From the onset of illness to potential recovery

Prudon, Roger Eric Karla

2024

DOI (link to publisher)
10.5463/thesis.335

document version
Publisher's PDF, also known as Version of record

From the onset of illness to potential recovery

Empirical economic analysis of health, disability and work

ISBN 978 90 361 0742 6

Cover photo: Generated from DALL-E 2
Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 835 of the Tinbergen Institute Research Series, established through cooperation between Rozenberg Publishers and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

VRIJE UNIVERSITEIT

From the onset of illness to potential recovery

Empirical economic analysis of health, disability and work

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philosophy aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de School of Business and Economics
op donderdag 29 februari om 13.45
in de aula van de universiteit,
De Boelelaan 1105

door

Roger Prudon

geboren te Maastricht

promotor:	prof.dr. P.W.C. Koning
copromotor:	dr. P. Muller
promotiecommissie:	prof.dr. V. Angelini
	prof.dr. W.H.J. Hassink
	prof.dr. M. Lindeboom
	prof.dr. A.H.O. van Soest
	prof.dr. N.R. Ziebarth

Acknowledgments

My PhD journey started on the 9th of February 2018, when I was still in the first year of my MPhil at the Tinbergen Institute, and received an e-mail from Pierre. Pierre, whom back then I still addressed as Professor Koning, asked whether I would be interested in a PhD position focussing on disability insurance. One year later I started working with Pierre and Paul on what would become Chapter 5 of this dissertation, and now, four years later, the dissertation is finished.

The acknowledgements to this dissertation should therefore undoubtedly start with a big thank you to my advisors, Pierre Koning and Paul Muller. Your guidance these past four years has shaped me from a student who had no clue how to do research, to an actual researcher. From the moment we started working together, I could always walk into your office to talk about anything. By being open and supportive, but also critical when necessary you have been excellent supervisors, co-authors and colleagues. Even though my dissertation is now finalized, I hope and think that this is only the beginning of our joint research efforts, as I have truly enjoyed working with you. Finally, I would also want to thank you, Pierre and Paul, together with Maarten for your support during the job market. Your tips and tricks, recommendations and letter writing made sure that I could actually enjoy being on the job market and land a job.

I also want to thank all of my fellow PhD students and colleagues. In particular my old office mates, Hermon, Rik and Zichen, my new office mates, Alessandro and Yuchen, my labor/health/applied econ colleagues Bas, Nadine, Nursena, Pieter, Sabien and Wendy, and the Work in Progress group at the UvA. Thank you for all of your critical questions and advice on the various chapters of this dissertation and for making the breaks at VU enjoyable.

The research of this thesis is part of the "Geschikt of ongeschikt voor arbeid?" research program of Pierre. Instituut Gak is gratefully acknowledged for the financial support of this program. The various projects are based on data provided by UWV, and I would thus also like to thank Carla van Deursen, Marloes de Graaf and Marloes Lammers for their help in understanding the data and the complexities of the Dutch social security system. Chapter 4 is inspired by a policy report written by Wim Zwinkels, who I would like to thank for sharing his original code.

To my high school friends, Rik, Nicky and Mischa, thanks for the endless board games and even longer lasting vastelaovond adventures. Despite meeting infrequently, our get-togethers have been a great distraction from academia. Rik, I would like to thank you in particular. We went from sharing a desk during economics in high school to sharing an office at the VU and in those thirteen years, our friendship has become indispensable to me. I could not ask for a better paranymph.

I am grateful for the unconditional support I received from my parents Pia and Maurice, my brothers Guy and Gaston and my sister Claire. Mom and Dad, by always fighting for your children, by being hardworking and by being sincere, you've inspired me to do research that (hopefully) has an impact. Guy, thanks for keeping up with my running adventures and for supporting me in this academic race as my second paranymph. I hope there are still many more races (and marathons) to come which we will finish together. And lastly a word for the newest addition to the family, Tuur. In the next couple of years, your mom can read bedtime stories to you from this book which are guaranteed to make you fall asleep. Once you are old enough to understand, I hope that you are proud of the work I do, just as I am proud to be your godfather.

Finally, I want to thank those who have endured the most during my PhD. Chelsea, thank you for always being there for me in good and bad times. You were able to slow me down when I was going too fast, while also taking care of everything whenever I was travelling again. I might not often express my gratitude vocally, but I am very grateful for everything you do. And lastly, to my furry friend who has been with me ever since I moved out ten years ago. I could not have imagined going on this journey without you.

Roger Prudon
Amsterdam, June 2023

Contents

Amsterdam, June 2023

Chapter 1

Introduction

In the 1980s, almost one in eight individuals in the Dutch labor force were receiving disability insurance (DI) benefits. High inflow rates and low outflow rates of the DI system were considered one of the main socioeconomic challenges and spurred a series of reforms between 1996 and 2006. These reforms have been effective in reducing annual inflow rates, from approximately 1.5% to 0.5% of the insured working population. At the same time, outflow rates have remained fairly constant. Despite the substantial reduction in inflow, a significant share (7 %) of the labor force still receives disability benefits in 2023 and the risk to remain on DI until retirement continues to be high. From a policy perspective, it is therefore important to understand which groups are most at risk of entering DI, what measures can be taken to prevent this, and how to support them to return to work.

To gain insights into these questions, this dissertation follows the process from the onset of health issues, through the DI application and award process until potential recovery and work resumption. In Chapter 2, the focus is on the onset of mental health problems and the timely access to treatments for these health problems. Next, I turn to two groups with relatively high probabilities of reporting sick and subsequently applying for DI benefits in Chapters 3 and 4. Specifically, Chapter 3 examines the DI application behavior of individuals with temporary contracts, while Chapter 4 focuses on individuals receiving unemployment insurance (UI) benefits. Finally, Chapter 5 looks into the potential recovery of health issues and subsequent outflow from the DI system.

For an overview of every step of the DI process, three administrative datasets are combined with comprehensive information about the characteristics of workers, their healthcare utilization and their social insurance and employment histories. The first dataset, provided by Statistics Netherlands (CBS), contains characteristics such as age, gender, nationality and education level for the entire Dutch population. This data is linked with employment histories and healthcare utilization, also provided by Statistics Netherlands.

The third dataset, provided by the Dutch employment agency (UWV), contains information on all applications for sickness insurance (SI) and DI. Quasi-experimental methods are applied to answer the various research questions.

The first step in the process towards a DI application is the onset of health issues. Successful treatment of these health issues is crucial in reducing the inflow into DI. Chapter 2 focuses on treatments for mental health issues. Workers with mental health issues make up an increasingly large share (40%) of DI applicants. Nevertheless, the provision of mental health treatment has been under intense pressure in the last decade. This is mirrored by a strong increase in waiting lists for these treatments. Chapter 2 of this dissertation investigates to what extent the delayed provision of mental health treatment affects employment and inflow into the DI system.

Exploiting plausibly exogenous variation in the congestion of the mental healthcare system across municipalities, the estimation results show that increased waiting time for mental health treatment leads to a substantial reduction in employment. A two-month (one standard deviation) increase in waiting time reduces the probability of being employed by 4 percentage points while increasing the likelihood to receive DI benefits by 2 percentage points. Albeit all groups in the population are susceptive to increased waiting times, the largest burden falls upon vulnerable groups with a low education level or a migration background. The impact of one additional month of waiting time is almost twice as large for them, and in addition, they have longer waiting times on average. Timely provision of (mental) health treatment could substantially reduce DI inflow, especially among lower-educated individuals and those with a migration background.

While Chapter 2 shows that health issues are important in understanding DI inflow, there are large differences in DI application behavior between groups with similar reported health issues. Chapters 3 and 4 analyse two specific groups that have a relatively high probability of applying for DI benefits: those with a fixed-term contract and those on unemployment insurance (UI). The first group, those with a fixed-term contract, are approximately 50% more likely to apply for DI benefits, compared to workers with a permanent contract. To better understand this risk differential, Chapter 3 decomposes the difference in DI application risk into contributions from the following four mechanisms that are relevant at distinctive moments in time; (1) differences in characteristics of workers with permanent and fixed-term contracts, (2) the causal impact of contract type on health, (3) employer incentives and obligations during the sickness period, and (4) labor market prospects during the sickness period. As to the latter explanation, the idea is that fixed-term workers are more susceptive to labor market conditions than permanent workers that are still formally employed, and therefore more likely to file a DI application. Previous research has found that these mechanisms likely impact DI applications, but their relative contribution to the high DI application risk of temporary workers is not yet known. Understanding the relative importance of the mechanisms is crucial to designing appropriate policies to lower the disability risk of temporary workers.

The results show that there are significant differences in the characteristics of workers with permanent and fixed-term contracts, especially in terms of their age and job type. However, these compositional differences do not explain the observed DI risk premium. On the contrary, if permanent and temporary workers would have similar characteristics, the DI risk premium would be even larger. Additionally, the probability to experience either a mental health shock or a physical health shock – measured as a strong increase in medical consumption – is very similar for both contract types. Accordingly, the second mechanism does not explain the DI risk premium either. In contrast, employer incentives during the sickness period, and labor market prospects of sick individuals do explain more than 80% of the observed DI risk premium. The decomposition shows that while two groups may have very similar probabilities of falling ill, they can still have very different probabilities of applying for DI due to economic incentives inherent with contract types.

A second group with a particularly high DI application risk premium are unemployment insurance (UI) recipients, who are approximately four times as likely to apply for DI as workers with fixed-term contracts. In particular, there is a clear spike in sickness reporting of UI recipients in the last month of their UI entitlement period. Chapter 4 zooms into this spike of inflow into sickness insurance (SI) and considers two potential explanations for this phenomenon. First, individuals whose UI entitlement is about to expire might consider SI as an option to extend the total duration of benefit receipt. In this case, the spike would be driven by rationally optimizing individuals with relatively mild health issues ("moral hazard"). Alternatively, the spike might also mirror a catch-up of initial non-take-up. That is, UI recipients with health issues might not be aware of their eligibility for SI benefits. These individuals receive a letter 6 weeks before the end of their UI entitlement period, informing them of the fact that they can call in sick if they are currently experiencing health problems and should do so before their UI eligibility terminates. Under this scenario, those in the spike (the "spike cohort") would have similar health issues, compared to those who call in sick earlier (the "pre-spike cohort").

To determine which of these two explanations is most relevant, Chapter 4 follows a two-step approach. In the first step, the spike-cohort and the pre-spike cohort are followed during the two-year sickness period that may result in a DI application. During these two years, there are two external screening moments so if the spike is driven by moral hazard, one would expect a larger share of the spike-cohort to be screened out during these screening moments. The findings suggest this is not the case: workers in the spike cohort are more likely to reach the DI application and subsequently be granted benefits than individuals in the pre-spike cohort. In the second step, compositional differences in demographic characteristics, health conditions and labor market outcomes prior to unemployment between the cohorts are examined. While the spike cohort appears to have a weaker labor market attachment before entering UI, their healthcare utilization is comparable to the pre-spike cohort and compositional differences do not explain the difference in DI inflow. High inflow probabilities at the end of UI eligibility therefore appear not to be driven by moral hazard, but by catch-up of initial non-take-up.

