0.055)	(0.053)
Number of applications = 13,866		Number of caseworkers = 162	

Note: The caseworker award rate and caseworker prepayment grant rate are computed using leave-out means. All regressions include controls for age dummies, gender, nationality, labor market history, UI exhaustion, and team fixed effects interacted with year fixed effects. Standard errors in parentheses are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In panel B we conduct similar regressions, but now with the probability of being awarded benefits within 8 weeks as the dependent variable. In line with expectations, from column (3) we conclude that only our instrumental variable caseworker speed significantly predicts the probability of being award benefits within 8 weeks and that the award rate is not important in the combined model.

In panel C, we check whether caseworkers differ in the likelihood that they grant prepayments and whether this is correlated with their application processing speed. In column (1), we find that our instrumental variable is negatively correlated with the probability of receiving prepayments. This is not surprising, since the need for prepayments is smaller in case of fast applications. Column (2) shows that having a caseworker with a higher propensity to grant prepayments is predictive for the probability of receiving prepayments. Column (3) shows that in a combined model, the processing speed and the prepayment rate together predict the proba-

Table 5.4: Effect of caseworker processing speed on the monthly level of welfare benefit payments

Dependent variable:	First payment (1)	Second payment (2)	Third payment (3)	Fourth payment (4)
Caseworker speed	-1,293***	-10	-27	-12
	(88)	(22)	(21)	(19)
Dependent mean (s.d.)	1,884 (1,360)	858 (378)	868 (385)	861 (335)
Number of applicants	13,288	11,597	11,319	10,775

Note: All regressions include controls for age dummies, gender, nationality, labor market history, UI exhaustion, and team fixed effects interacted with year fixed effects. Standard errors in parentheses are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

bility of receiving prepayments. This is a potential threat to the exclusion restriction. Therefore, we will check the robustness of our main results by including the award rate and prepayment rate as additional regressors to our empirical model in Subsection 5.5.1.

As a second test on the validity of the exclusion restriction, we empirically test the possibility that caseworkers use their discretion to impact the level of awarded benefits. Table 5.4 shows the relationship between the caseworker processing speed and the benefit payments if awarded benefits. Since overdue payments stemming from the application period are paid retroactively and potential prepayments are deducted from the first payment, we are interested in the welfare benefit levels after the first payment.[22] Columns (2)-(4) show that these subsequent payments are unaffected by caseworker speed.[23] (As expected, column (1) shows that the first welfare payments are on average higher and that caseworker speed is strongly negatively correlated with the level of the first welfare payment.)

[22]Note that the first welfare payment does not necessary coincide with the first month after application.

[23]Conditional on household status, there is little variation in the level of welfare benefits. Some (downward) changes in the benefit level might originate from household income, inhabiting children or outstanding claims. Most of the variation in the income of welfare recipients stems from the different fiscal income supplements that we do not observe, such as housing subsidies, child subsidies, and health insurance subsidies. The caseworkers assessing the applications do not decide on these supplements.

Instrument relevance. Figure 5.3 shows the distribution of the caseworker speed instrument both unconditional (a) and conditional (b) on UI exhaustion and interacted team and year fixed effects. The right-hand-side panel also adds a local linear regression that describes the relationship between the residual variation in caseworker speed and the residual variation in our treatment dummy T_{ict} (following Dahl et al. (2014) and Bhuller et al. (2020)). The unconditional distribution shows large variation in the processing speed of the caseworkers. About half of the applications are awarded welfare benefits within 8 weeks. The conditional caseworker speed roughly follows a normal distribution. After conditioning on UI exhaustion and fully interacted team and year fixed effects, the standard deviation of the distribution decreases from 0.209 to 0.166. The local linear regressions show that applicants assigned to a caseworker in the 5th percentile have about a 35% probability of being awarded benefits within 8 weeks, compared to a 65% probability if assigned to a caseworker in the 95th percentile. The figure thus shows the first evidence for a strong instrument.

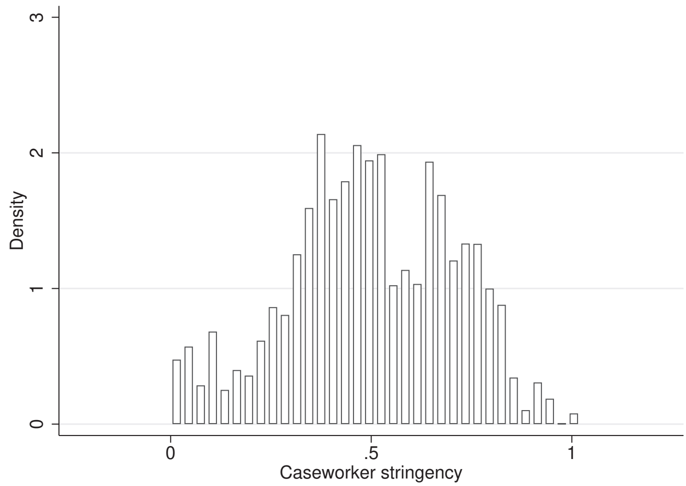
Table 5.5 shows the parameter estimates of λ of the first-stage Equation (5.2) for the full sample and for subsamples of applicants with different characteristics. The first-stage estimate for the full sample is 0.805 (with a standard error of 0.019), which indicates that the probability of benefits awarded within 8 weeks largely depends on the average processing time of the caseworker (F-statistic = 1,752). Caseworker speed is therefore a relevant and strong instrument for the applicants' probability of being awarded welfare benefits within 8 weeks of application.[24] Additionally, Figure 5.7 in section 5.A shows that the differences in caseworker speed are persistent over time (conditional on UI exhaustion and interacted team and year fixed effects). This indicates that some caseworkers are systematically faster than others.

Monotonicity. In the current setting, the monotonicity assumption states that applicants assigned to a caseworker with long processing times (low speed) who started receiving welfare benefits within 8 weeks would also

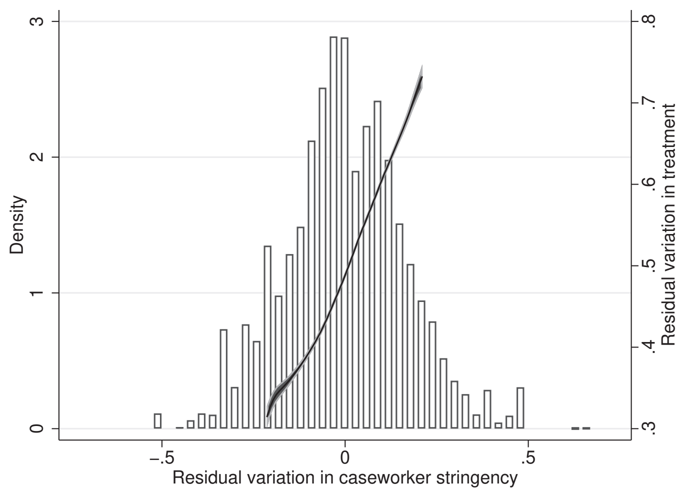
[24] In Table 5.10 in Appendix 5.A we show the robustness of the first-stage to different thresholds for the minimum and maximum number of applications per caseworker and year and to the inclusion of control variables. Consistent with the expectations, we only see a small change in the coefficients when we control for exhaustion of UI benefits.

Figure 5.3: Distribution of caseworker speed (a) and conditional on UI exhaustion and team and year fixed effects (b)

(a) Caseworker speed



(b) Caseworker speed (demeaned)



Note: The histograms show the density of caseworker speed along the left y-axis (both figures). Residual variation in the treatment probability (inflow within 8 weeks) stems from a regression of the treatment on all variables listed in Table 5.1 and fully interacted caseworker team and year fixed effects. The demeaned caseworker speed is conditional on UI exhaustion and team and year fixed effects. The probability of treatment (inflow within 8 weeks) is plotted on the right y-axis (right-hand-side figure) against leave-out mean caseworker speed along the x-axis. The solid line shows a local linear regression of residual variation in the treatment dummy on demeaned caseworker speed. Grey area shows 90% confidence intervals.

Table 5.5: First-stage estimates of caseworker speed on fast inflow by subgroups

	Coefficient	S.e.	F-stat	N	Dependent Mean
	(1)	(2)	(3)	(4)	(5)
Full sample					
Full sample	0.805***	(0.019)	1,752	13,866	0.498
Gender					
Female	0.798***	(0.030)	690	6,640	0.505
Male	0.813***	(0.024)	1,108	7,226	0.492
Age					
Age 27–35	0.791***	(0.026)	899	7,161	0.503
Age 36–49	0.822***	(0.032)	674	6,705	0.493
Nationality					
Native	0.861***	(0.042)	424	3,087	0.518
First generation migrant	0.804***	(0.026)	959	7,388	0.482
Second generation migrant	0.753***	(0.040)	351	3,391	0.516
Welfare receipt in preceding year					
In welfare in preceding year	0.789***	(0.040)	380	4,426	0.576
Not in welfare in preceding year	0.809***	(0.025)	1,089	9,440	0.462
Employment history					
Work in preceding year	0.795***	(0.031)	656	6,257	0.505
No work in preceding year	0.814***	(0.027)	932	7,609	0.493

Note: All regressions include controls for age dummies, gender, nationality, labor market history, UI exhaustion, and team fixed effects interacted with year fixed effects. Standard errors in parentheses are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

have received welfare benefits within 8 weeks if assigned to a caseworker with short processing times (high speed), and vice versa. The monotonicity assumption is violated if different caseworkers are strict to different groups of applicants (De Chaisemartin 2017). Although this is not directly testable, Table 5.5 shows that the first-stage estimates for all subsamples are strongly and positively significant and of comparable magnitude. Although this is not conclusive evidence, positive first-stage coefficients for all different subsamples supports the monotonicity assumption (Imbens and Angrist 1994). As the estimates do not differ much between individuals with different characteristics, we can also conclude that there is no specific group of compliers.

Under the monotonicity assumption, a back-of-the-envelope calculation provides us with an indication of the share of compliers. In the unconditional distribution (left panel of Figure 5.3), the caseworker at the 5th percentile awarded benefits within 8 weeks for 11% of the applications, while this is 82% for the caseworker in the 95th percentile. Roughly speaking, about 11% of the applicants are always-takers (are awarded benefits within 8 weeks from all caseworkers), 18% are never-takers (none of the caseworkers award them benefits within 8 weeks) and the remaining 71% are compliers.

Results 5.5

Effects of fast application processing times 5.5.1

Figure 5.4 graphically presents the local average treatment effects of short welfare applications times on the four main labor market outcomes, with t=0 as the moment of the application. The effect on welfare receipt spikes in the first months after the application. Specifically, the probability of welfare receipt for the applicants with short application times is at most 30 percentage points higher than for those with longer application times (graph (a)). Six months after the application this effect reverses. This negative effect is smaller, but lasts for a longer period. A similar pattern is observed for income from welfare benefits, with a positive spike immediately after the application and a negative spike a few months later (graph (c)). The effects on welfare receipt and welfare benefits reflect a predominantly mechanical effect of comparing applicants with short and long application times, where the former group enters welfare quickly and the latter group catches up in the later months. Perhaps more strikingly, the two right-hand graphs of Figure 5.4 show that the length of the application has no significant effect on the employment probability or earnings of the applicants. This suggests that on average a fast application process does not increase employment (e.g. because job search monitoring and welfare to work services start earlier). On the other hand, on average a fast application process does not reduce employment either (as longer waiting times could encourage workers to resume to work).

Figure 5.4 shows substantial timing effects on welfare receipt, but does not provide insight into the accumulated effects of application processing times. Table 5.6 shows the estimated effects on the cumulative outcomes one and two years after the application. Even though there are large effects on the timing of benefit receipt, the estimates show that the application processing times do not affect the total time in welfare or employment. Similarly, the cumulative amounts of welfare benefits, earnings, and total income are on average not significantly affected by the application processing time. Table 5.15 in Appendix 5.C shows the cumulative estimates after one year for different groups of applicants. Most estimates are insignificant. Women and younger applicants, however, experience a (weakly) significant increase in earnings. They may benefit from fast monitoring and welfare to work services, which significantly increases total income of young applicants on average with 806 euros in the first year after application. Second-generation migrants experience a weakly significant decrease in welfare benefits, accompanied with a smaller and not significant increase in earnings.

In Subsection 5.4.3 we stated that the exclusion restriction might not hold when the caseworker speed and the caseworker propensity to grant prepayments together predicted the probability of receiving prepayments. We now test the robustness of our results by including the award rate and the propensity to grant prepayments of the caseworker as additional regressors to the model. Table 5.11 in Appendix 5.B shows that all estimated effects are robust. From this we conclude that there is no reason to believe that the exclusion restriction is violated or that our results suffer from measurement bias. Additionally, Table 5.12 in Appendix 5.B shows that our results are robust to the use of quarter fixed effects instead of year fixed effects.

The results up to this point show that the length of the welfare application process has – apart from large mechanical effects on the timing of welfare receipt – on average no significant effects on the labor market outcomes of the applicants. For women, young applicants, and second generation applicant we do find weakly significant effect on earnings,

Table 5.6: Effects of fast application processing time on cumulative outcomes - instrumental variable estimates

Dependent variable:	Welfare receipt	Work	Welfare benefits	Earnings	Total income
	(1)	(2)	(3)	(4)	(5)
One year after application	0.153	0.149	-253	408	159
	(0.226)	(0.237)	(260)	(386)	(389)
Dependent mean	9.382	3.120	7,531	4,023	11,887
Number of workers			13,866		
Two years after application	-0.422	0.076	-696	794	56
	(0.506)	(0.481)	(512)	(910)	(870)
Dependent mean	16.191	7.290	13,337	10,972	24,962
Number of workers			13,426		

Note: Time in welfare and employment are measured in months. Total income includes benefits from welfare, UI and DI, earnings and income from self-employment. All regressions include controls for age dummies, gender, nationality, labor market history, UI exhaustion, and team fixed effects interacted with year fixed effects. Standard errors in parentheses are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

total income, and/or welfare benefits. Using the extended model in the next section we study the effect of benefit prepayments among applicants with longer application times. As explained in Subsection 5.2.4, benefit prepayments both reduce the financial incentives to search for work as well as the financial strain.

Results of the extended model 5.5.2

Table 5.7 shows the estimation results on the cumulative outcomes two years after the application for the extended model (Equation (5.3)).[25] With this extended model, we aim to uncover two potential mechanisms. First,

[25]The estimation results for the outcomes one year after the application are shown in Table 5.13 in Appendix 5.B. The general conclusions remain the same.

Table 5.7: Estimation results of the extended model on cumulative outcomes two years after application – instrumental variable estimates

Dependent variable:	Welfare receipt	Work	Welfare benefits	Earnings	Total income
	(1)	(2)	(3)	(4)	(5)
No inflow	-11.867*	3.957	-11,269*	10,423	1,180
	(6.336)	(6.369)	(6,651)	(12,600)	(10,873)
Slow with prepayments	1.964***	0.445	1,887**	-660	1,205
	(0.752)	(0.760)	(789)	(1,500)	(1,289)
Slow without prepayments	1.531	-3.700**	2,290	-7,454**	-5,366
	(1.846)	(1.859)	(1,934)	(3,674)	(3,145)
Dependent mean	16.191	7.290	13,337	10,972	24,962
Number of workers			13,426		

Note: Time in welfare and employment are measured in months. The baseline is the group with application processing times shorter than 8 weeks with prepayments. Total income includes benefits from welfare, UI and DI, earnings and income from self-employment. All regressions include controls for age dummies, gender, nationality, labor market history, UI exhaustion, and team fixed effects interacted with year fixed effects. Standard errors in parentheses are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

we investigate the extent to which the effects described in Subsection 5.5.1 are driven by withdrawal effects and by the effects of application times on awarded welfare applicants. Second, we investigate the importance of prepayments among applicants awaiting welfare benefits.

The coefficients in Table 5.7 suggest that applicants that have withdrawn their application indeed spend less time in welfare and receive less income from benefits as compared to those awarded benefits. These effects are however imprecisely estimated. While the effects on employment are less substantial than the effects on welfare receipt, the applicants appear to compensate the loss of welfare benefits by increasing their earnings. The results from the extended model also suggest that the effects of short processing times may be partially offset by the withdrawal effects on a subsample of applicants.[26] In other words, a reduction in welfare dependency among applicants with short processing times may go unnoticed in the earlier analysis due to a (similar) reduction among applicants who withdrawn their application, as these effects balance each other out.