Chapters 2, 3 and 4 focus on workers that are at risk of applying for DI benefits. Chapter 5 considers workers that are enrolled in DI. A large share of all DI recipients in the Netherlands is assessed as being partially and/or temporarily disabled and are expected to recover and resume work at some point in time. Despite the series of reforms between 1996 and 2006, outflow rates out of DI have remained at a low level and the vast majority of recipients remain on DI until they reach their retirement age. This raises the question of to what extent the DI system hinders the return to the workforce once DI recipients recover. Chapter 5 answers this question by comparing the response to an improvement in health for individuals just above the DI threshold (those with partial DI benefits) to that of individuals just below this threshold (those without DI benefits). As both groups are close to the DI threshold, they are very similar in terms of their employment status prior to recovery. Once medical treatment ends (a proxy for recovery), their employment trajectories diverge. A difference-in-difference approach confirms that DI benefits do in fact reduce the return to work after recovery by about 50%.

Chapter 6 concludes and provides a summary of all chapters.

Chapter 2

Is delayed mental health treatment detrimental to employment?

2.1 Introduction

Over the last several decades, the prevalence of mental health problems has been high and increasing in most OECD countries. Approximately half of all individuals suffer from mental health issues at some point in their lifetime (Hewlett & Moran, 2014). The onset of mental health problems has been found to have large negative effects on employment, with a corresponding decrease of 10 to 30 percentage points in the probability that an individual is employed (Frijters et al., 2014). Conversely, appropriate mental health treatment has been found to be effective in mitigating these negative impacts (Biasi et al., 2021; Shapiro, 2022).

Despite this, many countries struggle with providing sufficient mental healthcare capacity, leading to decreased access to treatment in the form of long waiting times. The Covid-19 pandemic has aggravated this issue, as emotional distress increased while access to treatment was reduced or even eliminated due to lockdowns. As a result, waiting times have been increasing in many countries, such as the UK, Australia, the US, and the Netherlands (Campbell, 2020; Kinsilla, 2021; Caron, 2021; n.d., 2021). Given the time it can now take before appropriate treatment can start, a crucial question is whether treatment remains effective in reducing the negative impact of mental health problems on employment if individuals have to wait several weeks or months before it can commence. While some correlational evidence now exists that increased waiting times are associated with worse mental health outcomes (Reichert & Jacobs, 2018), causal estimates of the impact on employment are yet to be documented. This knowledge gap leaves policymak-

[1]This chapter is based on Prudon (2023).

ers in the dark when it comes to quantifying the micro- and macro-level consequences of inadequate mental healthcare provision.

In this chapter, I attempt to fill in this gap by investigating whether increased waiting times for specialized mental health treatment negatively affect labor market outcomes. Given that there exist large disparities in the propensity to receive mental health treatment between various demographic groups (Sentell et al., 2007), I additionally analyze whether certain groups are affected to a greater extent by these increased waiting times. To answer these questions, I use administrative data from the Netherlands regarding the usage of mental health treatments and the corresponding waiting times. Using anonymized citizen service numbers, I merge this with data on labor market outcomes and demographic characteristics at an individual level. The main analysis focuses on waiting times for any kind of mental health problem. As the spectrum of mental health issues is broad, I additionally distinguish between treatments for the four main categories of mental health issues: personality disorders, mood disorders, anxiety disorders, and other disorders. The time window under consideration runs from 2012 to 2019.

To begin with, I introduce a benchmark for the potential effects of increased waiting times. This benchmark is estimated using an event-study specification comparing individuals starting mental health treatment to those who do not undergo treatment. Receiving treatment could be seen as an imperfect proxy of experiencing mental health problems, and the benchmark estimates should thus be interpreted as the net effect of the onset of mental health problems and the subsequent treatment of these problems. Due to potential reverse causality and time-varying confounders, the resulting estimated correlations should not be taken at face value, but serve as a point of comparison for the remaining analyses instead. With this in mind, the onset of mental health problems is associated with a nine percentage point reduction in the probability that a given individual is employed two years after the start of treatment. The majority of those whose employment was terminated subsequently made use of sickness/disability insurance (seven percentage points), or social assistance (four percentage points).

Following this, I estimate the causal impact of increased waiting times on employment. This presents a methodological challenge: not only are waiting times likely endogenous with respect to the severity of a diagnosis, but individuals can also choose among providers based on expected waiting times. To account for this endogeneity, I instrument individual waiting time using regional waiting time. This IV approach exploits plausibly exogenous regional variations in the congestion of the mental health system as measured through regional waiting time on a municipality level. I subsequently find that a two-month (equal to one standard deviation) increase in waiting time decreases the probability of employment by approximately four percentage points while also increasing the probability of receiving sickness/disability benefits by two percentage points. Heterogeneity in the effect of waiting time is limited with respect to the type of diagnosis and gender, but the impact of increased waiting time is noticeably higher for individuals with a migration background and those with lower educational attainment.

Given the negative effects that I find for delayed treatment – particularly for certain vulnerable groups – a crucial follow-up question is to what extent an individual's access to mental healthcare is impacted by their demographic characteristics. The final part of this chapter therefore examines differences in waiting times based on gender, migration background, and educational attainment. As with the effects of delayed treatment, differences in waiting times based on gender are small, while those based on migration background and educational attainment are relatively large. Specifically, the average waiting time of individuals with a migration background is 7-11 days longer than that of individuals without a migration background. For less educated individuals, this gap is 3-13 days with respect to their higher-educated counterparts. These estimates are all on top of any differences in a rich set of observable characteristics, which include other demographics, job characteristics, provider fixed effects, and mental health diagnoses. It is important to stress that these differences in waiting time are therefore not caused by selection based on the municipality of residence, pre-treatment labor market status or differences in the severity of mental health problems.

The effects of increased waiting times are substantial relative to the changes in labor market status around the start of treatment. A two-month increase in waiting time results in a four percentage point reduction in the probability of employment, as compared to a "total" employment effect around the start of treatment of about nine percentage points. On top of this, vulnerable groups experience both longer average waiting times and larger negative effects, meaning that the differential impact of reduced access to mental health treatment could be substantial. If policymakers wish to protect economically vulnerable individuals and combat inequality, my results suggest that greater availability of mental health resources could be a valuable tool.

The results contribute to several lines of research. First, there is extensive literature on the effects of mental health problems on employment. By their very nature, most mental health problems are interrelated, with a wide range of both observable and unobservable characteristics and events affecting mental health. To perform causal inference, early studies used cross-sectional data with instruments based on early-life events. Examples of these instruments are parental psychological problems (Ettner et al., 1997; Marcotte et al., 2000; Chatterji et al., 2011), degree of religiosity, perceived social support, and participation in physical activity (Alexandre & French, 2001; Hamilton et al., 1997; Ojeda et al., 2010) and past mental health issues (Ettner et al., 1997; Hamilton et al., 1997;

Chatterji et al., 2007, 2011). The IV estimates of these studies point to a decrease in the probability of being employed by between 10 and 30 percentage points due to the onset of mental health problems. While these early-life events have a clear impact on mental health, they might also affect other aspects of an individual's life, such as their motivation or time preferences, potentially leading to biased IV estimates.

An exception in this strand of literature is a more recent study by Frijters et al. (2014) which uses panel data in which the death of a friend is used as an instrument for mental health. This instrument is less likely to violate the exclusion restriction, but the shock considered is specific and the impact on mental health is relatively small; the death of a close friend decreases mental health by on average 0.04 standard deviation. The authors' IV estimates indicate that a one standard deviation worsening of mental health decreases the probability of being employed by 30 percentage points.

Even more recently, attention has shifted to the effects of treatment for mental health problems on labor outcomes by using plausibly exogenous variation in the availability of pharmaceuticals. Biasi et al. (2021) show that the availability of lithium as a treatment for bipolar disorder reduced the earnings penalty of bipolar disorder by approximately one-third. Similarly, Shapiro (2022) finds that increases in the number of advertisements for antidepressants reduce workplace absenteeism significantly. This mitigating impact of treatment for mental disorders on employment suggests that there is a negative impact of mental health problems themselves on employment. I add to this literature in two ways. First, I use a broader notion of treatment which includes both the use of pharmaceuticals and psychotherapy. Second, I do not examine variation in the availability of treatment, but variation in the time individuals have to wait before receiving it.

A small but growing strand of literature focuses on waiting times for various treatments. At the time of writing, the only study to investigate the effect of waiting times for mental health treatment finds moderate effects (Reichert & Jacobs, 2018). However, this study only examines correlations between waiting time and mental health itself and does not consider labor market outcomes. The impact of waiting time for other medical treatments has been examined using a similar estimation approach as used in this chapter. Godøy et al. (2022) and Williams & Bretteville-Jensen (2022) estimate the causal impact of waiting times for orthopedic surgery and substance abuse treatment, respectively, on employment. Both studies use regional variation in waiting times as instruments to obtain causal estimates. Godøy et al. (2022) find no health effects, but strong employment effects of increased waiting time for orthopedic surgery. Williams & Bretteville-Jensen (2022) on the other hand, find both health and employment effects of increased waiting times for substance abuse treatment.

The effect of waiting times for non-medical treatment on employment has been considered by Autor et al. (2015) and Hauge & Markussen (2021). Autor et al. (2015) study increased processing times for disability insurance (DI) applications in the US and find that a 2.1-month (one standard deviation) increase in waiting time reduces the probability of employment by 3.5%. In contrast, Hauge & Markussen (2021) consider reduced waiting times for vocational rehabilitation programs for individuals on temporary DI in Norway, and find no significant effects of reduced waiting times. I add to this literature on waiting times by estimating the causal impact of waiting time for one of the most prevalent types of treatment, namely, treatment for mental health problems.

The final related strand of literature concerns inequality in both access to and use of mental healthcare. Previous research on this subject has mainly focused on the US context. There we see that large differences exist, with minority groups being up to 80% less likely to use mental healthcare (Sentell et al., 2007; Cook et al., 2017). Sentell et al. (2007) find that one of the main reasons for this reduced access is limited English proficiency, i.e., language barriers. However, little is known about differences in access conditional on seeking treatment. Furthermore, the differential impact of mental health problems on minority groups is also under-investigated. I fill these gaps in the literature by examining differences in waiting times conditional on seeking treatment, and by estimating differential impacts on various groups.

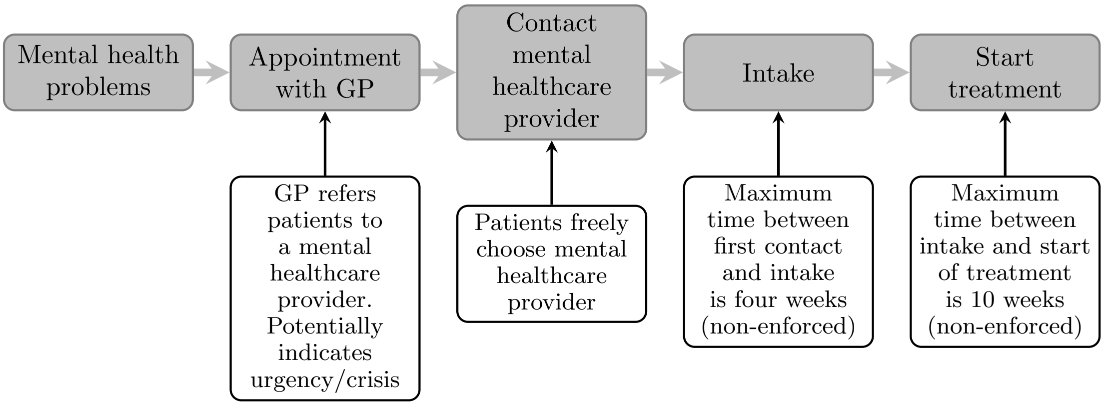
The remainder of this chapter is organized as follows: the institutional setting and the data are described in Sections 2.2 and 2.3. Sections 2.4, 2.5 and 2.6 discuss the analyses on the onset of mental health problems, the exacerbating effects of waiting times, and unequal access to mental health treatment respectively. Section 2.7 concludes.

2.2 Mental healthcare in the Netherlands

Figure 2.1 illustrates the process individuals in the Netherlands go through from the moment they experience mental health problems, until the start of their treatment. Mental health problems as discussed in this chapter range from mild depression to severe personality disorders. Treatment for all mental health problems is covered by universal health insurance. Individuals experiencing mental health problems first contact their general practitioner (GP). The GP is the gatekeeper of the mental healthcare system and makes the first assessment of the severity of mental health problems. In case of mild mental health problems, the GP can either decide to treat the individual within their GP practice or refer them to a provider of basic mental healthcare. If the problems are more severe, the GP will refer to specialized mental healthcare, which is the focus of this chapter. Individuals can receive some form of treatment from their GP while waiting for specialized mental healthcare.

GPs can influence the waiting time by indicating the urgency of the case. In cases

Figure 2.1: Timeline from the onset of mental health problems to the start of treatment



with high urgency, mental healthcare providers can schedule the intake sooner. In crisis situations, treatment starts as soon as possible (within a few days). A GP can refer to a specific care provider, but individuals are free to choose a different provider. To help individuals choose an appropriate mental healthcare provider, the government publishes general information about every provider, including average waiting times.

After an individual has contacted a mental healthcare provider, the intake takes place. During the intake, a first assessment is made of the (severity of) the diagnosis, and a treatment plan is made. After the intake, treatment commences as soon as the provider has the required capacity. In order to decrease the waiting time of patients, the Dutch Ministry of Health, Welfare and Sport has set norms for the maximum waiting times. Once an individual has contacted a mental healthcare provider, the intake should take place within four weeks and treatment should start within 10 weeks after the intake, implying a total waiting time of at most 14 weeks. Compliance with these norms is limited, as no immediate action is taken once the norms are exceeded. As shown in the next subsection, individual waiting times can be significantly longer than the norms.

2.3 Data

To obtain individual time series on mental healthcare usage and a range of labor market outcomes, several administrative datasets provided by Statistics Netherlands covering the entire Dutch population are linked. These time series are complemented with data on both individual- and municipality-level characteristics. Linkage of datasets was done using anonymized citizen service numbers.

2.3.1 Data on (mental) healthcare

The mental healthcare data contains all treatment-related specialized mental healthcare events occurring between 2011 and 2019. Mental healthcare treatment is defined as having real-life or virtual contact with a mental healthcare provider. Treatment could be a combination of some form of therapy and pharmaceuticals, but the use of pharmaceuticals is not reported in the data. For all treatment-related events, I observe the date of the event, the type of the event (first contact/intake/treatment/administrative, etc.), the number of contact minutes with a patient, the mental health diagnosis, the type of treatment provider (psychologist, psychiatrist or other) and anonymized identifiers for the patient and provider.

Table 2.1 shows the sample selection steps to obtain the final sample of individuals that started specialized mental health treatment and for whom waiting time is observed. Waiting times were calculated for all individuals who started mental health treatment between 2012 and 2019. Individuals receiving mental health treatment in 2011 are excluded, as it cannot be determined whether they started mental health treatment in 2011 or whether they were already being treated in 2010. Approximately one-and-a-half million individuals started mental health treatment in 2012-2019. Of these, waiting time is observed for 1,268,211 individuals. Excluding individuals with waiting times longer than one year or without an intake reduces the sample to 1,016,127 individuals. As this chapter focuses on labor market outcomes, only individuals within the working-age range of 18 to 65 are included. Finally, to ensure that individuals are not merely continuing previous treatments, I exclude all individuals with mental healthcare spending in the three years prior to the start of treatment. The final sample comprises 524,707 individuals who started mental health treatment in the period under consideration.

Figure 2.2 shows the distributions of time until intake (left) and total waiting time

Table 2.1: Sample selection steps

Inclusion criteria	Remaining sample
Start of mental health treatment 2012-2019[a]	1,537,153
Waiting time observed[b]	1,268,211
0 < Waiting time < 365 days[c]	1,062,563
Intake observed	1,016,127
18-65 years old during the first moment of contact	759,041
No mental healthcare spending in three years prior to treatment	524,707

(a) Individuals in the mental healthcare data in year t but not in year t-1; (b) Waiting time is observed if the first moment of contact is recorded and if there is at least one treatment activity; (c) Waiting times of more than one year are unlikely and could be caused by measurement error in the first moment of contact.

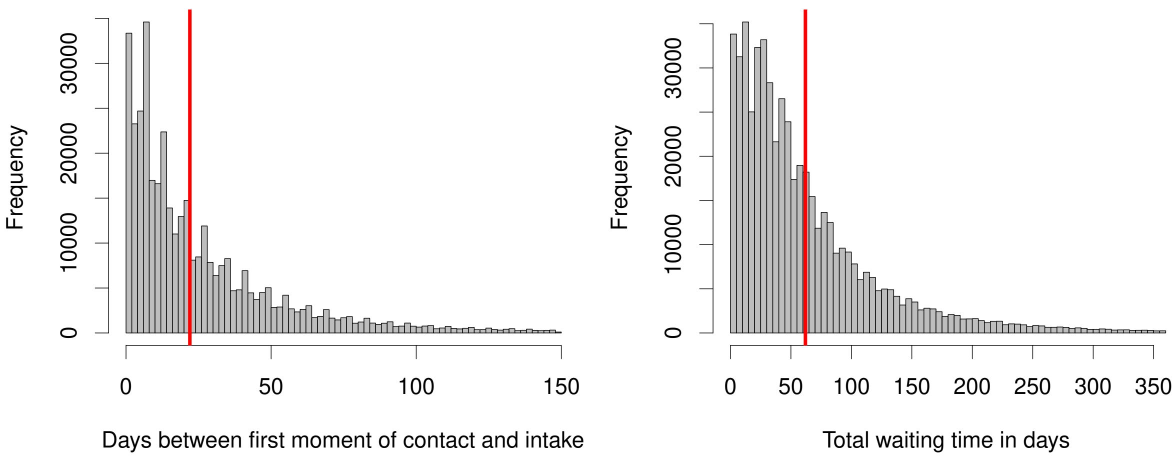
(right) for the entire sample. Time until intake is defined as the number of days between the first moment of contact and the intake while waiting time equals the number of days between the first moment of contact and the start of treatment. Both time until intake and waiting time are right-skewed with an average time until intake of 22 days, and an average waiting time of 62 days (averages in red). The observed waiting times are likely to be under-reported given that there is considerable bunching at time until intake of 0 days 27.3% of all observations, not shown in Figure 2.2). IV estimation should account for potential measurement error bias caused by this under-reporting of waiting times.[2]

A secondary source of data on healthcare usage is obtained through the health insurance system. Statistics Netherlands provides the yearly healthcare expenditures covered by basic health insurance for the years 2009 through 2020. Given the compulsory nature of health insurance in the Netherlands and its broad coverage, the data covers the vast majority of all healthcare. Spending is reported in various subcategories, which allows the distinction between mental and non-mental healthcare expenditures and spending on pharmaceuticals.[3] Furthermore, the data on mental healthcare spending also contains spending on basic mental healthcare, which is not included in the primary data on specialized mental healthcare. Healthcare expenditures are used as additional outcome measures to determine whether waiting time also has an impact on healthcare usage.

Figure 2.2: Distribution of time until intake (left) and waiting time (right) excluding observations with zero days until intake. Mean values indicated in red.

(a) Intake time: First contact until intake

(b) Waiting time: First contact until treatment



[2]Zero days of time until intake implies that the day of first contact and intake are identical which is highly unlikely. Official statistics furthermore show slightly higher average waiting times (NZA, 2021). As a robustness test, I alternatively exclude entries with zero days until intake; this yielded similar results to the primary analysis. This implies that under-reporting of waiting times does not bias the results.

[3]See Appendix Section 2.A for classification of healthcare spending categories.

2.3.2 Data on labor market outcomes

The labor market outcomes of the sample include monthly measures of employment, labor earnings, working hours, and the receipt of unemployment benefits, sickness/disability benefits, and social assistance. The labor market panel spans the period from 2004 up until 2021. Data on monthly labor earnings and working hours are available from 2009 onward. For all analyses, the time series are converted to time relative to the first moment of contact with a mental healthcare provider. I follow individuals starting six years prior to treatment until eight years after the start of treatment (for an unbalanced panel).

The panel data on health and labor market outcomes are enriched with administrative records from Statistics Netherlands on the year of birth, gender, migration background, level of education, and the municipality of residence. For all 422 municipalities, I observe the distribution of income, the proportion of inhabitants receiving various social benefits, real-estate characteristics, population densities, the ethnic background of the population, and gender.[4]

2.3.3 Descriptive statistics

The first column of Table 2.2 shows descriptive statistics of the individuals starting mental health treatment. For comparison, the second column shows descriptive statistics of the full Dutch population aged between 18 and 65 who do not receive any mental health treatment between 2009 and 2019 and the third column shows statistics of a sample matched one-to-one based on the propensity to start mental health treatment. The propensity score is estimated using only the demographic characteristics of the individuals. The matched sample will be used as a comparison group in the analysis of the effects of the onset of mental health problems. I discuss the matching procedure in detail in Section 2.4.