The results of the extended model concur with the idea that longer processing times, both with and without prepayments, have an impact on welfare receipt. We find that, after excluding applicants who do not enter welfare, applicants with long processing times receive greater amounts of welfare benefits and over a longer period compared to those with short processing times. While these effects are only statistically significant for applicants who received prepayments during their long processing times, the effects for those without prepayments are of similar magnitude.

Next, we investigate the effects of prepayments among applicants with longer processing times. We already concluded that the effects on welfare receipt are not significantly different between applicants with long processing times with and without prepayments. However, the results indicate that applicants with long processing times who do not receive prepayments work and earn significantly less than applicants with fast processing times. As a consequence, among applicants with long processing times, those who did not receive prepayments have worse labor market outcomes than those who received prepayments.

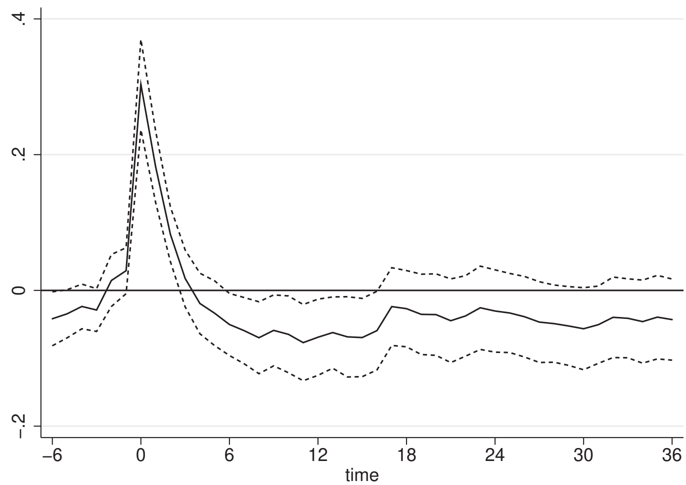
Table 5.14 in Appendix 5.C provides additional evidence that applicants with long processing times have better labor market outcomes when receiving prepayments. In this appendix we estimate an instrumental variables model on the subsample of applicants with long processing times. The prepayments are instrumented with the caseworker prepayment grant rate, which is shown to be uncorrelated with individual characteristics (F-stat=0.46, p-value=0.904). To control for the potential effect of the processing times, we also account for the observed processing times. Although the coefficients are not precisely estimated, they indicate that applicants with long processing times receive less welfare benefits and work more if they receive prepayments. These results are robust to alternative model

[26]Even though the point estimates of the withdrawal effects are substantial, the overall impact on the results in Subsection 5.5.1 are expected to be limited as the size of the population of the compliers (the applicants who potentially withdraw their application) is rather small.

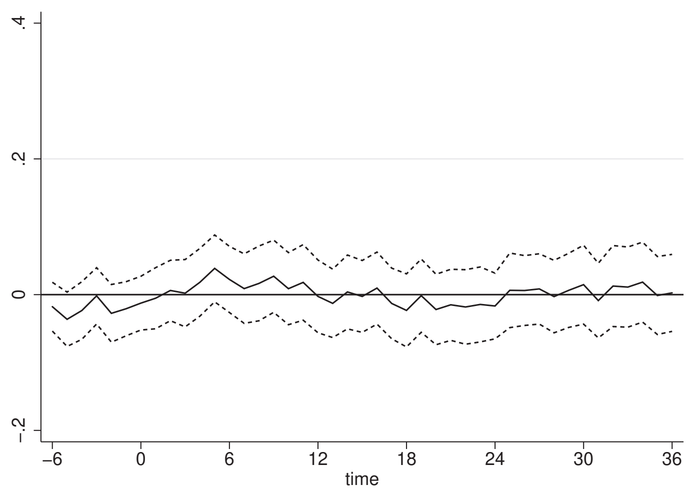
specifications, including those in which caseworker speed is controlled for instead of observed processing times. Both the results in Table 5.7 and Table 5.14 show that prepayments have a positive impact on the labor market outcomes of applicants.

Figure 5.4: Effects of fast application processing time – instrumental variable estimates

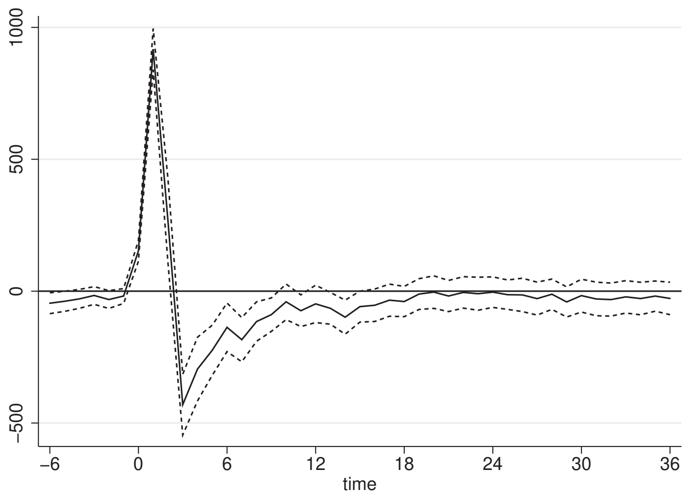
(a) Welfare receipt



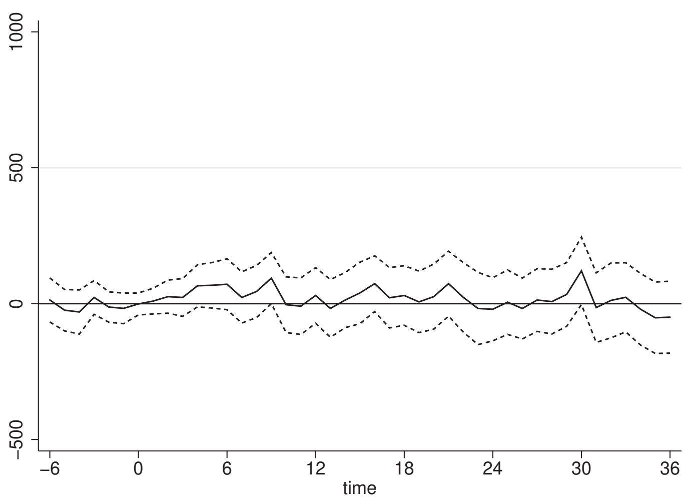
(b) Employment



(c) Welfare benefits



(d) Earnings



Note: Vertical axis displays the probability (top panels) and amounts in Euros (bottom panels). Regressions include controls for age dummies, gender, nationality, labor market history, UI exhaustion, and team fixed effects interacted with year fixed effects. $N = 13,866$ for all estimated effects in all panels. Dashed lines display 95% confidence interval based on standard errors clustered on caseworker level. $t = 0$ is the welfare benefit application date.

Conclusion 5.6

The present paper aims to examine the causal impact of welfare application processing times in Rotterdam, the second-largest city in the Netherlands. Unemployed workers submit their welfare applications to the city offices, providing detailed information regarding their living situation, income, wealth, and other relevant factors. Applications are then evaluated by caseworkers, who determine the applicants' eligibility and level of benefits. Caseworkers may request additional information from the applicant in some cases, leading to an extension of the formal application period and delayed award decision. To estimate the effects of application processing times, we exploit variation in application processing speed among caseworkers and the quasi-random assignment of applications to caseworkers.

Our estimation results reveal a substantial mechanical effect of processing times (proxied as dummy indicating award within 8 weeks) on the timing of welfare receipt, but merely economically small and statistically insignificant effects on the total income from or time in welfare or employment. Additional analyses suggest that the (absence of) effects can be attributed to two compensating effects of long processing times on applicants. On the one hand, some applicants have withdrawn their welfare application due to longer processing times. These applicants spend less time in welfare, but are able to compensate the lost benefit income with increased earnings. On the other hand, applicants who entered welfare after experiencing longer processing times tend to remain longer in welfare as compared to applicants who had faster applications. The unintended (overall) income effects are largest for those who are required to wait for an extended period without replacing income from prepayments.

One possible interpretation of our results is that barriers to enroll in welfare, such as long processing times, screen out applicants with better labor market prospects. However, the improved targeting comes at the cost of worse labor market outcomes for another group of applicants, namely the awarded applicants who experience longer processing times. This is consistent with Autor et al. (2015), who find long-lasting reductions in employment and earnings due to long processing times for SSDI applicants. Our findings show that much shorter processing times compared to those for SSDI applicants can already lead to significant reductions in post-application employment and earnings.

Our findings provide a novel perspective on the important trade-off between providing timely income security and ensuring the accuracy of benefit award decisions. While the idea of providing immediate benefits combined with ex post eligibility checks may seem appealing, it inhibits the risk of diluting the deterrence effect of the application processing times and causes financial stress for those ultimately deemed ineligible for benefits. Our findings show that long application processing times without any form of income can have detrimental effects on the (eligible) applicants facing severe financial constraints. Specifically, applicants who face longer processing times work less and have a significantly lower income if they are not receiving benefit prepayments. These results align with previous studies highlighting the potentially detrimental effects severe liquidity constraints can have on individuals (Dobbie and Song 2015, Gathergood 2012). The positive effects of prepayments are thus not limited to resolving liquidity constraints, but they can also improve the labor market prospects of vulnerable individuals.

5.A Justification of the assumptions: Additional tables and figures

Figure 5.5: Distribution of mean age of the applicants by caseworker

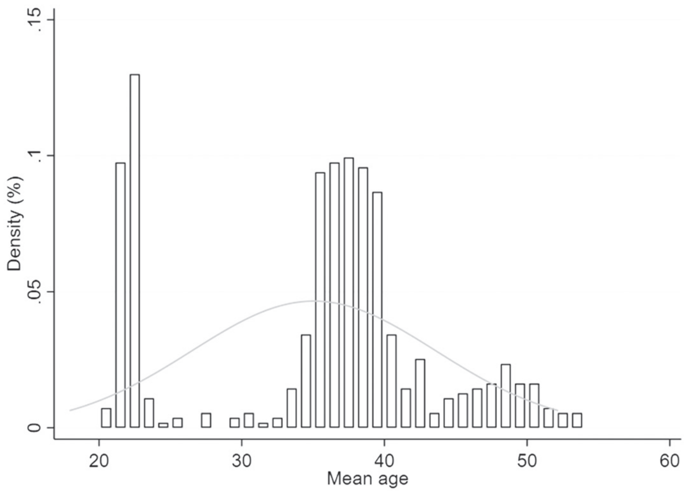


Table 5.8: Sample selections and descriptives of the samples

	All first applicants	Applicants aged 27-49	+ Non-missing caseworker	+ Caseworkers with 25-400 applicants (main sample)	p-value difference: Main sample vs. All first applicants	Main sample vs. Aged 27-49	Main sample vs. Aged 27-49 + Non-missing
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Demographics							
Age 27–31	0.175	0.312	0.324	0.324	0.00	0.00	0.82
Age 32–36	0.130	0.232	0.233	0.232	0.00	0.90	0.77
Age 37–41	0.098	0.176	0.171	0.171	0.00	0.05	0.80
Age 42–46	0.099	0.178	0.173	0.174	0.00	0.09	0.72
Age 47–49	0.057	0.101	0.099	0.099	0.00	0.21	0.99
Female	0.464	0.463	0.481	0.479	0.00	0.00	0.37
Native	0.229	0.213	0.221	0.223	0.04	0.00	0.49
First generation migrant	0.513	0.567	0.532	0.533	0.00	0.00	0.85
Second generation migrant	0.258	0.220	0.246	0.245	0.00	0.00	0.37
Labor market history and previous benefit eligibility							
Welfare benefit receipt	0.281	0.317	0.310	0.319	0.00	0.36	0.00
Employed	0.394	0.398	0.448	0.451	0.00	0.00	0.14
Exhaustion of UI benefits	0.140	0.184	0.219	0.221	0.00	0.00	0.28
Observations	47,596	26,615	19,492	13,866			

Note: The p-values in columns (5-7) apply to t-tests of different means for the specific subs-sample and full sample. The main sample refers to the sample used in the main analysis in this paper. ‘Welfare benefit receipt’ and ‘Employed’ are dummies indicating the status in the year preceding the application. ‘Exhaustion of UI benefits’ is based on the variable indicating the application reason in the application data of the city.

Figure 5.6: Distribution of mean of the applicants who applied after exhaustion of UI benefits by caseworker

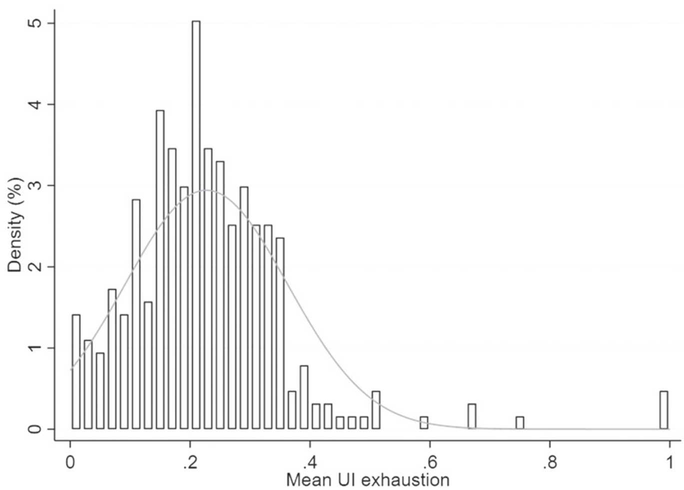
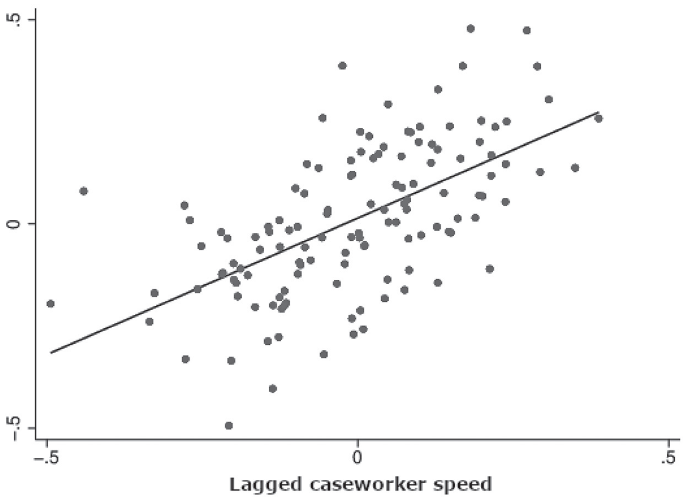


Figure 5.7: Caseworker speed in period t and $t-1$



Note: The black line shows the calculated prediction of the linear relationship between the caseworker speed and its lagged value.