Individuals receiving mental health treatment are on average younger than the rest of the population, which is mainly caused by a high prevalence of mental health problems for individuals aged 20 to 40. Furthermore, individuals receiving mental health treatment are more likely to be Dutch natives, and they tend to have completed a lower level of education.[5] By construction, the matched sample is almost identical to the treatment sample in terms of demographics. As expected, the treatment population has high mental healthcare spending, but their spending on both non-mental healthcare and pharmaceuti-

[4] The population size of Dutch municipalities ranges from approximately 1,700 to one million inhabitants, with an average population size of approximately 44,000.

[5] The large percentage of unknown education level in the general population is due to the fact that the Dutch education registry started in the 1980s. The education level is unknown for most individuals in cohorts that graduated earlier. The difference in unknown education level between the sample with and without mental health treatment is mainly driven by the age difference.

Table 2.2: Descriptive statistics of individuals starting treatment, a matched sample not receiving treatment and the general Dutch population not receiving treatment

	Start MH treatment 2012-2019[a]	Matched sample[b]	No MH treatment 2010-2019[c]
Demographics[d]:			
Age	37.8	37.8	42.9
Female	54.0%	53.9%	53.8%
Dutch native	72.8%	73.1%	68.4%
Education unknown	20.9%	20.9%	47.5%
Education[e]:			
Low	22.9%	22.9%	18.7%
Middle	43.4%	43.1%	39.6%
High	33.7%	34.0%	41.8%
Annual healthcare expenditures[f]:			
Mental healthcare (in €)	4,582	28	27
Physical healthcare (in €)	2,082	989	1,026
Pharmaceuticals (in €)	31	16	14
Mental healthcare treatment			
Main diagnosis:			
Mood	29.5%		
Anxiety	22.7%		
Personality	8.4%		
Other	39.5%		
Treatment provider:			
Psychologist[h]	43.2%		
Psychotherapist[h]	12.2%		
Psychiatrist[h]	20.1%		
Other[h]	24.5%		
Crisis	4.6%		
Treatment minutes[g]	117.2		
Number of individuals	524,707	524,707	14,674,592

(a) All individuals who start mental health treatment between 2012 and 2016 aged 18-65; (b) All individuals in the Dutch population who do not receive any mental health treatment between 2012 and 2019 aged 18-65; (c) Sample of the Dutch population who do not receive mental health treatment between 2012 and 2019, matched one-to-one on the propensity to follow treatment with the mental treatment sample; (d) Demographics on January 2014; (e) Education level if known; (f) Yearly healthcare expenditures in the year of first contact with a mental healthcare provider; (g) Average number of treatment minutes in the first month of treatment; (h) Treated by a psychologist or psychiatrist during the first contact.

cals is also almost twice as high as that of the other samples. This indicates the presence of co-morbidities and/or the interplay between mental and non-mental health.

To understand the impact of mental health problems and the delays in receiving treatment, it is instructive to examine which mental health problems individuals face and what treatment entails for them. The sample of individuals starting mental health treatment covers the full spectrum of mental health problems. The majority of them are diagnosed with mood disorders (29.5%) or anxiety disorders (22.7%), while personality disorders (8.4%) are less common. The remainder of the sample (39.5%) are diagnosed with some other disorder. The majority of all patients (43.2%) are treated by psychologists, while 12.2% and 20.1% are treated by psychotherapists and psychiatrists. Psychiatrists are allowed to prescribe medication and often treat more severe mental health problems, while psychotherapists and psychologists are not allowed to prescribe medication. Approximately one in twenty individuals that start treatment is reported to be in a crisis situation. These individuals are fast-tracked and treatment usually starts within days after the first moment of contact. In general, they receive more intensive treatment.

The intensity of monthly treatment decreases as treatment progress. The average number of treatment minutes is 117 in the first month and decreases to 37 and 19 minutes after one and two years respectively. The decrease in treatment minutes is mainly driven by a decrease in the number of individuals who continue treatment (extensive margin) and not by a decrease in the number of treatment minutes per treated individual.[6]

2.4 The association between onset and treatment of mental health problems and employment

As a benchmark for the effect of waiting times, I first provide a correlational measure of the net effect of mental health problems and their subsequent treatment on labor market outcomes. By their very nature, mental health problems are interrelated with a wide range of observable and unobservable characteristics and events. As discussed in the introduction, previous literature has used early-life events as instruments for mental health. While these events have a clear impact on mental health, they also affect other aspects of an individual's life, such as motivation or time preferences, potentially leading to biased IV estimates.

Given the scarcity of convincing instruments for mental health, I used an event-study approach in which individuals undergoing mental healthcare treatment were compared to individuals not receiving treatment. This approach thus shows the net effect of both the

[6]See Appendix Figure 2.A.1 for the distribution of treatment minutes at the start of treatment and after one and two years.

onset of mental health problems and their treatment. By using an event-study setup, I am able to control and test for pre-treatment differences caused by unobserved confounders. However, the event-study setup does not control for reverse causality or time-varying unobserved confounders, and the resulting estimates should therefore not be interpreted as causal effects. Instead, the estimates will be used to benchmark the effects of waiting times for mental health treatment by indicating what effects would be expected given average waiting times.

2.4.1 Methodology

The event-study compares individuals starting mental health treatment to a control group that does not undergo treatment. As shown in Table 2.2, individuals receiving mental health treatment are different from individuals not receiving treatment in terms of their age and gender. I therefore construct a control group using one-to-one matching on the propensity to start mental health treatment.[7] The propensity is estimated based on the municipality of residence, gender, age, migration background, and education level. The matched sample is very similar to the treatment sample in terms of these demographics but very different in terms of healthcare usage, as shown in Table 2.2. Comparisons to either the full population (no matching) or comparisons to the siblings of the patients, as proposed by Biasi et al. (2021), yield similar results.

To avoid comparisons between not-yet-treated and already-treated units, I use time relative to the first moment of contact with a mental healthcare provider.[8] For individuals in the control group, the counterfactual first moment of contact is not observed. I therefore use the first moment of contact of the matched treatment individual. Given that individuals are matched on demographics, the baseline specification does not include these characteristics as control variables.[9] The time window ranges from 72 months prior to the first moment of contact until 96 months after the first moment of contact. The event-study specification looks as follows;

$$E_{it} = \alpha_t + \sum_{l=-71}^{96} \beta_l MH_i \, I_{t=l} + \varepsilon_{it} \tag{2.1}$$

[7]Matching directly on all observable demographics yields similar results.

[8]Recent literature has shown that using calendar time and a two-way fixed effects estimator can lead to biased results in cases of staggered treatment implementation or dynamic treatment effects Goodman-Bacon (2021); Callaway & Sant'Anna (2021); Borusyak et al. (2021). By using time relative to the first moment of contact, a single treatment group (those starting mental health treatment) is compared to a single control group that is never treated (those never receiving mental health treatment) and thus these concerns do not apply (see for example Baker et al. (2021)).

[9]Including observable characteristics as control does not affect the β_l estimates as the control variables do not change over time.

in which i subscripts the individual and t denotes the time relative to the first moment of contact (with $t = 0$ being the month of the first moment of contact). E_{it} is a labor market status outcome, MH_i is an indicator for receiving mental health treatment, and $I_{t=l}$ indicates whether an observation is in month l relative to the first moment of contact. α_t captures the evolution over time for individuals who do not receive mental health treatment while β_l, the parameters of interest, capture deviations over time for individuals who do receive mental health treatment. β_l runs from 71 months prior to the first moment of contact until 96 months after the first moment of contact. The difference between those who do receive mental health treatment and those who do not is thus normalized to zero at month -72.

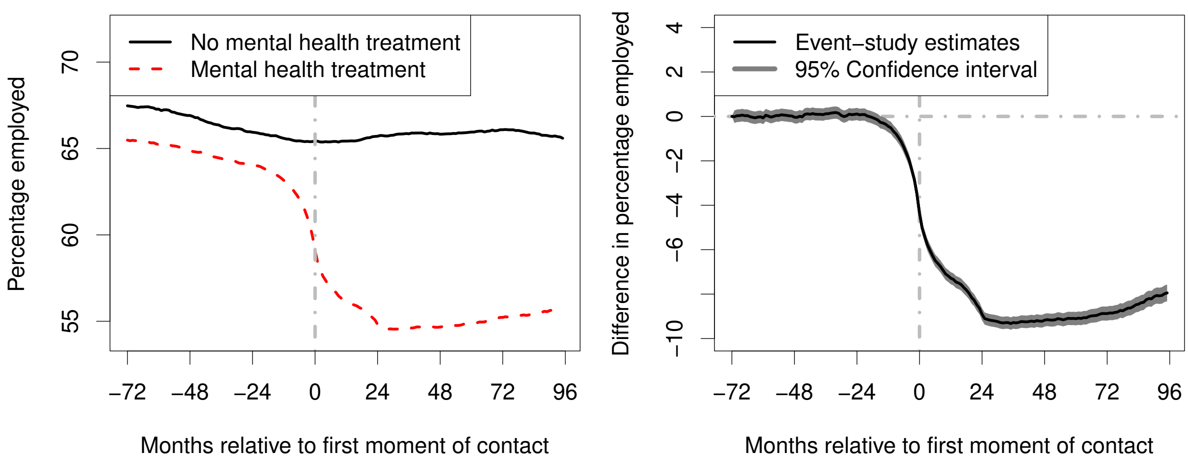
2.4.2 Results

The left panel of Figure 2.3 shows the employment rates for the group who start mental health treatment, and the matched control group relative to the first moment of contact. The figure on the right shows the corresponding event-study estimate, i.e., the difference between the two groups. Figure 2.4 show similar estimates for the alternative labor market outcomes.[10] Even though the treatment and control groups are matched based on propensity scores, the pre-treatment labor market status of individuals receiving mental health treatment is significantly different from the pre-treatment labor market status of individuals not receiving treatment. Six years prior to the first moment of contact, individuals in the treatment group are less likely to be employed but more likely to receive

Figure 2.3: Probability of employment (left) and corresponding event-study estimates (right) comparing individuals with and without mental health treatment

(a) Employment rates

(b) Event-study estimates



[10]Trends of these outcomes, equivalent to Figure 2.3 (a), can be found in Appendix Figure 2.A.2.

various social benefits. The trends furthermore show that most of these differences become larger in the years leading up to the first moment of contact.