Table 5.9: Testing for random assignment of caseworker instruments used in the extended model

Instrumental variable:	Discouragement		Slow with prepayments		Slow without prepayments	
	Coefficient Estimate	Standard Error	Coefficient Estimate	Standard Error	Coefficient Estimate	Standard Error
	(1)	(2)	(3)	(4)	(5)	(6)
Demographics						
Age 27–31	—		—		—	
Age 32–36	-0.0010	(0.0007)	0.0003	(0.0035)	-0.0010	(0.0015)
Age 37–41	-0.0004	(0.0009)	-0.0014	(0.0030)	0.0002	(0.0018)
Age 42–46	-0.0003	(0.0009)	-0.0005	(0.0030)	-0.0004	(0.0015)
Age 47–49	0.0006	(0.0011)	0.0055	(0.0037)	0.0027	(0.0020)
Female	-0.0007	(0.0007)	-0.0032*	(0.0019)	-0.0019	(0.0014)
Native	—		—		—	
First generation migrant	0.0003	(0.0008)	0.0006	(0.0025)	-0.0256	(0.0017)
Second generation migrant	-0.0013	(0.0009)	0.0014	(0.0027)	-0.0020	(0.0017)
Labor market history and previous benefit eligibility						
Welfare benefit receipt	-0.0004	(0.0008)	-0.0061**	(0.0028)	-0.0012	(0.0014)
Employed	-0.0002	(0.0006)	0.0044**	(0.0021)	-0.0002	(0.0013)
F-stat for joint significance	0.60		1.69		0.93	
[p-value]	[.795]		[.091]		[.495]	
	Number of applications = 13,866			Number of caseworkers = 162		

Note: OLS estimates of caseworker instruments on individual characteristics. All regressions include controls for exhaustion of UI benefits and team fixed effects interacted with year fixed effect. Standard errors are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.10: First-stage estimates using different sample selections on caseworkers and different controls

Sample selection	20-400	25-400†	30-400	25-300	25-500
	(1)	(2)	(3)	(4)	(5)
Panel A. No controls					
Caseworker speed	0.809***	0.824***	0.837***	0.824***	0.825***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.018)
F-stat. (Instrument)	1,905	1,972	1,896	1,972	2,010
Panel B. Add exhaustion of UI benefits					
Caseworker speed	0.798***	0.813***	0.827***	0.813***	0.814***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)
F-stat. (Instrument)	1,678	1,699	1,636	1,699	1,720
Panel C. Add demographic controls					
Caseworker speed	0.797***	0.812***	0.826***	0.812***	0.813***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)
F-stat. (Instrument)	1,703	1,722	1,669	1,722	1,743
Panel D. Add labor market history controls					
Caseworker speed	0.793***	0.805***	0.819***	0.805***	0.805***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)
F-stat. (Instrument)	1,750	1,752	1,658	1,752	1,760

Note: †The baseline analysis uses caseworkers meeting 25-400 benefits recipients. All regressions include local office fixed effects interacted with month fixed effect. The demographic controls are age groups, gender and migration background dummies. The labor market history controls are previous welfare receipt and employment dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.B Robustness of the results

Table 5.11: Effects of fast application processing time on cumulative outcomes – instrumental variable estimates with additional caseworker stringency controls

Dependent variable:	Welfare receipt	Work	Welfare benefits	Earnings	Total income
	(1)	(2)	(3)	(4)	(5)
One year after application	0.018	0.147	-431	574	164
	(0.318)	(0.340)	(360)	(573)	(555)
Dependent mean	9.382	3.120	7,531	4,023	11,887
Number of workers			13,866		
Two years after application	-0.961	0.292	-1,176	1,252	28
	(0.737)	(0.699)	(749)	(1,390)	(1,273)
Dependent mean	16.191	7.290	13,337	10,972	24,962
Number of workers			13,426		

Note: All regressions include controls for age dummies, gender, nationality, labor market history, UI exhaustion, and team fixed effects interacted with year fixed effects. Additional caseworker stringency controls are the award rate and the prepayment rate, which are computed as leave-out means. Time in welfare and employment are measured in months. Total income includes benefits from welfare, UI and DI, earnings and income from self-employment. Standard errors in parentheses are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.12: Effects of fast application processing time on cumulative outcomes – instrumental variable estimates with quarter fixed effects

Dependent variable:	Welfare receipt	Work	Welfare benefits	Earnings	Total income
	(1)	(2)	(3)	(4)	(5)
One year after application	0.016	0.179	-266	520	228
	(0.247)	(0.260)	(293)	(424)	(419)
Dependent mean	9.382	3.120	7,531	4,023	11,887
Number of workers			13,866		
Two years after application	-0.680	0.080	-734	936	165
	(0.556)	(0.544)	(564)	(1,024)	(952)
Dependent mean	16.191	7.290	13,337	10,972	24,962
Number of workers			13,426		

Note: All regressions include controls for age dummies, gender, nationality, labor market history, UI exhaustion, and team fixed effects interacted with quarter year fixed effects. Time in welfare and employment are measured in months. Total income includes benefits from welfare, UI and DI, earnings and income from self-employment. Standard errors in parentheses are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.13: Estimation results of the extended model on cumulative outcomes one year after application – instrumental variable estimates

Dependent variable:	Welfare receipt	Work	Welfare benefits	Earnings	Total income
	(1)	(2)	(3)	(4)	(5)
No inflow	-5.471*	-1.662	-4,207	2,706	-746
	(3.124)	(3.717)	(3,592)	(6,103)	(5,388)
Slow with prepayments	0.838***	0.296	859**	-250	643
	(0.324)	(0.372)	(372)	(624)	(555)
Slow without prepayments	-0.208	-2.344**	1,122	-4,011***	-3,163**
	(0.797)	(0.918)	(918)	(1,538)	(1,357)
Dependent mean	9.382	3.120	7,531	4,023	11,887
Number of workers			13,866		

Note: Time in welfare and employment are measured in months. The baseline is the group with application processing times longer than 8 weeks with prepayments. Total income includes benefits from welfare, UI and DI, earnings and income from self-employment. All regressions include controls for age dummies, gender, nationality, labor market history, UI exhaustion, and team fixed effects interacted with year fixed effects. Standard errors in parentheses are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.C Results: additional tables

Table 5.14: Effects of prepayments on cumulative outcomes – instrumental variable estimates on the subsample with processing times longer than 8 weeks

Dependent variable:	Welfare receipt	Work	Welfare benefits	Earnings	Total income
	(1)	(2)	(3)	(4)	(5)
One year after application	-0.590	1.699	757	2,077	3,080**
	(1.075)	(1.312)	(1,385)	(1,715)	(1,559)
Dependent mean	9.097	2.991	7,426	3,763	11,475
Number of workers			6,159		
Two years after application	-2.191	4.604*	-1,420	5,268	3,982
	(2.506)	(2.672)	(2,938)	(4,371)	(3,528)
Dependent mean	16.170	6.850	13,464	10,031	24,094
Number of workers			5,949		

Note: All regressions control for observed applications processing times. The regressions additionally include controls for age dummies, gender, nationality, labor market history, UI exhaustion, and team fixed effects interacted with year fixed effects. Time in welfare and employment are measured in months. Total income includes benefits from welfare, UI and DI, earnings and income from self-employment. Standard errors in parentheses are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.15: Effects of fast application processing time on cumulative outcomes after one year for different demographic groups – instrumental variable estimates

Dependent variable:	Welfare receipt	Work	Welfare benefits	Earnings	Total income
	(1)	(2)	(3)	(4)	(5)
A. GENDER:					
1. Female					
Estimate	0.015	0.176	-315	924*	662
(s.e.)	(0.334)	(0.340)	(367)	(484)	(421)
Dependent mean	9.484	3.172	7,568	3,639	11,521
Number of workers			6,640		
2. Male					
Estimate	0.136	0.150	-243	-24	-315
(s.e.)	(0.340)	(0.298)	(391)	(597)	(575)
Dependent mean	9.289	3.073	7,498	4,377	12,223
Number of workers			7,226		
B. AGE:					
1. Aged 27-35					
Estimate	-0.005	0.410	-424	1,175**	806*
(s.e.)	(0.335)	(0.311)	(361)	(568)	(484)
Dependent mean	9.135	3.427	7,220	4,538	12,050
Number of workers			7,161		
2. Aged 36-49					
Estimate	0.292	-0.098	-137	-350	-542
(s.e.)	(0.295)	(0.313)	(344)	(577)	(615)
Dependent mean	9.647	2.793	7,864	3,474	11,712
Number of workers			6,705		
C. NATIONALITY:					
1. Native					
Estimate	0.402	0.178	49	-94	-51
(s.e.)	(0.279)	(0.251)	(303)	(440)	(484)
Dependent mean	9.113	3.664	7,056	5,197	12,634
Number of workers			3,087		
2. First generation migrant					
Estimate	0.036	0.060	-255	199	-85
(s.e.)	(0.307)	(0.318)	(379)	(499)	(453)
Dependent mean	9.518	2.858	7,842	3,339	11,502
Number of workers			7,388		
3. Second generation migrant					
Estimate	0.095	0.538	-850*	622	-189
(s.e.)	(0.457)	(0.463)	(440)	(800)	(692)
Dependent mean	9.331	3.196	7,288	4,445	12,043
Number of workers			3,391		

Note: All regressions include controls for age dummies, gender, nationality, labor market history, UI exhaustion, and team fixed effects interacted with year fixed effects. Time in welfare and employment are measured in months. Total income includes benefits from welfare, UI and DI, earnings and income from self-employment. Standard errors in parentheses are robust and clustered at the caseworker level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6 General discussion

Aims 6.1

This thesis contributes to the existing literature on the impact of social insurance on labor market outcomes. Social insurance must target those in need while maintaining incentives that promote work and reduce the risk of unemployment originating from moral hazard. Therefore, designing social insurance is a complex matter as the goals of insurance and the reduction of unemployment risk are conflicting objectives. Specifically, a more generous program will provide more insurance, but reduces the work incentives and increases moral hazard. Additionally, the incentives driving these effects may differ among groups of individuals, and may be influenced by contextual factors or program design differences. Given these complexities, the continuous investigation of the interplay between social insurance programs and the labor market through empirical analysis is crucial to enhance the overall understanding of these topics. This knowledge can then assist in making evidence-based and welfare-improving policy decisions that strike a better balance between the goals of providing insurance and maintaining work incentives. Therefore, the aim of this thesis is to contribute empirical evidence on the relationship between social insurance and the labor market in the context of the Netherlands.

This final chapter of the thesis presents a synthesis of the findings from the four distinct studies concerning the aim of this thesis. This synthesis focuses on the questions of which selection of individuals is more responsive to (changes in) social insurance program parameters and how this relates to the targeting and welfare-maximizing objectives of the programs. To do so, the synthesis follows two main steps. First, it aims to provide empirical evidence on the general importance of selection into (part-time) work in the Netherlands. The selection of individuals that engage in work and part-time work depends on various factors that influence both individual labor supply decisions and the demand for labor. While understanding this selection process is valuable in gaining insights into labor market participation and work intensity, it is challenging to disentangle the impact of all relevant contributing factors, including individual preferences and social insurance programs. Therefore, the second part of this synthesis zooms in to the role of specific social insurance program parameters in the selection of individuals into work and social insurance, as well as the targeting of social insurance and its welfare effects. This focused analysis of particular social insurance program parameters offers two major advantages. One is that it allows for causal interpretation of the results, the other that results can be directly related to the trade-off between insurance and work incentives.

The remainder of this final chapter is structured as follows. The subsequent section discusses the primary economic theories related to social insurance and employment. Sections 6.3 and 6.4 then link the findings discussed in the preceding chapters to the two primary aims of this thesis and the economic theories presented earlier. Considering the significance of these findings for the design of future social policies, the chapter concludes by offering policy implications. Furthermore, as the findings are not exhaustive, the final section provides recommendations for future research to further deepen our understanding of the interplay between social insurance programs and labor market outcomes.

6.2 Theory on social insurance and employment

6.2.1 Theoretical objectives of social insurance

Social insurance addresses several market failures – including imperfect information, non-rational behavior and incomplete markets – that prevent the private market to achieve the objectives of the welfare state due to adverse selection, liquidity constraints, and other inefficiencies. The objectives of the welfare state are equity, efficiency and administrative feasibility (Barr 2020). The objectives have several aspects on their own, which play an important role throughout this thesis. Starting with the former, the aspects of equity are relieving poverty, reducing inequality and social inclusion. Achieving these goals, particularly the goal of eliminating poverty, requires an effective system capable of identifying those in need. Additionally, this system should be able to provide them with benefits, without barriers that may keep individuals from benefits take-up – e.g. administrative barriers, long processing times or liquidity constraints. The take-up of (welfare) benefits may allow for more social inclusion. Of course, the different social insurance programs play an important role in reducing income inequality (Caminada et al. 2021), but the selection in (part-time) work has also important implications for inequalities related to part-time contracts, gender and having a migration background.

Next, the aspects of efficiency include macro- and micro-efficiency, consumption smoothing, risk sharing and incentives to reduce adverse effects. Macro- and micro-efficiency imply that people should among other things be kept healthy, educated and socially included, that an efficient fraction of GDP should be devoted to social insurance, and that these resources are efficiently distributed between the different programs (Barr 2020). For example, the situation in the disability insurance in the Netherlands in the 1980s (also sometimes referred to as the "Dutch disease"), in which a substantial fraction of the insured individuals received disability benefits, was clearly in conflict with efficiency goals. Where the role of social insurance programs in consumption smoothing and the prevention of large drops in living standards (risk sharing) is apparent, the selection in (part-time) work is also important in that regard, for example through the accumulation of savings, pensions and wealth. Incentives to reduce adverse effects of social insurance on labor supply and saving – such as moral hazard – are important for among other things the sustainability of the welfare state and maximizing both overall and individual welfare.

Finally, administrative feasibility includes intelligibility and absence of abuse. Intelligibility implies that the system should be easy to understand and be provided at low cost. The former implies among other things that uncertainty about income effects of changing working hours and red tape preventing (quick) benefit take-up should be minimized. Absence of abuse implies that those ineligible are successfully excluded from take-up, or stated differently, that moral hazard in minimized.

6.2.2 Employment effects of social insurance

The remainder of the present section discusses the behavioral implications of social insurance – such as unemployment insurance (UI), disability benefits (DI), or welfare – on the employment outcomes of individuals. For simplicity, the generosity of insurance can be interpreted in a broad sense, and may concern policy parameters including the replacement rate, the length of benefit entitlement, and the eligibility conditions. These parameters may all increase the value of social insurance. The theoretical implications highlight that the different goals of social insurance are competing objectives.

Social insurance exerts its influence on the labor market by impacting individuals' labor supply decisions. Given that both consumption and leisure are typically assumed to be normal goods (e.g. Cogan 1980, Heckman 1974a,b), more generous social insurance is expected to have income and substitution effects. Income effects follow from the fact that social insurance increases individuals' income while not working, enabling them to increase both consumption and leisure. In order to do so, they reduce their labor supply. Substitution effects follow from the fact that social insurance decreases individuals' marginal benefit from employment as compared to that of leisure or non-work time. Thus, individuals substitute paid work for more leisure time. Consequently, the overall participation effect of more generous social insurance is negative. However, as reservation wages increase, workers may accept better jobs.

The degree to which individuals respond to changes in the generosity of social insurance depends on their preferences. Policy makers face the challenge of leveraging this heterogeneity in preferences to enhance the targeting of social insurance. Hence, they aim to design social insurance in such a way that it provides income insurance to those in need while it encourages active job search among those able to work, thereby reducing moral hazard. Policy makers often try to achieve this by reducing the generosity of social insurance, while not adversely affecting those in need. Typically, good risks have better labor market prospect or a comparative advantage and are thus expected to be more responsive to changes in social insurance and the first ones to select into employment. For instance, introducing barriers to entry diminishes the expected returns of social insurance, but those in need may be less responsive to these barriers and remain within the social insurance system due to limited labor market options, while able individuals are screened out. Successful implementation of such measures can reduce moral hazard without excluding those in need and lead to an improvement in targeting.

The importance of selection in (part-time) work 6.3

Given that the participation effect of social insurance on labor force participation largely depends on individual preferences, it becomes imperative to explore the patterns of selection into work or non-work. In this regard, Chapter 2 contributes to this exploration by investigating the role of unobserved heterogeneity, such as leisure-time preferences and health, stemming from intensive labor supply choices in the selection process for both part-time and full-time work in the Netherlands, focusing on men and women, respectively. This investigation builds upon an extensive body of literature on selection into (part-time) work (e.g. Heckman 1979, Hotchkiss 1991, Myck 2010, Nakamura and Nakamura 1983, Solon 1988, Zabalza et al. 1980).

The findings of Chapter 2 reveal intriguing gender-specific differences in the selection into employment. Specifically, selection into employment holds greater importance for women than for men. Women with more affluent characteristics – such as preferences for work, education or effort – tend to self-select into both part-time and full-time employment. This implies that the labor supply decisions of women are largely shaped by their preferences. On the other hand, men with more affluent characteristics exhibit selection primarily into part-time employment and not into full-time employment. This gender-based distinction underlines the substantial variations in preferences between men and women in the Netherlands, suggesting that the factors shaping their labor supply decisions may significantly differ.

Even though the insights gained from Chapter 2 show substantial gender-specific differences in the selection into employment, they do not provide definitive conclusions on the gender-specific differences in the complex interplay between social insurance and labor supply decisions. Nonetheless, the higher responsiveness of women to factors influencing their labor supply decisions may also imply a greater sensitivity to social insurance compared to men. Consequently, social insurance policies could have a more pronounced impact on women's labor market behavior, potentially influencing their decision to participate in the workforce or their type of employment. This is consistent with previous research on labor supply in the Netherlands (see Evers et al. 2008, for a meta-analysis of empirical estimates of uncompensated labour supply elasticities for the Netherlands). Importantly, even though the literature documents that much of the variation stems from differences in household compositions (e.g. Jongen et al. 2015), we show that (when controlling for various individual characteristics, including the household composition) unobserved heterogeneity still plays an important role in labor supply decisions, especially among women.[1] Although the Chapters 3, 4 and 5 are not explicitly focused on gender-differences in employment reactions on social insurance parameters, they do confirm that there is substantial heterogeneity in labor market responses between men and women.

6.4 Selection, targeting and welfare effects of social insurance

The present section discusses the direct link between social insurance and the labor market behavior in the Netherlands. It discusses the causal evidence on the effects of social insurance reported in Chapters 3, 4 and 5

[1] Consistent with previous findings, the estimation results for the selection equations in Chapter 2 show that having children, marital status, and having a partner past the early retirement age are indeed more important for the employment decision for women than for men.

in relation with the effects of social insurance parameters on selection, targeting and welfare. Several major conclusions follow from these findings.

First, the empirical findings confirm theoretical predictions, showing that reductions in the (expected) value of social insurance lead to lower take-up rates. This was the case for DI reforms that increased the work incentives upon benefit receipt and increased the costs of benefit applications by requiring more rehabilitation efforts, extending the waiting period, and increasing the threshold for benefit receipt (Chapter 3), caseworker meetings for UI recipients that increase the cost of receiving unemployment benefits (Chapter 4), as well as longer welfare application processing times that increase the benefit application costs (Chapter 5). The findings for the Netherlands are consistent with existing evidence that shows that higher costs for applying or staying in social insurance lower the use of social insurance (Deshpande and Li 2019, Kleven and Kopczuk 2011, Ko and Moffitt 2022, Nichols and Zeckhauser 1982), and that this decline is largely translated into higher employment rates (Blank 2002, Card et al. 2010).

Second, selection effects have likely improved the targeting of the social insurance programs. This is reassuring as there is also evidence from other studies pointing at increases in non-take-up of benefits among individuals in greatest need and, thus, worse targeting (Currie 2006, Deshpande and Li 2019, Finkelstein and Notowidigdo 2019, Ko and Moffitt 2022).[2] Specifically, the Dutch DI reforms have led to increased self-screening among workers, excluding those with residual earnings capacities who rely on benefits to supplement their labor income. Together with that the DI enrollment was internationally high before the reform and attributed to moral hazard (Burkhauser and Daly 2011, Koning and Lindeboom 2015), the reduced inflow by increased self-screening points at improved targeting.

Similarly, the welfare application processing times screened out workers with better labor market characteristics. This follows from an additional heterogeneity analysis (not reported in this thesis, results are available

[2]Only recently, the Dutch Committee Social Minimum stated in their report that recipients of social benefits and income supplements, for example rent supplements, are reluctant to changing their labor supply choices as they do not fully understand the consequences of those choices for their income and fear significant repayments (Commissie Social Minimum 2023).

upon request) that shows applicants with better labor market characteristics (e.g. males, individuals that did not receive welfare before, and natives) are more responsive to the length of processing times and thus more likely to self-select out of social insurance and into employment. That these screened-out applicants are in fact on average able to replace the lost benefits with increased wage earnings and that this group concerns only a small part of the total sample of applicants suggests that these effects improved the targeting of the welfare program.

Although the selection effects of the caseworker meetings are not directly being tested, there is reason to believe that these meetings are also improving the targeting of the UI program. There is a substantial literature that argues that moral hazard is presence in large numbers in UI (Chetty 2008, Gruber 1997, Krueger and Meyer 2002) and that unemployed workers generally have biased beliefs giving them too high reservation wages (Belot et al. 2019, Krueger and Mueller 2016, Mueller et al. 2021). When the caseworker meetings reduce the reservation wage and increase the job search effort without the presence of liquidity effects at the side of the individual worker, this can be interpreted as a reduction of moral hazard (Chetty 2008). As caseworker meetings have no direct financial consequences, the findings indeed point at improved targeting.

Third, improved targeting does not necessarily translate into welfare improvement from societal perspective.[3] This is demonstrated by the effects of caseworker meetings for unemployed workers and welfare application processing times.[4] These interventions not only changed the targeting of social insurance, but also had adverse income effects on individuals. Caseworker meetings reduced reliance on unemployment benefits, but unemployed workers were unable to fully compensate for the reduction in benefits with increased earnings, leading to a net decrease in their income.[5] Similarly, the processing times reduced the employment

[3]Note that welfare improvement does not require Pareto improvement and may mean deterioration of either individual or government outcomes.

[4]It is unlikely that the DI reforms are not welfare-improving given the large targeting improvements, even though there are also adverse effects as substitution towards UI (Borghans et al. 2014, Koning and Van Vuuren 2010) and probably higher mortality rates (García-Gómez and Gielen 2018).

[5]Another adverse effect could arise, namely a reduction of the job search effect. The job search effect implies that more generous social insurance results in higher reservation

outcomes of the applicants that are awarded welfare benefits after long application times. This effect fully offsets the positive earnings effect that the processing times have by screening out the individuals with stronger labor market positions and suggest that periods of inactivity resulting from longer processing times reduce human capital (see also Fasani et al. 2021, Marbach et al. 2018).

Fourth, the findings highlight the substantial impact caseworkers have on the outcomes of individuals in the Dutch social insurance system (Bolhaar et al. 2020, Garcia-Gomez et al. 2023, Verlaat et al. 2021). Caseworker discretion allows for more tailored treatment and improved targeting, but increases the element of luck or horizontal inequality (Kahneman et al. 2021). The findings in this thesis confirm that caseworkers are indeed able to some extent to identify the workers for whom the returns are the expected to be the highest (Chapter 4). However, the results in this thesis also indicate that the element of (bad) luck is present (Chapters 4 and 5).

Finally, the findings in this thesis often underscore the complexity and multidimensionality of designing social insurance programs. More specifically, policy makers must consider constraints that can work counterproductively. For instance, job search requirements can constrain job seekers in their job search behavior and therefore lead to inefficient job search, increased reliance on social insurance and decreases in employment (Chapter 4). Similarly, the positive employment effects of benefit prepayments among welfare applicants suggest that financial stress induced by liquidity constraints can hinder job search (Chapter 5).[6]

6.5 Policy implications and future research

The findings presented in this thesis have implications for the design of policies related to income inequalities, social insurance programs, and active labor market policies.

wages, which in turn increases the quality of the post-unemployment matches. Additional analyses (not reported in this thesis, but available upon request) show that the caseworker meetings have negligible effects on the job characteristics after unemployment.

[6] In a related paper, we show with a similar sample that debt relief can have positive employment effects among debtors who faced more severe debt situations (de Bruijn et al. 2023). This result is consistent with that liquidity constraints can hinder job search.

The large positive selection into part-time work compared to full-time work (Chapter 2), for both men and women, has important consequences for the estimation of various inequality measures. Failure to correct for such selection would lead to an overestimation of earnings, especially of those in part-time employment. Consequently, previous estimates of the part-time wage gap in the Netherlands may have underestimated the true gap (Fouarge and Muffels 2009, Russo and Hassink 2008).[7] This finding also has implications for the gender wage gap (Blau et al. 2021, Weichselbaumer and Winter-Ebmer 2005) and the native/migrant wage gap (Anderson et al. 2006, Dustmann and Schmidt 2000, Neumark 2018, Pager 2007, Riach and Rich 2002), as women are more often employed in part-time jobs and individuals with different migration background may choose different working hours. Estimation biases in these gaps could result in unintended income inequality and affect policies related to savings, pensions, and wealth accumulation. Therefore, future research should focus on correctly estimating these wage gaps, while accounting for all relevant sources of heterogeneity.

Additionally, the findings in Chapter 2 confirm the presence of significant differences in work preferences both between and within genders in the Netherlands, which is consistent with previous research (e.g. Mastrogiacomo et al. 2013). Similarly, Chapters 3, 4 and 5 highlight substantial heterogeneity in labor market responses among different groups such as men and women, younger and older workers, and natives and migrants. These results emphasize the importance of conducting heterogeneous analyses on the Dutch labor market.

In accordance to Chapter 2, the findings in Chapter 3 highlight the importance of selection effects. Specifically, the findings show that the DI reforms in the Netherlands largely changed the composition of applicants, while reforms exhibited limited success in stimulating work among the benefit recipients. Therefore, it is imperative for policy makers to take these selection effects into account when designing future reforms to ensure that the intended impact is achieved among the appropriate individuals.

[7] Depending on the selection effects in other countries, the estimates of the part-time wage gap in the literature might also be prone to estimation bias (e.g. Aaronson and French 2004, Bardasi and Gornick 2008, Blank 1990a, Ermisch and Wright 1993, Hirsch 2005, Manning and Petrongolo 2008, O'Dorchai et al. 2007, Rodgers 2004).

Next, the adverse effects of imposing the broader search task (Chapter 4) seem to contradict results from earlier studies that show positive effects of stimulating a broader job search through 'information treatments' (Belot et al. 2019, Skandalis 2019). An important difference lies in the nature of the task, as the broader search task is part of a formal and mandatory program. Unemployed workers that comply with this mandatory task might be a distinct population from those responsive to information treatments, as information treatments might affect mainly the beliefs about the returns to the job search among unemployed workers who were too optimistic. This surmise is supported by the fact that unemployed workers who are most likely to receive the broader search requirements experience the largest adverse effects of the requirements. This is likely due to caseworkers assigning the task mainly to specialized workers who were optimising their job search before receiving the mandatory task and were therefore searching narrowly before the meeting. It is thus likely that the respondents of an information treatment are different from the group targeted by the caseworkers. Further investigation is needed to explore whether caseworker meetings focused on broader search without the mandatory task might have more beneficial effects, akin to information treatments. This can be accomplished relatively easily by investigating the recently changed broader search program in which the mandatory component has been eliminated.

Finally, the mixed effects of welfare application processing times on employment outcomes (Chapter 5) warrant careful consideration in the design of the review process for applications. Although processing times have positive employment effects by screening out those with more favorable labor market characteristics, they also decrease the employment of applicants awarded benefits after longer processing times. Similarly, the findings show that welfare benefit prepayments on average increase the employment of the applicants with longer processing times. The presence of this (local average treatment) effect implies that there are also (untreated) applicants without prepayments who might benefit from the prepayments as well. Furthermore, the findings also point at the existence of information and/or language barriers within the application process and when requesting prepayments. Taken together, these findings described above call for reflection on the welfare application review process. Policy makers should strike a balance between increased targeting and better employment outcomes for awarded applicants who also face longer processing times. Additionally, considering more generous welfare benefit prepayments for those experiencing longer processing times could alleviate financial stress and facilitate successful job search.[8] Addressing the information and/or language barriers could potentially serve as a means to decrease processing times or increase the use of prepayments.

[8]Additionally, policy makers should attempt to simultaneously prevent liquidity constraints and the emergence of new welfare debts. As shown in a complementary paper in which we employ a similar sample of individuals, many welfare debts originate from the application period, and that welfare debts may have adverse and long-lasting effects on subsequent employment outcomes (de Bruijn et al. 2023).

Bibliography

Aaronson, D. and E. French (2004): The Effects of Part-time Work of Wages: Evidence from Social Security Rules. *Journal of Labor Economics*, 22:329–352. Cited on page 216.

Aizer, A. and J. J. Doyle Jr (2015): Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges. *Quarterly Journal of Economics*, 130(2):759–803. Cited on pages 101, 116, 117, 160, 175, and 177.

Altmann, S., A. Falk, S. Jäger, and F. Zimmermann (2018): Learning about Job Search: A Field Experiment with Job Seekers in Germany. *Journal of Public Economics*, 164:33–49. Cited on pages 8, 101, 127, 133, and 141.

Anderson, L., R. Fryer, and C. Holt (2006): Discrimination: Experimental Evidence from Psychology and Economics. *Handbook on the Economics of Discrimination*, pages 97–118. Cited on page 216.

Arni, P. and A. Schiprowski (2019): Job Search Requirements, Effort Provision and Labor Market Outcomes. *Journal of Public Economics*, 169:65–88. Cited on pages 101, 102, 115, 119, 160, and 175.

Autor, D., M. Duggan, and J. Gruber (2014): Moral Hazard and Claims Deterrence in Private Disability Insurance. *American Economic Journal: Applied Economics*, 6(4):110–41. Cited on pages 161 and 166.

Autor, D., N. Maestas, K. J. Mullen, A. Strand, et al. (2015): Does Delay Cause Decay? The Effect of Administrative Decision Time on the Labor Force Participation and Earnings of Disability Applicants. Technical report, National Bureau of Economic Research. Cited on pages 160, 161, 166, and 195.

Autor, D. H. and M. G. Duggan (2003): The Rise in the Disability Rolls and the Decline in Unemployment. *Quarterly Journal of Economics*, 118(1):157–206. Cited on pages 5, 58, 60, and 90.

Averett, S. and J. Hotchkiss (1997): Female Labor Supply with a Discontinuous Nonconvex Budget Constraint: Incorporation of a Part-time / Full-time Wage Differential. *The Review of Economics and Statistics*, 79:461–470. Cited on page 13.

Baker, M. (1997): Growth-rate Heterogeneity and the Covariance Structure of Life-cycle Earnings. *Journal of Labor Economics*, 15(2):338–375. Cited on pages 12 and 44.

Baker, M. and G. Solon (2003): Earnings Dynamics and Inequality among Canadian men, 1976-1992: Evidence from Longitudinal Tax Records. *Journal of Labor Economics*, 21(2):267–288. Cited on pages 12 and 44.

Bardasi, E. and J. Gornick (2008): Working for Less? Women's Part-time Wage Penalties across Countries. *Feminist Economics*, 14(1):37–72. Cited on pages 18 and 216.

Barr, N. (2020): *Economics of the Welfare State*. Oxford University Press, USA. Cited on page 209.

Basten, C., A. Fagereng, and K. Telle (2014): Cash-on-hand and the Duration of Job Search: Quasi-experimental Evidence from Norway. *The Economic Journal*, 124(576):540–568. Cited on pages 161 and 166.

Been, J., M. Hurd, and S. Rohwedder (2021): Households' Joint Consumption Spending and Home Production Responses to Retirement in the US. *Review of Economics of the Household*, 19:959–985. Cited on page 36.

Been, J., M. Knoef, and H. Vethaak (2023): A Panel Data Sample Selection Model to Estimate Life-cycle Earning Profiles: How Important is Selection into Full-time and Part-time Employment? Technical report. Cited on page 11.

Bell, A., M. Fairbrother, and K. Jones (2019): Fixed and Random Effects Models: Making an Informed Choice. *Quality & Quantity*, 53(2):1051–1074. Cited on page 35.

Belot, M., P. Kircher, and P. Muller (2019): Providing Advice to Jobseekers at Low Cost: An Experimental Study on Online Advice. *Review of Economic Studies*, 86(4):1411–1447. Cited on pages 7, 8, 100, 101, 127, 133, 141, 214, and 217.

Benitez-Silva, H., R. Disney, and S. Jiménez-Martín (2010): Disability, Capacity for Work and the Business Cycle: an International Perspective. *Economic Policy*, 25(63):483–536. Cited on page 88.

Van den Berg, G. J. and B. Van der Klaauw (2006): Counseling and Monitoring of Unemployed Workers: Theory and Evidence from a Controlled Social Experiment. *International Economic Review*, 47(3):895–936. Cited on pages 102 and 124.

——— (2019): Structural Empirical Evaluation of Job Search Monitoring. *International Economic Review*, 60(2):879–903. Cited on pages 125, 126, and 131.

Berger, M. and D. Black (1992): Child Care Subsidies, Quality of Care, and the Labor Supply of Low-Income, Single Mothers. *The Review of Economics and Statistics*, 74(4):635–642. Cited on page 22.