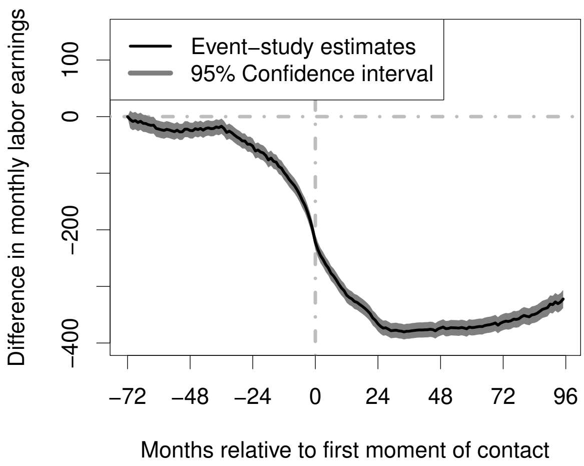
The level difference between both groups points to differences in unobservable characteristics. Furthermore, the divergence of trends could be driven by a number of reasons. First of all, the onset of mental health problems happens prior to the start of treatment. The divergence could however also be driven by reverse causality: a deterioration of labor market status could have a negative effect on mental health. Additionally, there could be unobserved time-varying confounders affecting both mental health and employment. Reverse causality and unobserved confounders would bias the event-study estimates and these estimates should therefore not be interpreted as causal effects of the onset of mental health problems on labor market status. However, since these factors would most likely upwardly bias the estimates, the event-study estimates can be used to obtain the upper bounds of the causal effects.

I find that the onset of mental health problems and the treatment of these problems is associated with a nine percentage point drop in employment. This estimate is close to the lower bound of estimates found in the IV studies, which range between 10 and 30 percentage points (Frijters et al., 2014; Ettner et al., 1997; Chatterji et al., 2011). The drop in employment rate corresponds to a drop in monthly labor earnings of approximately €400 and a 20-hour drop in the monthly number of hours worked (see Figure 2.4 (a) and (b)).

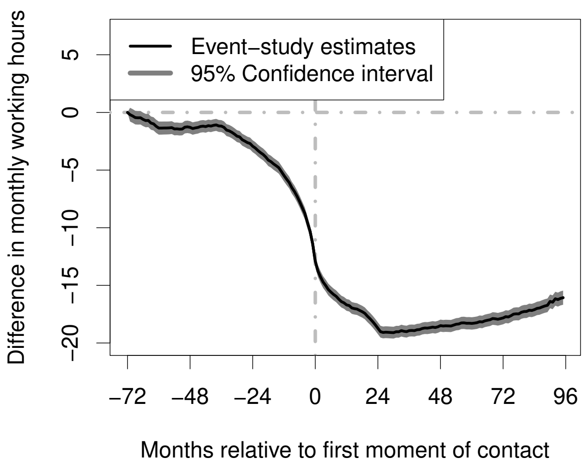
The drop in the employment rate is mirrored partly by an increase in the probability of receiving unemployment benefits of 1.5 percentage points (see Figure 2.4 (c)). The probability of receiving UI benefits drops shortly after the first moment of contact, caused by an inflow into sickness/DI benefits. The onset of mental health problems leads to a seven percentage point increase in the probability of receiving sickness and disability benefits (Figure 2.4 (d)). The increase in sickness and disability benefits is of a slightly smaller magnitude than the decrease in the employment rate. A similar pattern emerges for the probability of receiving social assistance, with an increase of approximately four percentage points (Figure 2.4 (e)).

Figure 2.4: Estimated event-study estimates and 95% CI comparing individuals with and without mental health treatment

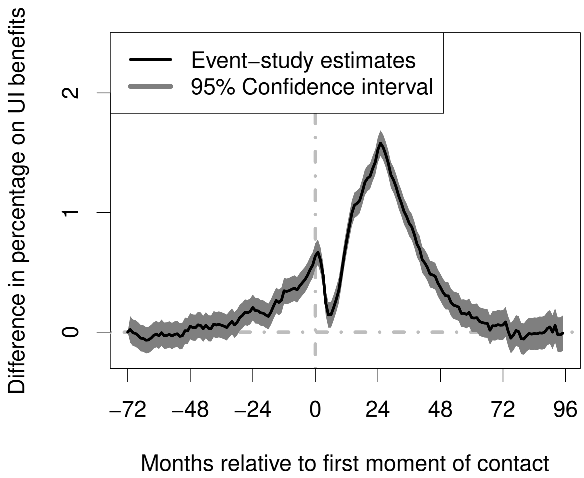
(a) Monthly labor earnings



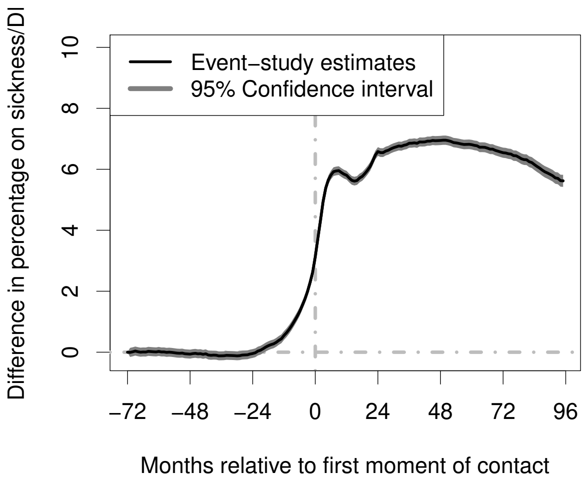
(b) Monthly working hours



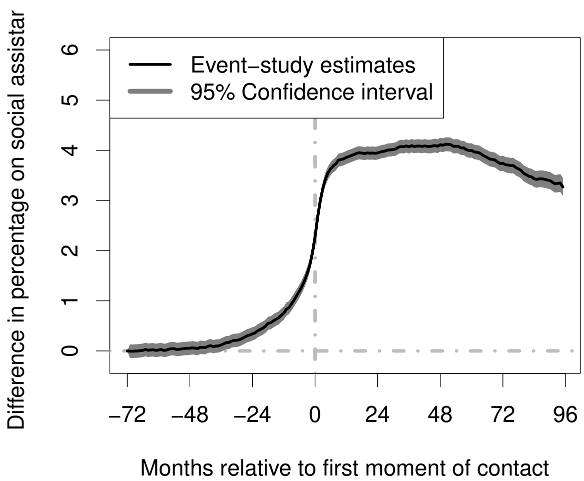
(c) UI benefits



(d) DI benefits



(e) Social assistance



2.4.3 Heterogeneity analysis

The estimates reported above are based on all individuals that start some form of mental health treatment in the Netherlands. I now investigate whether the impact differs for different groups of the population and for different types of mental health problems. I re-estimate the event-study specification by gender, age, migration background, and education categories. I also examine differences based on the mental health diagnosis and the type of provider. Treatment can be provided by psychologists, psychotherapists, psychiatrists, and other providers. The type of provider might signal the severity of the underlying condition being treated.

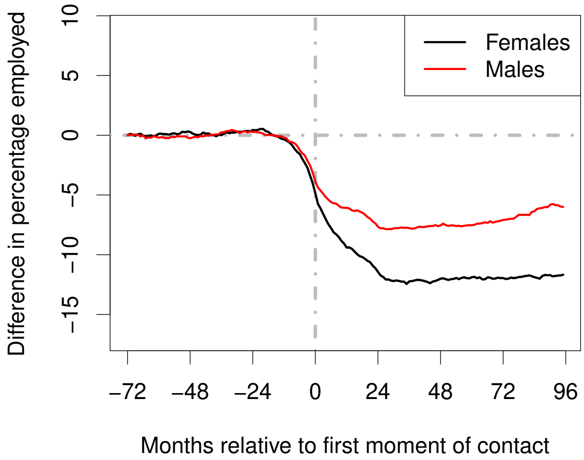
Figure 2.5 shows the event-study estimates of the impact of mental health problems on employment for the various subsamples.[11] Females and individuals with a migration background experience larger drops in employment around the first moment of contact with a mental healthcare provider, while heterogeneity by age and education level is limited. Additionally, there exists significant heterogeneity by mental health status. Treatment for anxiety and personality disorders result in the largest drops in employment. Heterogeneity by treatment provider is as expected; individuals being treated by psychiatrists and other providers experience more severe employment drops, in accordance with the fact that these providers tend to treat individuals with more severe mental health problems than psychiatrists and psychotherapists.

Summing up, the onset of mental health problems and subsequent treatment of these problems is associated with a decrease in employment of approximately nine percentage points and an increase in the probability to receive sickness/disability benefits and social assistance of approximately half that amount. Effects are larger for females than for males while individuals with a migration background are affected the most. As expected, being treated for more severe mental problems is associated with larger drops in employment.

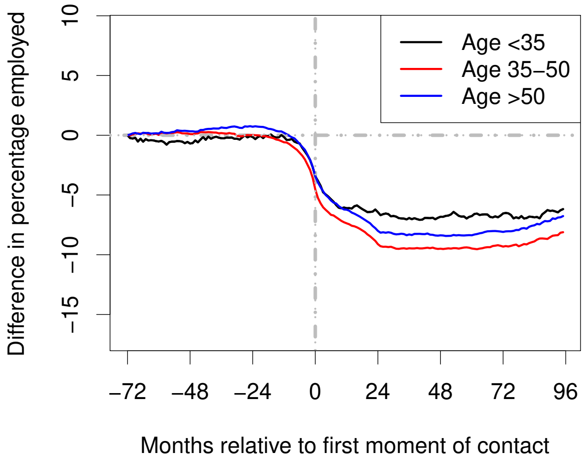
[11] Figures including confidence intervals are available upon request. Given the large sample size, the estimated impacts are almost always significantly different for the various groups.

Figure 2.5: Heterogeneity of event-study estimates on employment comparing individuals with and without mental health treatment