Bettendorf, L., E. Jongen, and P. Muller (2015): Childcare Subsidies and Labour Supply — Evidence from a Large Dutch Reform. *Labour Economics*, 36:112–123. Cited on page 22.

Bhuller, M., G. B. Dahl, K. V. Løken, and M. Mogstad (2020): Incarceration, Recidivism, and Employment. *Journal of Political Economy*, 128(4):1269–1324. Cited on pages 101, 115, 116, 117, 160, 175, 177, 181, and 184.

Blank, R. (1990a): Are Part-time Jobs Bad Jobs? In G. Burtless, editor, *A Future of Lousy Jobs? The Changing Structure of U.S. Wages*, pages 123–164. The Brookings Institution, Washington, D.C. Cited on page 216.

——— (1990b): Are Part-time Jobs Lousy Jobs? In G. Burtless, editor, *A Future of Lousy Jobs?*, pages 123–155. The Brookings Institute, Washington, DC. Cited on page 33.

Blank, R. M. (2002): Evaluating Welfare Reform in the United States. *Journal of Economic Literature*, 40(4):1105–1166. Cited on page 213.

Blau, F., L. Kahn, N. Boboschko, and M. Comey (2021): The Impact of Selection into the Labor Force on the Gender Wage Gap. IZA Discussion Paper No. 14335. Cited on page 216.

Blundell, R., T. Magnac, and C. Meghir (1997): Savings and Labor-market Transitions. *Journal of Business & Economic Statistics*, 15(2):153–164. Cited on page 161.

Bolhaar, J., N. Ketel, and B. van der Klaauw (2019): Job Search Periods for Welfare Applicants: Evidence from a Randomized Experiment. *American Economic Journal: Applied Economics*, 11(1):92–125. Cited on page 161.

——— (2020): Caseworker's Discretion and the Effectiveness of Welfare-to-work Programs. *Journal of Public Economics*, 183:104080. Cited on page 215.

Borghans, L., A. Gielen, and E. Luttmer (2014): Social Support Substitution and the Earnings Rebound: Evidence from a Regression Discontinuity in Disability Insurance Reform. *American Economic Journal: Economic Policy*, 6(4):34–70. Cited on pages 88 and 214.

Bosch, N., A. Deelen, and R. Euwals (2010): Is Part-time Employment Here to Stay? Working Hours of Dutch Women over Successive Generations. *Labour*, 24(1):35–54. Cited on page 22.

Bound, J. (1989): The Health and Earnings of Rejected Disability Insurance Applicants. *American Economic Review*, 79(3):482. Cited on pages 59 and 71.

Bound, J., R. V. Burkhauser, and A. Nichols (2003): Tracking the Household Income of SSDI and SSI Applicants. In *Worker Well-being and Public Policy*, pages 113–158. Emerald Group Publishing Limited. Cited on pages 5, 58, and 71.

Bound, J., S. Lindner, and T. Waidmann (2014): Reconciling Findings on the Employment Effect of Disability Insurance. *IZA Journal of Labor Policy*, 3(1):11. Cited on pages 5 and 58.

de Bruijn, E.-J., H. Vethaak, P. Koning, and M. Knoef (2023): Debt Relief for the Financially Vulnerable: Impact on Employment, Welfare Receipt, and Mental Health. Cited on pages 215 and 218.

Burkhauser, R. V. and M. Daly (2011): *The Declining Work and Welfare of People with Disabilities: What Went Wrong and a Strategy for Change*. AEI Press. Cited on page 213.

Caminada, C., M. Brakel, K. Goudswaard, H. Vethaak, J. Been, K. Caminada, E. Jongen, W. Bos, and F. Otten (2021): Inkomensongelijkheid en het Effect van Herverdeling. In *Inkomen Verdeeld, Trends 1977-2019*, pages 41–56. Centraal Bureau voor de Statistiek (CBS)/Universiteit Leiden. Cited on page 209.

Cammeraat, E., E. Jongen, and P. Koning (2022): Preventing NEETs during the Great Recession: the Effects of Mandatory Activation Programs for Young Welfare Recipients. *Empirical Economics*, 62(2):749–777. Cited on page 163.

Campolieti, M. (2006): Disability Insurance Adjudication Criteria and the Incidence of Hard-to-Diagnose Medical Conditions. *The BE Journal of Economic Analysis & Policy*, 5(1):1–25. Cited on page 58.

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*, 122(4):1511–1560. Cited on pages 161 and 166.

Card, D., J. Kluve, and A. Weber (2010): Active Labour Market Policy Evaluations: A Meta-analysis. *Economic Journal*, 120(548):F452–F477. Cited on pages 115, 140, 141, and 213.

Carneiro, P., J. J. Heckman, and E. Vytlacil (2010): Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin. *Econometrica*, 78(1):377–394. Cited on page 134.

Casanova, M. (2010): The Wage Process of Older Workers. Mimeo. Cited on page 20.

Charlier, E., B. Melenberg, and A. Van Soest (2001): An Analysis of Housing Expenditure using Semiparametric Models and Panel Data. *Journal of Econometrics*, 101(1):71–107. Cited on page 12.

Chen, S. and W. Van der Klaauw (2008): The Work Disincentive Effects of the Disability Insurance Program in the 1990s. *Journal of Econometrics*, 142(2):757–784. Cited on pages 59 and 71.

Chetty, R. (2008): Moral Hazard versus Liquidity and Optimal Unemployment Insurance. *Journal of Political Economy*, 116(2):173–234. Cited on pages 161 and 214.

Cogan, J. (1980): Fixed Costs and Labor Supply. Cited on page 210.

Commissie Social Minimum (2023): Een Zeker Bestaan. Naar een Toekomstbestendig Stelsel van het Sociaal Minimum. Cited on page 213.

Currie, J. (2006): The Take-up of Social Benefits. In *Public Policy and the Distribution of Income*, pages 80–148. Russell Sage Foundation. Cited on pages 9, 158, 161, and 213.

Dahl, G. B., A. R. Kostøl, and M. Mogstad (2014): Family Welfare Cultures. *The Quarterly Journal of Economics*, 129(4):1711–1752. Cited on page 184.

Daly, M., D. Hryshko, and I. Manovskii (2022): Improving the Measurement of Earnings Dynamics. *International Economic Review*, 63(1):95–124. Cited on pages 12 and 44.

De Chaisemartin, C. (2017): Tolerating Defiance? Local Average Treatment Effects without Monotonicity. *Quantitative Economics*, 8(2):367–396. Cited on pages 122 and 186.

De Groot, N. and B. Van der Klaauw (2019): The Effects of Reducing the Entitlement Period to Unemployment Insurance Benefits. *Labour Economics*, 57:195–208. Cited on pages 104 and 111.

De Groot, N. and P. Koning (2016): Assessing the Effects of Disability Insurance Experience Rating. The Case of The Netherlands. *Labour Economics*, 41:304–317. Cited on page 63.

De Jong, P., M. Lindeboom, and B. Van der Klaauw (2011): Screening Disability Insurance Applications. *Journal of the European Economic Association*, 9(1):106–129. Cited on pages 5, 58, 64, and 90.

De Nardi, M., G. Fella, M. Knoef, G. Paz-Pardo, and R. Van Ooijen (2021): Family and Government Insurance: Wage, Earnings, and Income Risks in the Netherlands and the US. *Journal of Public Economics*, 193:104327. Cited on page 27.

Deaton, A. and C. Paxson (1994): Savings, Growth and Aging in Taiwan. In *Studies in the Economics of Aging*, pages 331–362. National Bureau of Economic Research. Cited on page 76.

Deshpande, M. and Y. Li (2019): Who Is Screened Out? Application Costs and the Targeting of Disability Programs. *American Economic Journal: Economic Policy*, 11(4):213–248. Cited on pages 58, 158, and 213.

Dobbie, W. and J. Song (2015): Debt Relief and Debtor Outcomes: Measuring the Effects of Consumer Bankruptcy Protection. *American Economic Review*, 105(3):1272–1311. Cited on pages 161, 167, and 196.

——— (2020): Targeted Debt Relief and the Origins of Financial Distress: Experimental Evidence from Distressed Credit card Borrowers. *American Economic Review*, 110(4):984–1018. Cited on pages 161 and 167.

Doyle Jr, J. J. (2007): Child Protection and Child Outcomes: Measuring the Effects of Foster Care. *American Economic Review*, 97(5):1583–1610. Cited on pages 101, 160, and 175.

——— (2008): Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care. *Journal of Political Economy*, 116(4):746–770. Cited on pages 101, 160, and 175.

Duncan, A. and M. Weeks (1997): Behavioral Tax Microsimulation with Finite Hours Choices. *European Economic Review*, 41:619–626. Cited on page 29.

Dustmann, C. and M. Rochina-Barrachina (2007): Selection Correction in Panel Data Models: An Application to the Estimation of Females' Wage Equations. *Econometrics Journal*, 10:263–293. Cited on pages 12 and 13.

Dustmann, C. and C. Schmidt (2000): The Wage Performance of Immigrant Women: Full-time Jobs, Part-time Jobs, and the Role of Selection. IZA Discussion Paper Series, No. 233. Cited on pages xv, 14, 29, 32, 34, 36, 39, 40, 43, 56, and 216.

DWP (2022): Fraud and Error in the Benefit System: Financial Year 2021 to 2022 Estimates. Department for Work and Pensions, London. URL https://www.gov.uk/government/statistics/fraud-and-error-in-the-benefit-system-financial-year-2021-to-2022-estimates/fraud-and-error-in-the-benefit-system-financial-year-ending-fye-2022#contact-information Cited on page 158.

Ermisch, J. and R. Wright (1993): Wage Offers and Full-time and Part-time Employment by British Women. *The Journal of Human Resources*, 28:111–133. Cited on pages 12, 13, 29, 33, and 216.

Evers, M., R. De Mooij, and D. Van Vuuren (2008): The Wage Elasticity of Labour Supply: A Synthesis of Empirical Estimates. *De Economist*, 156:25–43. Cited on page 212.

Fasani, F., T. Frattini, and L. Minale (2021): Lift the Ban? Initial Employment Restrictions and Refugee Labour Market Outcomes. *Journal of the European Economic Association*, 19(5):2803–2854. Cited on pages 166 and 215.

Finkelstein, A. and M. J. Notowidigdo (2019): Take-up and Targeting: Experimental Evidence from SNAP. *Quarterly Journal of Economics*, 134(3):1505–1556. Cited on pages 158 and 213.

Fouarge, D. and R. Muffels (2009): Working Part-time in the British, German and Ddutch Labour Market: Scarring for the Wage Career? *Journal of Contextual Economics–Schmollers Jahrbuch*, 129(2):217–226. Cited on page 216.

Franses, P. and J. Cramer (2010): On the Number of Categories in an Ordered Regression Model. *Statistica Neerlandica*, 64(1):125–128. Cited on pages 29 and 33.

French, E. and J. Song (2014): The Effect of Disability Insurance Receipt on Labor Supply. *American Economic Journal: Economic Policy*, 6(2):291–337. Cited on pages 59, 101, and 160.

GAO (2022): Unemployment Insurance: Data Indicate Substantial Levels of Fraud during the Pandemic; DOL Should Implement an Antifraud Strategy. United States Government Accountability Office, Washington. URL https://www.gao.gov/assets/gao-23-105523.pdf Cited on page 158.

García-Gómez, P. and A. C. Gielen (2018): Mortality Effects of Containing Moral Hazard: Evidence from Disability Insurance Reform. *Health economics*, 27(3):606–621. Cited on page 214.

Garcia-Gomez, P., P. Koning, O. O'Donnell, and C. Riumallo Herl (2023): Selective Exercise of Discretion in Disability Insurance Awards. Cited on page 215.

Gathergood, J. (2012): Debt and Depression: Causal Links and Social Norm Effects. *The Economic Journal*, 122(563):1094–1114. Cited on pages 161, 167, and 196.

Godard, M., P. Koning, and M. Lindeboom (2022): Application and Award Responses to Stricter Screening in Disability Insurance. *Journal of Human Resources*, pages 1120–11323R1. Cited on pages 5, 58, 62, 64, and 90.

Gottschalk, P. and R. Moffitt (1994): The Growth on Earnings Instability in the US Labor Market. *Brookings Papers on Economics Activity*, (2):217–272. Cited on pages 12 and 44.

Greene, W. H. (2003): *Econometric analysis*. Pearson Education India. Cited on page 179.

Gruber, J. (1997): The Consumption Smoothing Benefits of Unemployment Insurance. *The American Economic Review*, 87(1):192. Cited on page 214.

Guvenen, F. (2009): An Empirical Investigation of Labor Income Processes. *Review of Economic Dynamics*, 12(1):58–79. Cited on pages 12 and 44.

Haller, A., S. Staubli, and J. Zweimuller (2020): Designing Disability Insurance Reforms: Tightening Eligibility Rules or Reducing Benefits. NBER Working paper 27602. Cited on page 90.

Hanoch, G. and M. Honig (1985): "True" Age Profiles of Earnings: Adjusting for Censoring and for Period and Cohort Effects. *The Review of Economics and Statistics*, 67(3):383–394. Cited on pages 12 and 13.

Heathcote, J., K. Storesletten, and G. Violante (2010): The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy*, 118(4):681–722. Cited on pages 12 and 44.

Heckman, J. (1974a): Life Cycle Consumption and Labor Supply: An Explanation of the Relationship between Income and Consumption over the Life Cycle. *The American Economic Review*, pages 188–194. Cited on page 210.

——— (1974b): Shadow Prices, Market Wages, and Labor Supply. *Econometrica: journal of the econometric society*, pages 679–694. Cited on page 210.

——— (1976): The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement*, 5(4):475–492. Cited on pages 31 and 46.

——— (1979): Sample Selection Bias as a Specification Error. *Econometrica*, 14(4):481–502. Cited on pages 4, 12, 13, 14, 31, 46, and 211.

Heckman, J. J. and E. Vytlacil (2001): Policy-relevant Treatment Effects. *American Economic Review*, 91(2):107–111. Cited on page 134.

Hirsch, B. (2005): Why do Part-time Workers Earn Less? The Role of Worker and Job Skills. *Industrial and Labor Relations Review*, 58(4):525–551. Cited on page 216.

Hotchkiss, J. (1991): The Definition of Part-Time Employment: A Switching Regression Model with Unknown Sample Selection. *International Economic Review*, 32(4):899–917. Cited on pages 13, 29, and 211.

Hullegie, P. and P. Koning (2018): Employee Health and Employer Incentives. *Journal of Health Economics*, 62:134–146. Cited on page 64.

Imbens, G. W. and J. D. Angrist (1994): Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2):467–475. Cited on pages 123 and 186.

Imbens, G. W. and D. B. Rubin (1997): Estimating Outcome Distributions for Compliers in Instrumental Variables Models. *Review of Economic Studies*, 64(4):555–574. Cited on page 138.

Johansson, P., L. Laun, and T. Laun (2014): Screening Stringency in the Disability Insurance Program. *The BE Journal of Economic Analysis & Policy*, 14(3):873–891. Cited on page 89.

Johnson, T. R. and D. H. Klepinger (1994): Experimental Evidence on Unemployment Insurance Work-search Policies. *Journal of Human Resources*, 29(3):695–717. Cited on page 102.

Jongen, E., H.-W. de Boer, and P. Dekker (2015): De Effectiviteit van Fiscaal Participatiebeleid. Cited on page 212.

Kahneman, D., O. Sibony, and C. R. Sunstein (2021): *Noise: A Flaw in Human Judgment*. Hachette UK. Cited on page 215.

Kalwij, A. and R. Alessie (2007): Permanent and Transitory Wages of British Men, 1975-2001: Year, Age and Cohort Effects. *Journal of Applied Econometrics*, 22:1063–1093. Cited on pages 14 and 34.

Kantarci, T., J.-M. van Sonsbeek, and Y. Zhang (2023): The Heterogeneous Impact of Stricter Criteria for Disability Insurance. *Health Economics, Forthcoming*. Cited on pages 65 and 84.

Kapteyn, A., R. Alessie, and A. Lusardi (2005): Explaining the Wealth Holdings of Different Cohorts: Productivity Growth and Social Security. *European Economic Review*, 49:1361–1391. Cited on page 29.

van der Klaauw, B. and H. Vethaak (2022): Empirical Evaluation of Broader Job Search Requirements for Unemployed Workers. IZA Discussion Papers, No. 15698. Cited on pages 99, 160, and 175.

Klepinger, D. H., T. R. Johnson, and J. M. Joesch (2002): Effects of Unemployment Insurance Work-search Requirements: The Maryland Experiment. *ILR Review*, 56(1):3–22. Cited on page 102.

Kleven, H. J. and W. Kopczuk (2011): Transfer Program Complexity and the Take-up of Social Benefits. *American Economic Journal: Economic Policy*, 3(1):54–90. Cited on pages 158, 161, and 213.

Kling, J. R. (2006): Incarceration Length, Employment, and Earnings. *American Economic Review*, 96(3):863–876. Cited on pages 101, 116, 160, and 175.

Ko, W. and R. A. Moffitt (2022): Take-up of Social Benefits. Technical report, National Bureau of Economic Research. Cited on pages 9, 158, 161, and 213.

Koning, P. and M. Lindeboom (2015): The Rise and Fall of Disability Insurance Enrollment in the Netherlands. *Journal of Economic Perspectives*, 29(2):151–172. Cited on pages 62, 64, 65, 90, and 213.

Koning, P. and J.-M. van Sonsbeek (2017): Making Disability Work? The Effects of Financial Incentives on Partially Disabled Workers. *Labour Economics*, 47:202–215. Cited on pages 5, 58, 65, 84, and 88.

Koning, P. and D. J. Van Vuuren (2010): Disability Insurance and Unemployment Insurance as Substitute Pathways. *Applied Economics*, 42(5):575–588. Cited on pages 88 and 214.

Koning, P. and H. Vethaak (2021): Decomposing Employment Trends of Disabled Workers. *The BE Journal of Economic Analysis & Policy*, 21(4):1217–1255. Cited on page 57.

Krueger, A. B. and B. D. Meyer (2002): Labor Supply Effects of Social Insurance. *Handbook of public economics*, 4:2327–2392. Cited on page 214.

Krueger, A. B. and A. I. Mueller (2016): A Contribution to the Empirics of Reservation Wages. *American Economic Journal: Economic Policy*, 8(1):142–179. Cited on pages 7, 100, 127, and 214.

Kyriazidou, E. (1997): Estimation of a Panel Data Sample Selection Model. *Econometrica*, 65(6):1335–1364. Cited on pages 12 and 13.

Lagakos, D., B. Moll, T. Porzio, N. Qian, and T. Schoellman (2018): Life Cycle Wage Growth across Countries. *Journal of Political Economy*, 126(2):797–849. Cited on pages 12, 43, and 44.

Lammers, M., H. Bloemen, and S. Hochguertel (2013): Job Search Requirements for Older Unemployed: Transitions to Employment, Early Retirement and Disability Benefits. *European Economic Review*, 58:31–57. Cited on page 102.

Liebert, H. (2019): Does External Medical Review Reduce Disability Insurance Inflow? *Journal of Health Economics*, 64:108–128. Cited on page 58.

Lillard, L. and Y. Weiss (1979): Components of Variation in Panel Earnings Data. *Econometrica*, 47(2):437–454. Cited on pages 12 and 44.

Lillard, L. and R. Willis (1978): Dynamic Aspects of Earnings Mobility. *Econometrica*, 46(5):985–1012. Cited on pages 12 and 44.

Maestas, N. (2019): Identifying Work Capacity and Promoting Work: A Strategy for Modernizing the SSDI Program. *The ANNALS of the American Academy of Political and Social Science*, 686(1):93–120. Cited on pages 5, 58, and 90.

Maestas, N., K. J. Mullen, and A. Strand (2013): Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt. *American Economic Review*, 103(5):1797–1829. Cited on pages 59, 60, 73, 101, 116, 117, 121, 160, 175, 177, and 179.

Maibom, J., M. Rosholm, and M. Svarer (2017): Experimental Evidence on the Effects of Early Meetings and Activation. *Scandinavian Journal of Economics*, 119(3):541–570. Cited on pages 102, 115, 140, and 141.

Manjunath, B. and S. Wilhelm (2012): Moments Calculation for the Doubly Truncated Multivariate Normal Density. ArXiv:1206.5387 [stat.CO]. Cited on pages 46 and 47.

Manning, A. and B. Petrongolo (2008): The Part-time Pay Penalty for Women in Britain. *The Economic Journal*, 118(526):F28–F51. Cited on page 216.

Manning, A. and H. Robinson (2004): Something in the Way She Moves: A Fresh Look at an Old Gap. *Oxford Economic Papers*, 56:168–188. Cited on page 33.

Marbach, M., J. Hainmueller, and D. Hangartner (2018): The Long-term Impact of Employment Bans on the Economic Integration of Refugees. *Science Advances*, 4(9):eaap9519. Cited on pages 166 and 215.

Markussen, S., K. Røed, and R. C. Schreiner (2018): Can Compulsory Dialogues Nudge Sick-listed Workers Back to Work? *The Economic Journal*, 128(610):1276–1303. Cited on page 58.

Mastrogiacomo, M., N. M. Bosch, M. Gielen, and E. L. Jongen (2013): A Structural Analysis of Labour Supply Elasticities in the Netherlands. Cited on page 216.

Meghir, C. and L. Pistaferri (2004): Income Variance Dynamics and Heterogeneity. *Econometrica*, 72(1):1–32. Cited on pages 12 and 44.

——— (2010): Earnings, Consumption and Lifecycle Choices. In O. Ashenfelter and D. Card, editors, *Handbook of labor economics*, pages 773–854. Elsevier Science, Amsterdam. Cited on pages 12 and 44.

Moffitt, R. and P. Gottschalk (2012): Trends in the Transitory Variance of Male Earnings. *Journal of Human Resources*, 47(1):204–236. Cited on pages 12 and 44.

Mortensen, D. T. (1986): Job Search and Labor Market Analysis. *Handbook of labor economics*, 2:849–919. Cited on page 161.

Moscarini, G. (2001): Excess Worker Reallocation. *Review of Economic Studies*, 68(3):593–612. Cited on pages 100, 102, 127, 133, 135, and 141.

Mueller, A. I., J. Spinnewijn, and G. Topa (2021): Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence, and Bias. *American Economic Review*, 111(1):324–363. Cited on pages 7, 100, 127, and 214.

Mundlak, Y. (1978): On the Pooling of Time Series and Cross Section Data. *Econometrica*, 46(1):69–85. Cited on page 30.

Myck, M. (2010): Wages and Ageing: Is there Evidence for the "Inverse-U" Profile? *Oxford Bulletin of Economics and Statistics*, 72(3):282–306. Cited on pages 12, 20, and 211.

Nakamura, A. and M. Nakamura (1983): Part-time and Full-time Work Behavior of Married Women: A Model with a Doubly Truncated Dependent Variable. *Canadian Journal of Economics*, 16:229–257. Cited on pages 13, 29, and 211.

NCSL (2023): Child Tax Credit Overview. https://www.ncsl.org/human-services/child-tax-credit-overview.
URL https://www.ncsl.org/human-services/child-tax-credit-overview Cited on page 158.

Neumark, D. (2018): Experimental Research on Labor Market Discrimination. *Journal of Economic Literature*, 56(3):799–866. Cited on page 216.

Nichols, A. L. and R. J. Zeckhauser (1982): Targeting Transfers through Restrictions on Recipients. *American Economic Review*, 72(2):372–377. Cited on pages 158 and 213.

Nichols, D., E. Smolensky, and T. N. Tideman (1971): Discrimination by Waiting Time in Merit Goods. *American Economic Review*, 61(3):312–323. Cited on page 161.

O'Dorchai, S., R. Plasman, and F. Rycx (2007): The Part-time Wage Penalty in European Countries: How Large is it for Men? *International Journal of Manpower*, 28(7):571–603. Cited on page 216.

OECD (2010): *Sickness, Disability and Work: Breaking the Barriers*.
URL https://www.oecd-ilibrary.org/content/publication/9789264088856-en Cited on page 57.

——— (2018): *Adequacy of Guaranteed Minimum Income Benefits*. Cited on page 163.

——— (2020): OECD Employment Outlook. OECD Publishing. Cited on pages 14 and 17.

Pager, D. (2007): The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future. *The Annals of the American Academy of Political and Social Science*, 609(1):104–133. Cited on page 216.

Parsons, D. O. (1991): The Health and Earnings of Rejected Disability Insurance Applicants: Comment. *American Economic Review*, 81(5):1419–1426. Cited on pages 161 and 166.

Petrongolo, B. (2009): The Long-term Effects of Job Search Requirements: Evidence from the UK JSA reform. *Journal of Public Economics*, 93(11-12):1234–1253. Cited on page 102.

Pischke, J. (1995): Measurement Error and Earnings Dynamics: Some Estimates from the PSID Validation Study. *Journal of Business & Economic Statistics*, 13(3):305–314. Cited on pages 12 and 44.

Riach, P. A. and J. Rich (2002): Field Experiments of Discrimination in the Market Place. *The economic journal*, 112(483):F480–F518. Cited on page 216.

Ridley, M., G. Rao, F. Schilbach, and V. Patel (2020): Poverty, Depression, and Anxiety: Causal Evidence and Mechanisms. *Science*, 370(6522):eaay0214. Cited on page 161.

Rochina-Barrachina, M. (1999): A New Estimator for Panel Data Sample Selection Models. *Annales d'Economie et de Statistique*, 55/56:153–181. Cited on pages 12, 13, 14, 15, 31, 33, 36, 39, 41, 43, and 46.

Rodgers, J. (2004): Hourly Wages of Full-time and Part-time Employees in Australia. *Australian Journal of Labour Economics*, 7:215–238. Cited on page 216.

Russo, G. and W. Hassink (2008): The Part-time Wage Penalty: A Career Perspective. *De Economist*, 156(2):145–174. Cited on page 216.

Schiprowski, A. (2020): The Role of Caseworkers in Unemployment Insurance: Evidence from Unplanned Absences. *Journal of Labor Economics*, 38(4):1189–1225. Cited on pages 102, 115, 140, and 141.

Semykina, A. and J. Wooldridge (2010): Estimating Panel Data Models in the Presence of Endogeneity and Selection. *Journal of Econometrics*, 157(2):375–380. Cited on page 12.

——— (2013): Estimation of Dynamic Panel Data Models with Sample Selection. *Journal of Applied Econometrics*, 28(1):47–61. Cited on page 12.

——— (2018): Binary Response Panel Data Models with Sample Selection and Self-selection. *Journal of Applied Econometrics*, 33(2):179–197. Cited on page 12.

Skandalis, D. (2019): Breaking News: The Role of Information in Job Search and Matching. Cited on pages 8, 101, 133, and 217.

Solon, G. (1988): Self-selection Bias in Longitudinal Estimation of Wage Gaps. *Economics Letters*, 28:285–290. Cited on pages 4, 12, and 211.

Stam, M., M. Knoef, and A. Ramakers (2020): The Effects of Welfare Receipt on Crime: A Regression Discontinuity and Instrumental Variable Approach. Technical report. Cited on page 163.

Stancanelli, E. and A. Van Soest (2012): Retirement and Home Production: A Regression Discontinuity Approach. *American Economic Review*, 102(3):600–605. Cited on page 36.

Storer, P. and M. A. Van Audenrode (1995): Unemployment Insurance Take-up Rates in Canada: Facts, Determinants, and Implications. *Canadian Journal of Economics*, pages 822–835. Cited on page 161.

Storesletten, K., C. Telmer, and A. Yaron (2004): Cyclical Dynamics of Idiosyncratic Labor Market Risk. *Journal of Political Economy*, 112(3):36–47. Cited on pages 12 and 44.

Tummers, M. and I. Woittiez (1991): A Simultaneous Wage and Labor Supply Model with Hours Restrictions. *The Journal of Human Resources*, 26:394–423. Cited on page 13.

Van Soest, A. (1995): Structural Models of Family Labor Supply: A Discrete Choice Approach. *The Journal of Human Resources*, 30:63–88. Cited on pages 13, 29, and 33.

Van Sonsbeek, J.-M. and R. H. Gradus (2012): Estimating the Effects of Recent Disability Reforms in the Netherlands. *Oxford Economic Papers*, 65(4):832–855. Cited on page 65.

Verlaat, T., S. Rosenkranz, L. F. Groot, and M. Sanders (2021): Requirements vs. Autonomy: What Works in Social Assistance? *Autonomy: What works in social assistance*. Cited on page 215.

Visser, J., T. Wilthagen, R. Beltzer, and E. Koot-Van der Putte (2004): The Netherlands: From Atypicality to Typicality. In S. Sciarra, P. Davies, and M. Freedland, editors, *Employment Policy and the Regulation of Part-time Work in the European Union: A Comparative Analysis*, pages 190–222. Cambridge University Press, Cambridge. Cited on page 16.

Voas, D. and M. Chaves (2016): Is the United States a Counterexample to the Secularization Thesis? *American Journal of Sociology*, 121(5):1517–1556. Cited on page 71.

Von Wachter, T., J. Song, and J. Manchester (2011): Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program. *American Economic Review*, 101(7):3308–29. Cited on pages 5, 58, 59, 60, 71, and 90.

Weichselbaumer, D. and R. Winter-Ebmer (2005): A Meta-analysis of the International Gender Wage Gap. *Journal of economic surveys*, 19(3):479–511. Cited on page 216.

Wooldridge, J. (1995): Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions. *Journal of Econometrics*, 68:115–132. Cited on pages 12, 13, 14, and 43.

——— (2002): *Econometric Analyses of Cross-section and Panel Data*. MIT Press, Cambridge. Cited on pages 29 and 34.

Zabalza, A., C. Pissarides, and M. Barton (1980): Social Security and the Choice between Full-time Work, Part-time Work and Retirement. *Journal of Public Economics*, 14(2):245–276. Cited on pages 13, 29, and 211.

Samenvatting (Dutch summary)

Empirische analyse van sociale zekerheid, werkprikkels en arbeidsmarktuitkomsten

Bij het ontwikkelen van sociale verzekeringen staan beleidsmakers voor de uitdaging om de balans te vinden tussen het bieden van inkomenszekerheid en het behouden van prikkels om te werken. Deze afweging heeft enerzijds betrekking op het bieden van inkomenszekerheid in het geval van werkloosheid, ziekte en arbeidsongeschiktheid, terwijl het anderzijds van belang is dat er genoeg prikkels overeind blijven die het risico op deze vormen van inactiviteit verminderen. Met andere woorden, dienen sociale verzekeringen toegankelijk te zijn voor degenen die deze nodig hebben, terwijl degenen die in staat zijn om te werken in redelijke mate gestimuleerd moeten worden om dat te doen.

Wat de optimale vormgeving van sociale verzekeringen is hangt uiteindelijk af van individuele en maatschappelijke voorkeuren, politieke keuzes, en sociale normen. Economen kunnen hierbij een toegevoegde waarde hebben door beleidsmakers te ondersteunen met empirisch bewijs over de effecten van (veranderingen in) bestaande regelgeving over sociale verzekeringen op verschillende uitkomsten. Deze uitkomsten omvatten arbeidsparticipatie, inkomen, gebruik van uitkeringen, vrije tijd, en substitutie tussen sociale verzekeringen. Dit bewijs kan beleidsmakers dan helpen bij het maken van *evidence-based* keuzes die beter geïnformeerd zijn en, hopelijk, de welvaart verbeteren.