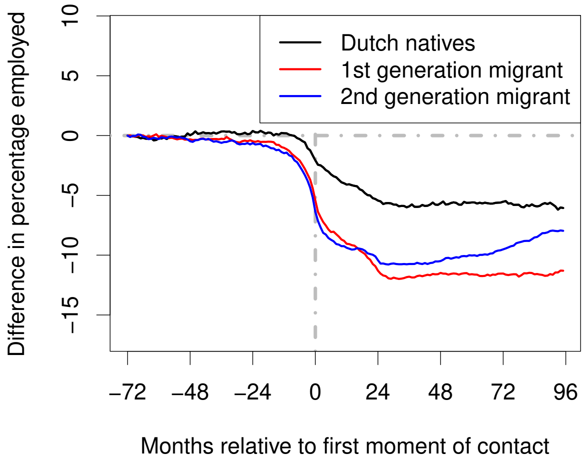
(a) Gender



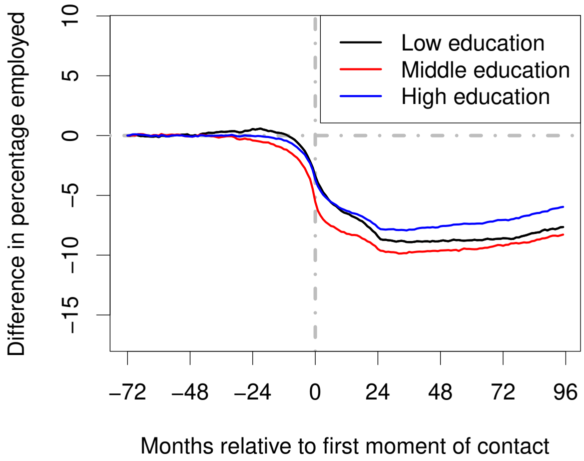
(b) Age



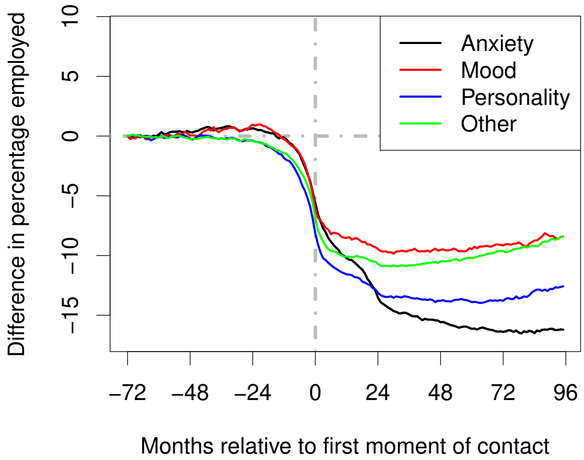
(c) Migration background



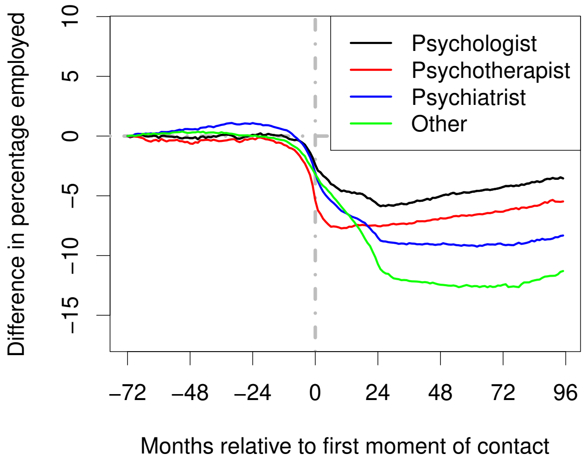
(d) Education level



(e) Main diagnosis



(f) Provider



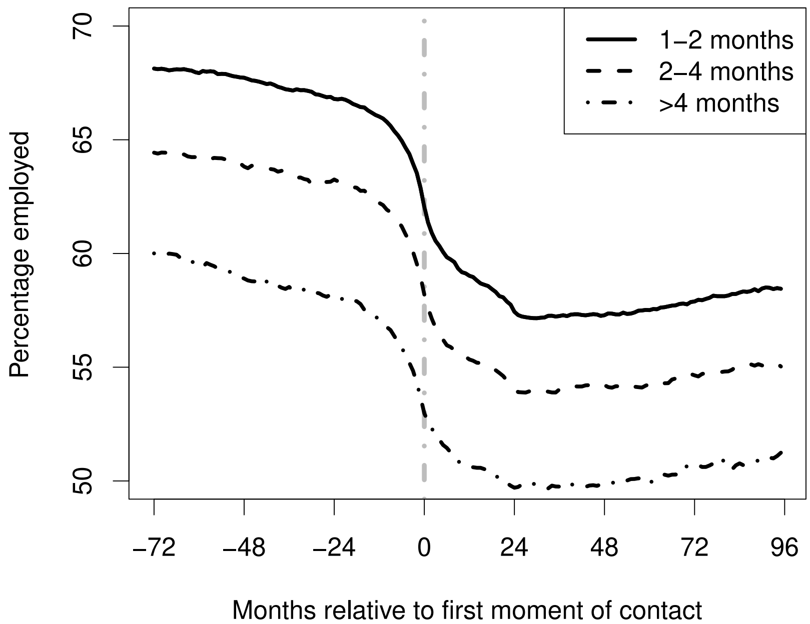
2.5 The impact of increased waiting times for mental health problems

The negative impact of the onset of mental health problems on labor market status shown in the previous section is the impact averaged over the entire waiting time distribution. It encompasses individuals whose treatment started a week after the first moment of contact and those who had to wait several months.

Given the process between the onset of mental health problems and the start of treatment as described in Section 2.2, waiting times are likely to be endogenous due to a number of reasons. First of all, individuals can freely choose mental healthcare providers and some individuals might base their decision on reported average waiting times. Furthermore, GPs can indicate crisis or urgency on the referral, fast-tracking patients with more severe mental health issues. Lastly, the severity of the mental health problems is partly determined during the intake. Based on the severity of the mental health problems, individuals might have to wait longer or shorter until treatment starts.

To illustrate the endogeneity of individual waiting times, Figure 2.6 shows the employment rate relative to the first moment of contact for three groups with different waiting times.[12] There is a drop in the employment rate around the first moment of contact, corresponding to the impact of mental health on employment as discussed in the previ-

Figure 2.6: Probability to be employed relative to the first moment of contact for three groups based on their individual waiting time



[12]Figures showing the trends of the other employment outcomes are shown in Appendix Figure 2.A.3

ous section. The drop in employment rate is similar for individuals with different waiting times but there is a level difference between the groups; individuals with longer individual waiting times have a lower probability to be employed, both prior to and after the first moment of contact. This holds when looking at raw averages (as in the figure), but also when controlling for a wide range of demographics. Individual waiting time thus correlates with both observable and unobservable characteristics, which also correlate with the probability to be employed. This implies that OLS estimates are biased and should not be interpreted as causal.

2.5.1 Methodology

To estimate the causal impact of waiting times, regional waiting time is used as an instrumental variable for individual waiting time. The intuition behind this instrument is that even though individuals can potentially choose mental healthcare providers based on expected waiting times, they are likely to choose providers within their region.[13] Longer regional waiting times should therefore result in longer individual waiting times, without being correlated to –for example– the severity of the individual's mental health problems. The IV approach exploits plausibly exogenous variations in the congestion of the mental healthcare system, as measured through regional waiting time at the municipality level. The first and second stage of the IV model looks as follows:

$$IW_i = \alpha_1 + \alpha_2 RW_i + \alpha_3 X_i + \alpha_4 R_i + \varepsilon_i \tag{2.2}$$

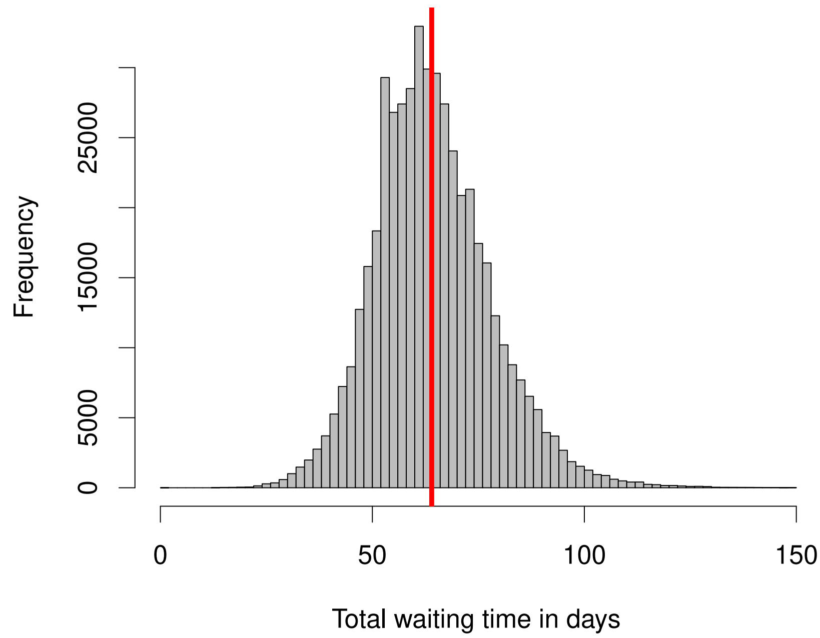
$$E_i = \beta_1 + \beta_2 \widehat{IW}_i + \beta_3 X_i + \beta_4 R_i + \mu_i \tag{2.3}$$

with IW_i and RW_i individual and regional waiting times, X_i individual characteristics, R_i regional characteristics (or regional fixed effects) and E_i the outcome of interest.[14] The following six labor market outcomes will be used; (1) employment, (2) monthly labor earnings, (3) monthly number of working hours (4) sickness/disability benefits, (5) UI benefits, (6) social assistance. Furthermore, I will also estimate the impact of waiting time on several measures of healthcare usage. The outcome of interest is measured at a specific point in time relative to the first moment of contact. The time window used starts six years prior to the first moment of contact and ends eight years after the first moment of contact. The IV estimates prior to the first moment of contact can be used as placebo tests for the exclusion restriction. Given that treatment has not commenced yet, waiting time should not have any effect on employment and the estimates should be close to zero.

[13]Individuals choosing mental healthcare providers in a different region weakens the first stage of IV, but do not bias the second-stage estimates.

[14]See Appendix Table 2.A.2 for a list of all control variables.

Figure 2.7: Distribution of leave-one-out regional waiting time. Mean indicated in red.

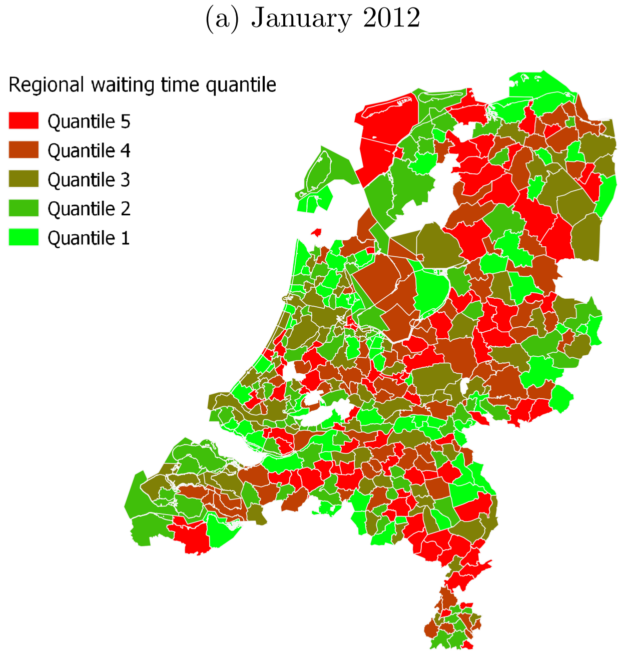


The average regional waiting time of an individual is computed using the leave-one-out principle. It is equal to the average waiting time of all individuals in that region who contacted a mental healthcare provider in the previous three months, excluding the individual under consideration.[15] Regions are defined at the municipality level, resulting in a total of 422 regions. The distribution of the leave-one-out regional waiting times is shown in Figure 2.7. Regional waiting times are almost symmetrically distributed, with a mean of 62 days and a standard deviation of 14 days. Further analyses show that variation between and within regions is mainly caused by variation in the number of individuals who terminated treatment in the preceding months. An increase in the number of individuals who stop treatment creates room for new treatments to start, reducing waiting times.