De uitdagingen waar beleidsmakers voor staan bij het ontwerpen van sociale verzekeringen zijn complex vanwege de multidimensionaliteit van dergelijke verzekeringen. De generositeit van dergelijk verzekeringen is afhankelijk van verschillende programmavoorwaarden, waaronder het vervangingspercentage (i.e. de uitkering als percentage van het laatstverdiende inkomen), de uitkeringsduur en de uitkeringsvoorwaarden. Deze voorwaarden zijn ook mede bepalend voor de sterkte van de prikkels om te werken. De voorwaarden kunnen dus worden aangepast om de juiste balans te vinden tussen inkomenszekerheid en deze werkprikkels, en zo de welvaart te verhogen. Verzekeringen voor arbeidsongeschiktheid hebben bijvoorbeeld doorgaans hoge drempels aan de poort, inclusief lange wachttijden, om het risico op *moral hazard* te verminderen. Uitkeringsgerechtigden van dergelijke verzekeringen hebben vervolgens echter recht op relatief genereuze uitkeringen, met een lange duur, en met weinig verplichtingen om te zoeken naar werk. Dit hangt samen met de verwachte kans op werkhervatting, de oorzaak van werkloosheid en de financiering van het programma middels premies. In vergelijking hiermee zijn bijvoorbeeld bijstandsuitkeringen doorgaans lager, hebben ze kortere wachttijden en strengere zoekverplichtingen, omdat de bijstand functioneert als een laatste vangnet.

Om de effecten van sociale verzekeringen beter te doorgronden, is het cruciaal om de prikkels waardoor deze effecten gestuurd worden te analyseren. Het effect van prikkels kan verschillen tussen groepen, zoals tussen armere en rijkere werknemers. Ook het effect van prikkels afhankelijk van de implementatie, bijvoorbeeld omdat werknemers anders reageren op verplichtingen dan op vrijblijvende adviezen van klantmanagers. Als gevolg van deze verscheidenheid kunnen nieuwe empirische bevindingen zowel breed geaccepteerde theoretische modellen als eerdere empirische studies bevestigen of ontkrachten. Het is dus cruciaal om de prikkels op de arbeidsmarkt en in sociale verzekeringen voortdurend empirisch te blijven onderzoek om daarmee de algehele kennis over deze onderwerpen te vergroten.

Toegepaste economen leveren bewijs voor beleidsrelevante relaties door gebruik te maken van verschillende methoden; zo ook in de verschillende hoofdstukken van dit proefschrift. Veel van deze methoden gebruiken exogene variatie in de onafhankelijke variabele om een causale relatie tussen de onafhankelijke en afhankelijke variabele te vinden. De ideale methode om causaal bewijs te verkrijgen is aan de hand van een experiment, omdat dit een directe vergelijking tussen de *treatment* en de controlegroep mogelijk maakt zonder het risico van mogelijke interfererende factoren (Hoofdstuk 4 van dit proefschrift is gedeeltelijk gebaseerd op een dergelijke experimentele methode). In veel gevallen is het echter niet haalbaar of ethisch om in het belang van het onderzoek individuen uit te sluiten. Daarom maken economen vaak gebruik van natuurlijke experimenten, waarbij (plotselinge) veranderingen in de tijd – zoals hervormingen – of exogene verschillen in beleidsparameters worden benut (zoals in Hoofdstuk 3). Een andere potentiële bron van exogene variatie kan voortkomen uit de willekeurige toewijzing van werkzoekenden aan klantmanagers die verschillen in de geneigdheid om een bepaalde *treatment* in te zetten. Zo kan door correctie op selectie het causale effect van interventies worden geschat (zoals in Hoofdstukken 4 en 5). Een sterk voordeel van de bovenstaande methoden is dat ze vaak beleidsrelevante uitkomsten genereren, omdat ze gebaseerd zijn op de groep die responsief is op de *treatment*, ook wel 'compliers' genoemd. Zelfs bij een vergelijking van niet-willekeurige steekproeven kunnen interessante schattingen worden verkregen door te corrigeren voor de selectie met behulp van selectiemodellen (zoals in Hoofdstuk 2). Hoewel er verschillende onderzoeksmethoden worden toegepast in de studies in dit proefschrift, hebben ze het gebruik van grootschalige administratieve datasets gemeen. Dergelijke datasets hebben meerdere voordelen, waaronder representativiteit, een grote steekproefomvang, hoge nauwkeurigheid, een panelstructuur en de mogelijkheid om ze te koppelen aan gegevens uit andere administratieve bronnen.

Dit proefschrift bestaat uit vier studies, beschreven in Hoofdstukken 2-5. Hoofdstuk 2 onderzoekt de arbeidsparticipatiebeslissingen van mannen en vrouwen in Nederland. Het hoofdstuk onderstreept het belang van selectie in de intensieve marge van het arbeidsaanbod, dat wil zeggen het aantal gewerkte uren. Deze intensieve margebeslissing is bij uitstek belangrijk in Nederland, waar aanzienlijke verschillen bestaan tussen de arbeidsparticipatie van mannen en vrouwen en tussen verschillende geboortecohorten. De overige studies beogen een dieper inzicht te krijgen in de impact van sociale verzekeringen en arbeidsmarktbeleid op de werkgelegenheid en andere arbeidsmarktuitkomsten van individuen in Nederland. Hoofdstuk 3 geeft inzicht in de determinanten van de arbeidsparticipatie onder arbeidsongeschikte werknemers in Nederland, die grotendeels zijn beïnvloed door hervormingen in de arbeidsongeschiktheidsregelingen. Hoofdstuk 4 maakt gebruik van een grootschalig veldexperiment om de effecten van gesprekken met klantmanagers en van bredere werkzoekverplichtingen op de uitkomsten van werkloosheidsuitkeringsontvangers te onderzoeken. Het laatste hoofdstuk onderzoekt de effecten van verwerkingstijden van bijstandsaanvragen en voorschotten op de uitkering op arbeidsmarktuitkomsten van bijstandaanvragers.

In Hoofdstuk 2 staat het schatten van inkomensmodellen centraal. Deze inkomensmodellen zijn belangrijk om arbeidsparticipatiebeslissingen en het gebruik van sociale verzekeringen te begrijpen. Het schatten van deze inkomensmodellen is echter gecompliceerd, aangezien inkomens alleen worden waargenomen bij werkende individuen. Het schatten van inkomensmodellen zonder rekening te houden met de niet-willekeurige selectie in werk leidt tot ernstig inconsistente schattingen van inkomens, zelfs in het geval van longitudinale data. Deze selectie leidt namelijk tot een overschatting van de inkomens van vrouwen en mannen die in deeltijd werken. Verschillende alternatieve methoden zijn mogelijk om het probleem van selectie aan te pakken. Een vaak gekozen methode is om inkomensmodellen uitsluitend te schatten op mannen tussen de 25 en 54 jaar, omdat zij het meest waarschijnlijk voltijd werken en minder geneigd zijn om af te wijken van deze participatiebeslissing. De conclusies die volgen uit dergelijke schattingen zijn echter mogelijk niet generaliseerbaar naar vrouwen, oudere mannen en mannen voor wie voltijd werken minder vanzelfsprekend is. Daarom heeft Hoofdstuk 2 van dit proefschrift als doel de vraag: *"Hoe belangrijk is selectie in voltijd- of deeltijdwerk?"* te beantwoorden.

Om deze vraag te beantwoorden, stelt Hoofdstuk 2 een nieuw selectiemodel voor dat twee stromingen in de literatuur combineert. De eerste stroming in de literatuur richt zich uitsluitend op de methodologische uitdaging van selectie, maar gaat ervan uit dat de selectie enkel plaats vindt in de extensieve marge. De tweede stroming in de literatuur richt zich op de intensieve marge van de arbeidsbeslissing, maar deze modellen hebben het nadeel dat ze enkel toepasbaar zijn op cross-sectionele data – data waarin personen niet over tijd gevolgd worden. Het nieuw voorgestelde model in Hoofdstuk 2 vult dit gat in de literatuur op door een nieuw model te introduceren dat zowel corrigeert voor selectie in de intensieve marge, als toepasbaar is op longitudinale data – data waarin dezelfde personen op meerdere momenten geobserveerd worden. De comparatieve voordelen van dit model worden onderzocht aan de hand van administratieve gegevens die representatief zijn voor Nederland getest. Het model doet recht aan de Nederlandse arbeidsmarkt, aangezien het aandeel deeltijders internationaal gezien hoog is, zowel onder mannen als vrouwen.

De toepassing van het model laat zien dat het van belang is om rekening te houden met selectie in de intensieve marge van arbeidsaanbod. Bij correctie voor selectie in de intensieve marge, vinden we positieve selectie in deeltijdwerk voor zowel mannen als vrouwen. Dit betekent dat individuen met een hogere productiviteit kiezen voor deeltijdwerk, en het niet corrigeren voor deze selectie kan leiden tot een overschatting van hun inkomen. Voor voltijdwerk vinden we alleen soortgelijke positieve selectie bij vrouwen. Als gevolg hiervan is de algemeen aangenomen afwezigheid van selectie bij mannen tussen de 25 en 54 jaar in de literatuur alleen geldig voor voltijdwerk. Bovendien kunnen resultaten gebaseerd op deze mannen niet direct worden vertaald naar vrouwen, oudere mannen en mannen die mogelijk deeltijdwerk verrichten.

In Hoofdstuk 3 wordt de aandacht verschoven naar arbeidsongeschikten. In de afgelopen decennia is de arbeidsparticipatie van arbeidsongeschikte individuen in veel OESO-landen gedaald. Deze trend wordt toegeschreven aan twee factoren. Ten eerste zijn het steeds vaker werknemers met kwetsbaardere posities op de arbeidsmarkt die een uitkering voor arbeidsongeschiktheid (AO) ontvangen. Ten tweede ervaren uitkeringsontvangers na toetreding tot de verzekering vaak geringere prikkels tot werken naast de uitkering of volledige werkhervatting. Nederland is geen uitzondering op deze trend, aangezien het ook een sterke daling heeft gezien van de arbeidsparticipatie onder AO-ontvangers. Tijdens de onderzoeksperiode van de in Hoofdstuk 3 beschreven studie werden in Nederland twee belangrijke hervormingen doorgevoerd, namelijk de *Wet verbetering Poortwachter* in 2003 en de invoering van de *Wet Werk en Inkomen naar Arbeidsvermogen* in 2006. Deze hervormingen hebben de groep van uitkeringsontvangers op twee verschillende manieren beïnvloed. Enerzijds leidden een strengere selectie aan de poort en een verhoogde drempel tot een selectievere groep werknemers in de AO-verzekeringen, met ernstigere aandoeningen en slechtere arbeidsmarktkansen. Anderzijds verhoogde de nieuwe arbeidsongeschiktheidswet de prikkels om te werken voor nieuwe cohorten van AO-ontvangers met resterende verdienmogelijkheden. Daarom heeft Hoofdstuk 3 als doel de vraag: *"In hoeverre wordt de werkgelegenheid van arbeidsongeschikte werknemers in Nederland beïnvloed door selectie effecten en veranderde prikkels om te werken?"* te beantwoorden.

Om deze vraag te beantwoorden, brengt Hoofdstuk 3 de arbeidsparticipatietrends van arbeidsongeschikte werknemers tussen 1999 en 2013 in Nederland in kaart. Daarbij worden zowel selectie effecten als effecten van financiële prikkels tot werken bestudeerd. Hiervoor maakt de studie gebruik van *Age-Period-Cohort* (APC) modellen, die het mogelijk maken om de aanvraagcohort-specifieke effecten te onderscheiden van de kalenderjaar ('*period*') effecten en de verstreken tijd sinds de aanvraag ('*age*') effecten. Aangezien de nieuwe regelgeving betrekking heeft op nieuwe cohorten van aanvragers van uitkeringen voor arbeidsongeschiktheid, zien we de selectie effecten en effecten van veranderde prikkels uitsluitend terug in de cohorteffecten die we schatten. Vervolgens worden deze cohorteffecten verder ontleed in selectie effecten en effecten van veranderde prikkels door APC-modellen te combineren met *difference-in-differences* (DiD) modellen. De DiD-modellen vergelijken afgewezen en geaccepteerde aanvraagcohorten vóór en na de hervormingen, waarbij wordt aangenomen dat de hervormingen de selectie van aanvragers in de twee groepen op gelijke wijze hebben veranderd. De resulterende DiD-schattingen laten daarom alleen de effecten van de door de hervormingen veranderde prikkels om te werken op de geaccepteerde aanvragers zien.

De resultaten tonen aan dat de afnemende arbeidsparticipatie van AO-aanvragers in Nederland grotendeels verklaard kan worden door veranderingen in de samenstelling van nieuwe cohorten van aanvragers van uitkeringen. Van het nieuwste cohort werkt 30 procentpunt minder dan het eerste cohort in de dataset. Daarentegen zijn tijdseffecten verwaarloosbaar, wat erop wijst dat conjunctuureffecten of seculiere trends die alle aanvraag-cohorten in gelijke mate beïnvloeden, niet zo belangrijk waren. De veranderingen in cohorteffecten lopen grotendeels parallel met de twee hervormingen, hoewel er ook een geleidelijke afname van de cohorteffecten wordt waargenomen na de tweede hervorming. De DiD-analyse toont aan dat de prikkels om te werken door de hervormingen nauwelijks veranderd zijn en dat de aanzienlijke cohorteffecten die samenvallen met de hervormingen dus bijna volledig toe te wijzen zijn aan selectie effecten. De hervormingen hebben de zelfselectie onder potentiële aanvragers vergroot, waardoor de resterende selectie van aanvragers vaker ongunstigere demografische en arbeidsmarktkarakteristieken heeft.

Hoofdstuk 4 bestudeert het zoekgedrag van werklozen met een werkloosheidsuitkering (WW). Werkloze werknemers zijn vaak te optimistisch over hun arbeidsmarktkansen. Dit optimisme draagt bij aan de langzame uitstroom uit werkloosheid en verklaart deels de langdurige werkloosheid onder werkzoekenden. Werkloze werknemers baseren hun reserveringsloon vaak op hun vorige loon en zoeken vaak naar werk dat lijkt op hun vorige baan. Het stimuleren van werkloze werknemers om breder te zoeken naar werk kan een positief effect hebben op hun arbeidsmarktuitkomsten, als werkzoekenden inderdaad hun verwachtingen naar beneden bijstellen. Gezien het stimuleren tegen geringe kosten gedaan kan worden, kan dit ook kostenefficiënt zijn voor de uitkeringsinstanties. Een toenemend aantal OESO-landen heeft daarom regelgeving geïntroduceerd welke werkloze werknemers die risico lopen op langdurige werkloosheid verplicht te zoeken naar banen die afwijken van hun vorige werk. Recente literatuur toont aan dat het *aanmoedigen* van werkloze werknemers om breder te zoeken inderdaad voordelen voor hen oplevert. De vraag die echter blijft staan is: *"Wat zijn de arbeidsmarkteffecten van bredere zoekverplichtingen?"*

Om deze vraag te beantwoorden, evalueert Hoofdstuk 4 de de uitvoering van extra dienstverlening ontwikkeld door Uitvoeringsinstituut Werknemersverzekeringen (UWV) dat een bredere werkzoekverplichting oplegt aan werkzoekenden die minstens zes maanden WW-uitkering ontvangen. De dienstverlening – het bredere zoekprogramma – begint met een extra gesprek met een klantmanager, waarbij het eerdere zoekgedrag van de werkzoekende wordt geëvalueerd. Als de klantmanager het eerdere zoekgedrag als te gericht beschouwt, heeft zij de bevoegdheid om de werknemer te verplichten zijn zoekgedrag te verbreden. De werkzoekende is verplicht om aan deze eis te voldoen, hetgeen wordt gecontroleerd door de klantmanager. In de praktijk betekent dit dat de werkzoekende actief moet solliciteren op banen die zich in andere sectoren bevinden, een grotere reisafstand hebben, een lager loon bieden of/en een lager opleidingsniveau vereisen.

Voor de evaluatie van het bredere zoekprogramma heeft UWV een grootschalig veldexperiment uitgevoerd door UWV. Een willekeurige steekproef van ongeveer 130 duizend werkloze werknemers is uitgenodigd voor een verplicht gesprek om hun zoekstrategieën te bespreken. De resultaten van het experiment laten zien dat deelname aan het programma gemiddeld genomen de arbeidsparticipatie verhoogt en de afhankelijkheid van WW-uitkeringen vermindert, waardoor het bredere zoekprogramma kosteneffectief voor UWV is. Vervolgens maken we gebruik van het feit dat klantmanagers aanzienlijk verschillen in hoe vaak ze bredere zoekverplichtingen opleggen tijdens het gesprek en dat werknemers willekeurig worden toegewezen aan klantmanagers binnen lokale UWV-kantoren. Door gebruik te maken van deze exogene variatie – verschillen in de strengheid van de klantmanager – als instrumentele variabele, kunnen we het (geïsoleerde) causale effect van de bredere zoekverplichtingen schatten voor de werkloze werknemers die op gesprek zijn geweest. De resultaten tonen aan dat het opleggen van een bredere zoekverplichting de arbeidsmarktuitkomsten niet verbetert. Integendeel, het vermindert de kans op het vinden van werk en verlengt de duur van de WW-uitkering. Bovendien zijn de baankenmerken minder gunstig na de verplichting om breder te zoeken naar werk, waarbij individuen minder kans hebben op een vast contract en minder uren per week werken. Dit impliceert dat de het gesprek met de klantmanager *an sich* het positieve effect van het bredere zoekprogramma bepaalt, terwijl de bredere zoekverplichting de effectiviteit van het programma vermindert.

De negatieve effecten van het opleggen van de bredere zoekverplichting lijken tegenstrijdig te zijn met de resultaten van eerdere studies die positieve effecten laten zien van het stimuleren van bredere zoeken naar werk. Een belangrijk verschil is echter dat de bredere zoekverplichting in ons onderzoek nageleefd moet worden, terwijl andere studies 'informatie-treatments' zonder verplichting onderzochten. Het is aannemelijk dat informatie-treatments voornamelijk de verwachtingen over de opbrengsten van het zoeken naar werk beïnvloeden bij een kleinere en selecte groep werkzoekenden die te optimistisch waren. Daarentegen impliceert het verplichte karakter van het breder zoekprogramma dat de behandelde groep groter is. Bovendien zullen klantmanagers de bredere zoekverplichting vaker inzetten op een andere groep dan de respondenten van een informatie-treatment. Onze resultaten tonen aan dat werkzoekenden die het meest waarschijnlijk geconfronteerd worden met de verplichting, de werkzoekenden die gericht zochten naar werk, de grootste negatieve effecten ondervinden. Dit kunnen werknemers zijn die voordeel hebben van gericht zoeken naar werk en die hun zoekgedrag al optimaliseerden voordat ze de verplichting kregen opgelegd. Het is waarschijnlijk dat deze werknemers minder responsief zouden zijn op een informatie-treatment. Samengevat, de groep waarop de bredere zoekverplichting gericht is vormt een verklaring voor negatieve effecten van deze verplichtingen, in tegenstelling tot de positieve effecten van informatie-treatments die door andere studies zijn gevonden.

Hoofdstuk 5 breidt de afweging tussen inkomenszekerheid en het behoud van financiële prikkels tot werk bij sociale verzekeringen uit naar informele en administratieve drempels. Eerder onderzoek heeft gevonden dat deze drempels het gebruik van en het vermogen om inkomenszekerheid te bieden van deze verzekeringen hebben verminderd. Tegelijkertijd is er ook een afweging tussen het verstrekken van tijdige uitkeringen en het waarborgen van een correctie beoordeling van een uitkeringsaanvraag. Snelle verstrekking van uitkeringen stelt aanvragers beter in staat om hun consumptie op peil te houden, maar dit kan ten koste gaan van de nauwkeurigheid van de beoordelingen. Deze onnauwkeurigheid kan op latere momenten worden gecorrigeerd, maar kan dan leiden tot aanzienlijke terugbetalingen die financiële stress kunnen veroorzaken. Hoofdstuk 5 draagt bij aan de literatuur door zich te richten op de vraag: *"Wat is het* *effect van de verwerkingstijd van uitkeringsaanvragen en van uitkeringsvoorschotten op de uitkerings- en arbeidsmarktuitkomsten van bijstandaanvragers?"*

In Nederland fungeert de bijstand als sociaal vangnet voor alle werkloze werknemers in huishoudens met onvoldoende inkomen en vermogen. Aanvragers voor een bijstandsuitkering moeten klantmanagers gedetailleerde informatie verstrekken over hun leefsituatie, inkomen, en vermogen. Het proces van het verzamelen en beoordelen van deze informatie kan variëren tussen aanvragers, met name wanneer de verstrekte informatie oorspronkelijk onvolledig was. In het geval van lange verwerkingstijden kunnen aanvragers een verzoek indienen om een voorschot op de uitkering te ontvangen om de periode zonder inkomen te overbruggen.

De empirische analyse in Hoofdstuk 5 maakt gebruik van gegevens over de bijstandsaanvragen in Rotterdam, die relatief de meeste bijstandontvangers in Nederland kent. De gegevens laten aanzienlijke variatie tussen aanvragers zien in de verwerkingstijden van de aanvragen. Deze variatie is gedeeltelijk te wijten aan verschillen in de verwerkingssnelheid van de quasi-willekeurig toegewezen klantmanagers die de aanvragen beoordelen. Net als in Hoofdstuk 4 wordt gebruik gemaakt van de variatie tussen klantmanagers en de quasi-willekeurige toewijzing van aanvragen aan klantmanagers om causale effecten te schatten. De resultaten tonen twee tegenstrijdige effecten van de verwerkingstijden van aanvragen. Enerzijds worden sommige aanvragers ontmoedigd om hun bijstandsaanvraag voort te zetten vanwege langere verwerkingstijden. Deze ontmoedigde aanvragers zitten dus minder lang in de bijstand, maar compenseren het verlies aan inkomsten uit de bijstand door meer inkomsten uit werk. Anderzijds blijven aanvragers die uiteindelijk instromen in de bijstand na langere verwerkingstijden langer afhankelijk van de uitkering dan degenen met snelle aanvragen. Met andere woorden, de bevindingen wijzen op een uitruil, omdat een selectievere groep van nieuwe uitkeringsontvangers gepaard gaat met slechtere arbeidsmarktuitkomsten voor dezelfde groep nieuwe uitkeringsontvangers. Tot slot tonen de resultaten aan dat voorschotten op de bijstandsuitkering de arbeidsparticipatie van toegekende aanvragers met lange verwerkingstijden verhogen. Dit suggereert dat het wegnemen van financiële stress een succesvolle zoektocht naar werk vergemakkelijkt.

Overview of author contributions

The PhD regulations outlined by the Doctorate Board of the Faculty of Law at Leiden University, requires to provide an overview of the author contributions of the (co-authored) chapters within this thesis. The current statement serves this requirement.

Chapter 1: Introduction
Heike Vethaak wrote the manuscript.

Chapter 2: A panel data sample selection model to estimate life-cycle earning profiles: How important is selection into full-time and part-time employment?
Heike Vethaak and Jim Been designed the study. Jim Been and Marike Knoef proposed the econometric model. Heike Vethaak and Jim Been implemented the model into the statistical software. Heike Vethaak constructed the dataset, conducted all analyses and did the data management. Heike Vethaak and Jim Been drafted the manuscript. Marike Knoef reviewed and revised the manuscript.

Chapter 3: Decomposing employment trends of disabled workers
Heike Vethaak and Pierre Koning designed the study. Pierre Koning acquired access to the data. Heike Vethaak constructed the dataset, conducted all analyses and did the data management. Heike Vethaak and Pierre Koning wrote the manuscript.

Chapter 4: Empirical evaluation of broader job search requirements for unemployed workers
This chapter resulted from the MSc thesis of Heike Vethaak of which Bas van der Klaauw was the supervisor. After finishing the thesis Heike Vethaak and Bas van der Klaauw concluded that there was potential to extend the thesis to make it into a research paper and designed the study. Heike Vethaak collected the additional data, conducted all analyses, and drafted the manuscript. Bas van der Klaauw revised the manuscript.

Chapter 5: The effects of application processing times and prepayments on welfare receipt and employment
Heike Vethaak and Ernst-Jan de Bruijn acquired access to the data. Heike Vethaak designed the study, constructed the dataset, conducted all analyses, did the data management and drafted the manuscript. Ernst-Jan de Bruijn, Pierre Koning and Marike Knoef reviewed and revised the manuscript.

Chapter 6: General discussion
Heike Vethaak wrote the manuscript.

Curriculum Vitae

Heike Vethaak was born in Purmerend, the Netherlands on the 16th of May, 1996. After obtaining a BSc degree in Economics and Business Economics, he obtained a MSc degree in Economics from the *Vrije Universiteit Amsterdam*. His MSc thesis, written during an internship at the Netherlands Employee Insurance Agency (UWV), forms the basis for Chapter 4 in this thesis. He also received formal education in Econometrics and Applied Microeconometrics at the Tinbergen Institute.

In 2018, Heike became a PhD candidate at the Department of Economics at Leiden Law School at Leiden University. His doctoral research was part of the *Self-reliance and Social Protection over the Lifecycle* research program, and received financial support from the Instituut GAK. The resulting papers were presented at various international conferences and workshops, such as the *EEA-ESEM*, *EALE*, *IZA* and *IPDC*. Policy relevant work was presented at the Netherlands Economists Day, the Netherlands Bureau for Economic Policy Analysis (CPB), the Dutch Employees Insurance Agency (UWV) and the Dutch Ministry of Social Affairs and Employment (SZW).

Besides his doctoral research, Heike taught several courses in Economics and supervised students in writing their thesis. He also contributed to three chapters in the book *Inkomen Verdeeld, trends 1977-2019*, the international project on inequalities Deaton Review, and articles in the *Economisch Statistische Berichten*. Heike also helps organizing the annual economics conference *KVS New Paper Sessions* since 2021.

In the range of books published by the Meijers Research Institute and Graduate School of Leiden Law School, Leiden University, the following titles were published in 2022 and 2023:

- MI-382 C.M.F. Mommers, *Voluntary return and the limits of individual responsibility in the EU Returns Directive*, (diss. Leiden), Amsterdam: Ipskamp Printing 2022, ISBN 978 94 6421 583 0
- MI-383 M.E. Franke, *Over de grens van de onrechtmatige daad. Een onderzoek naar de plaats van de rechtvaardiging in het buitencontractuele aansprakelijkheidsrecht*, (diss. Leiden), Den Haag: Boom juridisch 2022, ISBN 978 94 6290 152 0
- MI-384 B.L. Terpstra, *Instrumental and normative pathways to compliance. Results from field research on moped drivers*, (diss. Leiden), Amsterdam: Ipskamp Printing 2022
- MI-385 Q. Mauer, *Application, Adaptation and Rejection. The strategies of Roman jurists in responsa concerning Greek documents*, (diss. Leiden), Den Haag: Boom juridisch 2022, ISBN 978 94 6236 291 8, ISBN 978 90 5189 952 8 (e-book)
- MI-386 F. Tan, *The Duty to Investigate in Situations of Armed Conflict. An examination under international humanitarian law, international human rights law, and their interplay*, (diss. Leiden), Amsterdam: Ipskamp Printing 2022
- MI-387 G.G. Lodder, *Recht doen of Recht hebben. Een analyse van de rechten van de migrant op bescherming door de staat tegen arbeidsuitbuiting*, (diss. Leiden), Amsterdam: Ipskamp Printing 2022, ISBN 978 94 6421 679 0
- MI-388 M. Mannan, *The Emergence of Democratic Firms in the Platform Economy: Drivers, Obstacles and the Path Ahead*, (diss. Leiden), Amsterdam: Ipskamp Printing 2022
- MI-389 A.M. Bouland, *'Please Give Me My Divorce'. An Ethnography of Muslim Women and the Law in Senegal*, (diss. Leiden), Amsterdam: Ipskamp Printing 2022
- MI-390 B.J. Braak, *Overcoming ruptures: Zande identity, governance, and tradition during cycles of war and displacement in South Sudan and Uganda (2014-9)*, (diss. Leiden), Amsterdam: Ipskamp Printing 2022
- MI-391 S.R. Varadan, *Article 5 of the UN Convention on the Rights of the Child. Parental guidance and the evolving capacities of the child*, (diss. Leiden), Amsterdam: Ipskamp Printing 2022, ISBN 978 94 6421 700 1
- MI-392 J.M.W. Pool, *De rol van de curator bij de aanpak van onregelmatigheden – Een empirisch-juridisch onderzoek naar de rol van de curator in de praktijk bij de aanpak van onregelmatigheden voor en tijdens faillissement*, (diss.Leiden), Amsterdam: Ipskamp Printing 2022
- MI-393 D.S. Verkroost, *'Met zachte drang'. Uitgangspunten voor jeugdhulpverlening op het snijvlak van vrijwilligheid en dwang*, (diss. Leiden), Den Haag: Boom juridisch 2022, ISBN 978 94 6212 735 7, ISBN 978 94 0011 197 4 (e-book)
- MI-394 D. Mândrescu, *The application of EU antitrust law to (dominant) online platforms*, (diss. Leiden), Amsterdam: Ipskamp Printing 2022
- MI-395 A.H.A. Mohammad, *De normering van academisch ondernemerschap. Perspectieven vanuit het onderwijs(bekostigings)recht, het Europees staatssteunrecht en de academische vrijheid & wetenschappelijke integriteit*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023
- MI-396 T.L. Masson-Zwaan, *Widening the Horizons of Outer Space Law*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023, ISBN 978 94 6421 977 7
- MI-397 W. Zhang, *Achieving decent work in China. A case study of decent working time*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023
- MI-398 T.M. Vergouwen, *The effect of directives within the area of direct taxation on the interpretation and application of tax treaties*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023
- MI-399 R. de Massol de Rebetz, *Beyond the Dichotomy between Migrant Smuggling and Human Trafficking. A Belgian Case Study on the Governance of Migrants in Transit*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023
- MI-400 W. Ruijs, *Pandbelening*, (diss. Leiden), Den Haag: Boom juridisch 2023, ISBN 978 94 6212 819 4, ISBN 978 94 0011 288 9 (e-book)
- MI-401 C. Sriporm, *Franchising Legal Frameworks. A Comparative Study of the DCFR, US Law and Australian Law Regarding Franchise Contracts*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023, ISBN 978 94 6473 088 3
- MI-402 N. Amin, State-building, Lawmaking, and Criminal Justice in Afghanistan: *A case study of the prison system's legal mandate, and the rehabilitation programmes in Pul-e-charkhi prison*, (diss. Leiden), Den Haag: Eleven International Publishing 2023, ISBN 978 90 4730 156 1
- MI-403 J. Tobing, *The Essence of the 1999-2002 Constitutional Reform in Indonesia: Remaking the Negara Hukum. A Socio-Legal Study*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023

MI-404 W. Zhang, *Protection of aviation security through the establishment of prohibited airspace*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023, ISBN 978 94 6473 103 3
MI-405 A.J. Pasma, *Re-entry support from prison-based and community-based professionals*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023
MI-406 K.R. Filesia, *Speaking the same language. De invoering van de Anglo-Amerikaanse trust in het Nederlandse recht*, (diss. Leiden), Deventer: Kluwer 2023
MI-407 H.A. ten Oever, *Zorginkoopovereenkomst. De rechtsverhouding tussen de zorgverzekeraar en de zorgaanbieder in contractenrechtelijk perspectief*, (diss. Leiden), Den Haag: Boom juridisch 2023, ISBN 978 94 6212 859 0, ISBN 978 94 0011 339 8 (e-book)
MI-408 F.H.K. Theissen, *Sincerely believing in freedom. A Reconstruction and Comparison of The Interpretation of The Freedom of Religion and Belief on The Canadian Supreme Court, The South African Constitutional Court and the European Court of Human Rights*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023, ISBN: 978 94 6473 247 4
MI-409 R.M.S. van Es, *The mind in the courtroom. On forensic mental health reports in judicial decision-making about guilt and sentencing in the Netherlands*, (diss. Leiden) Alblasserdam: Ridderprint 2023
MI-410 I. Kokorin, *Intra-group financing and enterprise group insolvency: Problems, principles and solutions*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023
MI-411 J.T. Tegelaar, *Single Supervision, Single Judicial Protection? Towards effective judicial protection in Single Supervisory Mechanism composite procedures*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023
MI-412 A.M.H. van der Hoeven, *Met de vlam in de pijp door Europa. De arbeidssituatie van internationale vrachtwagenchauffeurs: constructies en percepties*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023
MI-413 H.T. Vethaak, *Empirical analysis of social insurance, work incentives and employment outcomes*, (diss. Leiden), Amsterdam: Ipskamp Printing 2023