The IV approach exploits variation in regional waiting time between regions, and/or variation within the same region over time. To illustrate that there is indeed variation in both dimensions (time and region), Figure 2.8 shows a heatmap of the regional waiting time for all Dutch municipalities. Panel (a) shows a snapshot of January 2012, while panel (b) shows regional waiting times in February 2012.[16] Coloring is based on five quantiles of regional waiting time. Regions depicted in white have less than 10 individuals starting treatment in that month. For these region-month observation, regional waiting time is not computed. Bright red indicates a region has an average regional waiting time in the highest quantile (long waiting times), while bright green indicates a region with a regional

[15]Using regional waiting time using a time window of one, two or four months gives similar results.
[16]A timelapse of regional waiting time between 2012-2019 can be found on rogerprudon.com/research.

Figure 2.8: Heatmap of regional waiting time in January (left) and February (right) 2012



waiting time in the lowest quantile (short waiting times). As can be seen in the figure, there is indeed variation in time and variation between regions.

The IV specification assumes a linear relationship between waiting time and labor market status outcomes. To determine whether a linear relationship is likely, Appendix Figure 2.A.4 shows non-parametric estimates of the association between employment, as measured 12 months after the first moment of contact, and waiting time. For waiting times up to approximately 200 days (28 weeks), a linear specification seems valid.

To be able to interpret the obtained estimates as causal, regional waiting time should only influence labor market outcomes through individual waiting times. The next subsection discusses potential violations of this exclusion restriction and presents various tests. The estimates of the impact of waiting time are presented in the subsequent subsections.

2.5.2 Potential violations of the exclusion restriction

A potential concern with using regional waiting times as an instrument is that regions with longer waiting times could be different from regions with shorter waiting times. The regions might have different living- and labor-market conditions, potentially violating the exclusion restriction. To account for differences between regions, two different specifications are used. The first specification controls for a wide range of regional characteristics. By doing so, similar individuals in similar regions are compared, while exploiting variation both between regions and within regions over time. In the second specification, regional fixed effects are included instead of regional controls. By doing so, similar individuals in the same region at a different point in time are compared to each other, solely exploiting variation over time. Including regional fixed effects instead of regional controls increases the standard errors significantly as it only uses variation in waiting times within regions. At the same time, however, it eliminates any potential endogeneity based on unobserved differences between regions. Given their respective advantages and disadvantages, both specifications are used.

A second concern might be that changes in regional waiting time are driven by local labor market shocks. If these shocks directly affect the mental health of the population, IV estimates will be biased. This issue is specific to mental health and less relevant for the treatment discussed by Godøy et al. (2022), as the underlying health issues are less likely to be caused by employment shocks. To test whether local labor market shocks affect regional waiting time, Appendix Table 2.A.3 shows the estimated impact of the (lagged) (un)employment rate in a region on the regional waiting time in that region. The (un)employment rate in a region is not significantly associated with the regional waiting time. Including more lags (or leads) of the unemployment rate gives similar results. To further rule out that estimated effects are driven by local labor market shocks, current and lagged regional employment rates are included as controls in the IV regressions. The inclusion of these controls does not affect the IV estimates, confirming that the results are not driven by local labor market shocks.

Additionally, the exclusion restriction would also be violated if regional waiting time acts as a gatekeeper for the mental healthcare system. Longer waiting times might deter relatively healthy individuals from seeking treatment, resulting in differences in the composition of patients flowing into the mental healthcare system. The first way of testing this is assessing whether the probability to have an intake and/or actually start treatment, conditional on contacting a mental healthcare provider, is affected by regional waiting time. If regional waiting time were to act as a gatekeeper, one might expect that – in regions with longer waiting times – more individuals would flow out of the mental healthcare system before starting treatment, increasing the fraction of patients who contact the mental healthcare provider without starting treatment. To test whether this is the case, I estimated the impact of regional waiting time on the probability of either reaching the intake or starting actual treatment, conditional on contacting a mental healthcare provider (See Appendix Table 2.A.4). Regional waiting time has no significant effect on either of these measures.

A second, more direct way of testing whether increased waiting time acts as a gatekeeping mechanism is by examining the composition of patients that contact mental healthcare providers. If, for example, long waiting times would deter relatively healthy individuals from starting mental health treatment, then the average health of those individuals who did start treatment would be worse. The second panel of Appendix Table 2.A.4 shows only small correlations between regional waiting time and the composition of patients starting treatment, both in terms of demographics and in terms of the type of mental health diagnosis/treatment provider.

A final potential problem is that regional waiting time might affect the extent to which individuals try to reduce the waiting time. If this would be the case, the monotonicity assumption could be violated. The most straightforward way to reduce the waiting time would be to search for a provider with a shorter waiting time. Unfortunately, it cannot be inferred which provider individuals would go to if they would not try to reduce their waiting time. However, if individuals would broaden their search for providers with shorter waiting times, the probability of going to a provider outside of one's municipality of residence should increase. This provides an indirect test on the assumption of monotonicity. Regressing regional waiting time on an indicator for seeking care outside of one's municipality of residence shows that this is not the case.

2.5.3 First stage: the impact of regional waiting time on individual waiting time

The second assumption of the IV estimation concerns the strength of the instrument. To assess the strength, Table 2.3 shows the first-stage estimates. A one-day increase in regional waiting time, on average, increases individual waiting time by 0.4 days. The inclusion of regional fixed effects instead of regional controls does significantly decrease the estimate. However, even after the inclusion of regional fixed effects, regional waiting time still has a large and significant effect on individual waiting time. This is also reflected in the large F-statistics for both first-stage estimates.

Table 2.3: First-stage estimates of the impact of regional waiting time on individual waiting time using regional controls or regional fixed effects

	Individual waiting time	Individual waiting time
Regional waiting time	0.396**	0.289**
	(0.011)	(0.012)
F-statistic	1200	531
Regional controls	X	
Regional fixed effect		X

Standard errors shown in parentheses; *significant at a 10% significance level; **significant at a 5% significance level; The reported F-statistic compares models with and without the instrument.

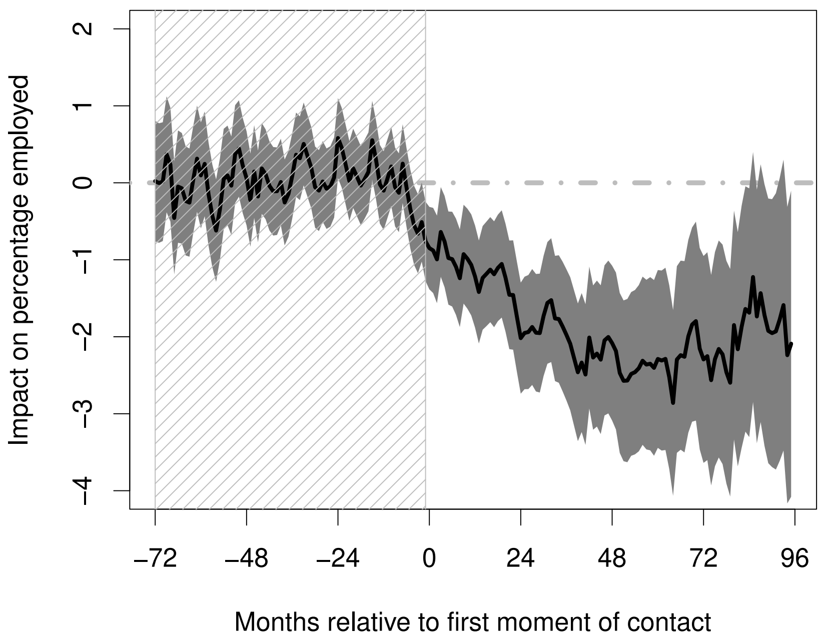
2.5.4 Second stage: the impact of waiting time on employment and healthcare usage

Turning to the causal impact of increased waiting time, Figure 2.9 shows the IV estimates and the corresponding 95% confidence intervals of one additional month of waiting time on employment. Note that the estimates are not event-study estimates, but are obtained through repeated IV-estimation instead. The estimate at time t is obtained through regressing instrumented individual waiting time on the employment status of an individual at time t.

The estimates for the six years prior to the first moment of contact, highlighted in grey, are placebo estimates; the outcome is measured prior to the first moment of contact, and waiting time should therefore not have any effect. As discussed, OLS estimation does yield significant placebo estimates, signifying the correlation between pre-treatment labor market status and individual waiting time as shown in Figure 2.6.[17]. In contrast, the IV placebo estimates do not differ significantly from zero, increasing the credibility of the IV approach.

The impact of waiting time starts to show in the four months prior to the first moment of contact, but this is likely due to measurement error. As discussed in Section 2.3, for a considerable share of the sample the moment of intake is also the first moment of

Figure 2.9: Estimated impact and 95% CI of one additional month of waiting time on probability to be employed



[17]The OLS estimates are shown in Appendix Figure 2.A.5

contact. Given that it is impossible to have an intake on the day in which individuals contact a mental healthcare providers, the actual first moment of contact is likely to be earlier. Measurement error in the first moment of contact implies that some individuals are already waiting in the months prior to $t = 0$, potentially explaining the observed effects in the months preceding the observed first moment of contact.

After the first moment of contact, increased waiting time has a negative and significant effect on the probability of being employed. A one-month increase in waiting time reduces the probability of employment by approximately two percentage points.[18]. The effects persist for at least eight years. OLS estimation points to a significant, but smaller negative effect of waiting time of approximately 0.3 percentage points (see Appendix Figure 2.A.5). OLS thus underestimates the negative effect of waiting time, as it does not take into account that individuals with more severe mental problems, and hence worse labor market outcomes, are more likely to receive treatment quicker. Using regional fixed effects instead of regional controls yields similar point estimates, but slightly larger confidence intervals given that less variation is used (See Appendix Figure 2.A.7). The similarity between the point estimates of the two specifications implies that the estimated impacts using regional controls are not driven by unobserved differences between regions.

Figure 2.10 shows the estimated impact on other labor market outcomes. The two percentage point reduction in the probability to be employed translates into a reduction in monthly labor earnings of approximately €100 (panel (a)), and the average number of working hours per month is reduced by approximately three (panel (b)). These impacts are comparable to the impact on the probability to be employed. This indicates that employment is mostly affected on the extensive margin. Panels (c)-(e) show that individuals whose employment is terminated flow into DI and social assistance, while the inflow into UI is unaffected. The effects persist for at least eight years.[19]

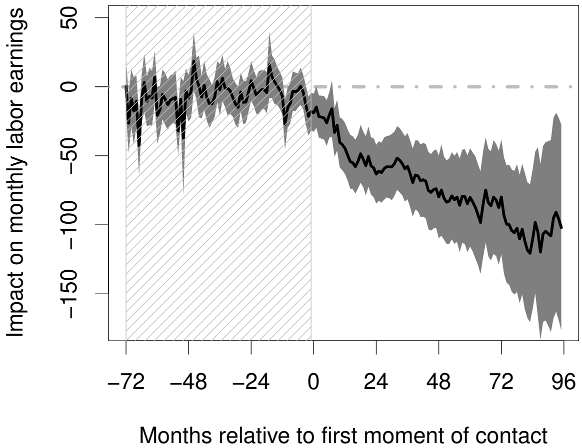
To interpret the magnitude of the causal impact of waiting time on labor market status, the effect sizes can be compared to the estimated effects of the onset of mental health problems from Section 2.4. The onset of mental health problems is associated with a drop in the probability of being employed of approximately nine percentage points, an increase in the probability to receive sickness/disability benefits of seven percentage points and an increase in the probability of receiving social assistance of four percentage points. A two-month (one standard deviation) increase in waiting time decreases the employment rate by four percentage points and increases the receipt of DI benefits by two percentage

[18]Estimation using time until intake instead of total waiting time yields larger estimates, as shown in Appendix Figure 2.A.6. If time until intake increases, all individuals are affected whereas mental healthcare providers are able to allocate increased waiting time to less severe cases if total waiting time increases. This potentially explains the larger effects of time until intake.

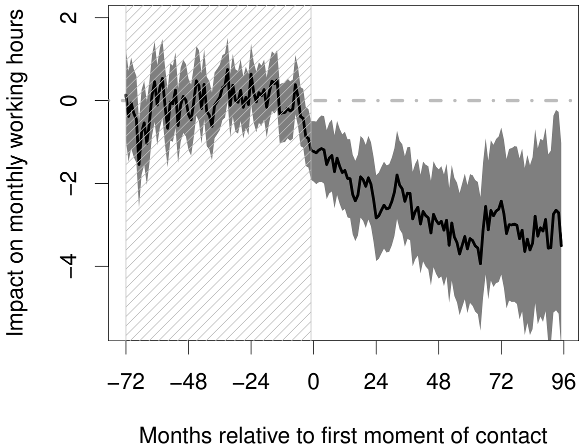
[19]In later years, the sample size decreases, which decreases the precision of the estimates.

Figure 2.10: Estimated impact and 95% CI of one additional month of waiting time on various labor market outcomes

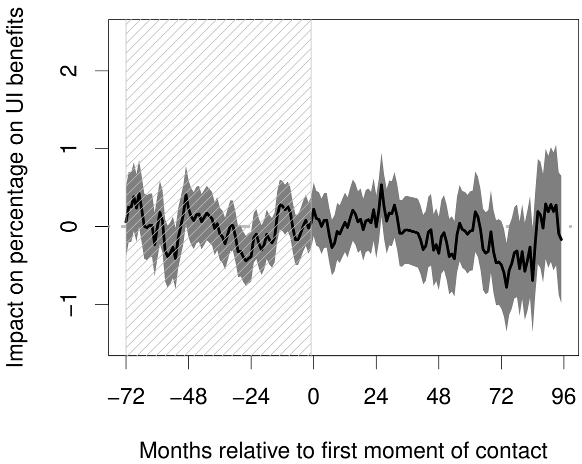
(a) Monthly labor earnings



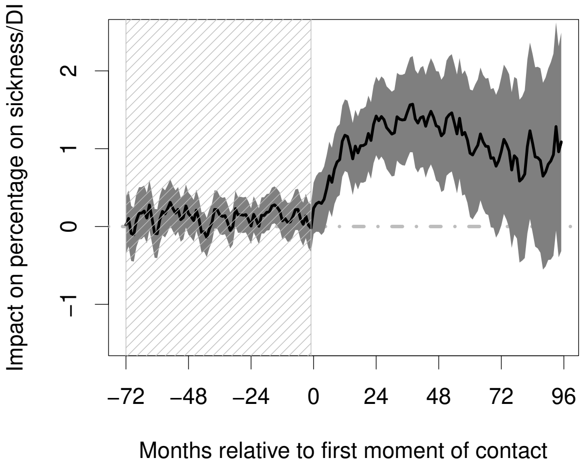
(b) Monthly working hours



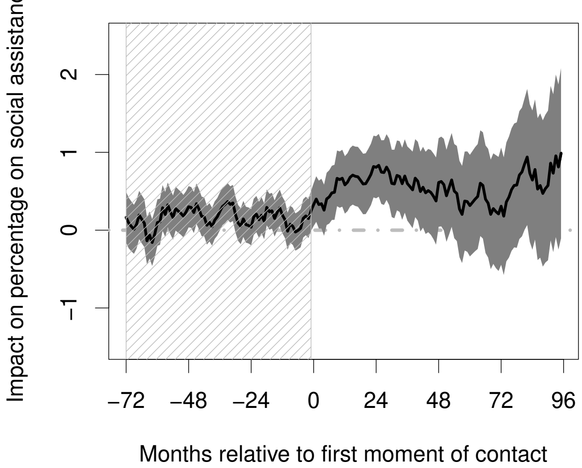
(c) UI benefits



(d) DI benefits



(e) Social assistance



points and the receipt of social assistance by one percentage point. This is almost half of the average effect of the onset of mental health problems. While the receipt of unemployment benefits is also affected by the onset of mental health problems, increased waiting time does not affect the probability to receive these benefits.

There are various potential explanations for the negative impact of waiting time on labor market status, one being a deterioration of (mental) health. Unfortunately, objective measures of (mental) health are not available; that being said, I do observe the number of treatment minutes individuals receive on a monthly basis, and the amount of annual spending on mental healthcare, non-mental healthcare, and pharmaceuticals. One additional month of waiting time increases the cumulative amount of treatment minutes and mental healthcare spending significantly. In the first eight years after the first moment of treatment, the cumulative amount of treatment minutes increases by 150 and spending on mental healthcare increases by €875. Relative to average healthcare utilization of 1,650 minutes and €7,574, this corresponds to an increase of nine and eleven percent in healthcare utilization. Estimates for spending on non-mental healthcare, and pharmaceuticals are insignificant (see Appendix Figure 2.A.8 for all results). Based on these impacts on observed healthcare utilization, increased waiting time for treatment seems to have an effect on (mental) health itself. The negative effects observed on labor market status might be driven by this impact on mental health, or by other mechanisms, such as increased distance to the labor market.

2.5.5 Heterogeneity analysis

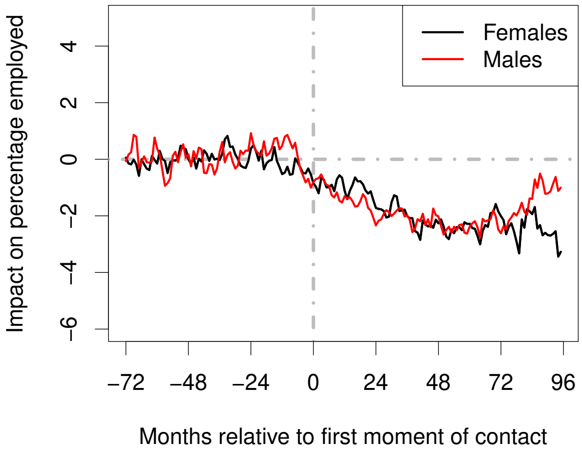
To investigate heterogeneous impacts, I split the sample based on various demographic characteristics and types of mental health problems. This is done both for the first stage and second stage of the IV estimation. Heterogeneity of the first-stage estimates indicates which groups are compliers in the IV setup as it shows which groups have to wait longer when regional waiting time increases. The heterogeneity of the second-stage estimates indicates a differential impact of increased waiting time on labor market status.

Heterogeneity in the first stage is limited (see Appendix Table 2.A.5), indicating that all subgroups are affected to a similar degree by increases in regional waiting time. Figure 2.11 shows heterogeneity of the impact of waiting time on employment by gender, age, migration background, education level, and mental health status.[20] Heterogeneity based on gender and age is limited whereas heterogeneity by migration background and education level is more pronounced. First of all, individuals with a migration background

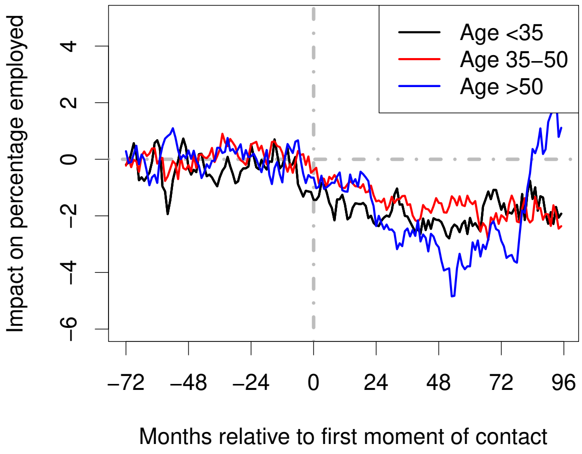
[20]Estimates including confidence intervals are available on request. Given the smaller sample sizes of the subgroups, the estimates are not significantly different from each other.

Figure 2.11: Heterogeneity of the impact of one additional month of waiting time on employment

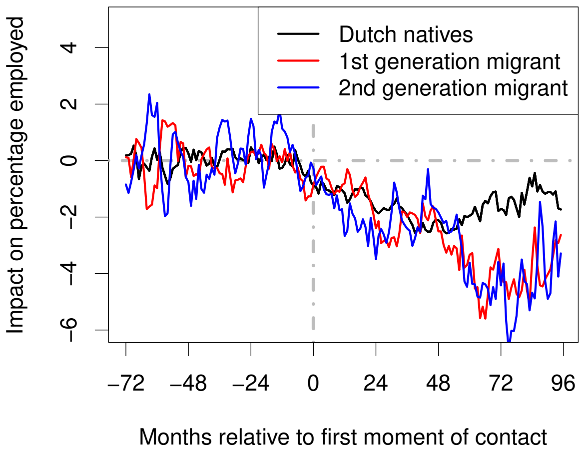
(a) Gender



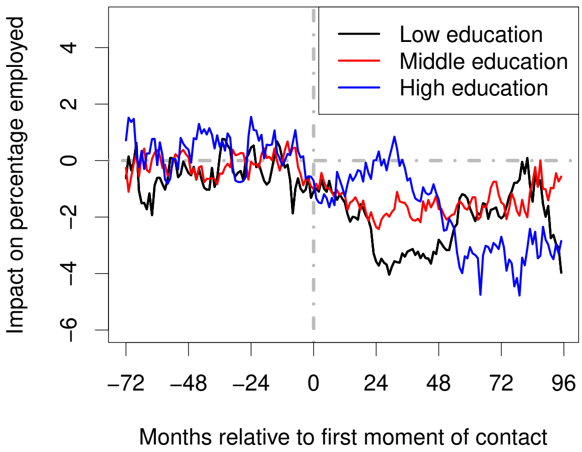
(b) Age



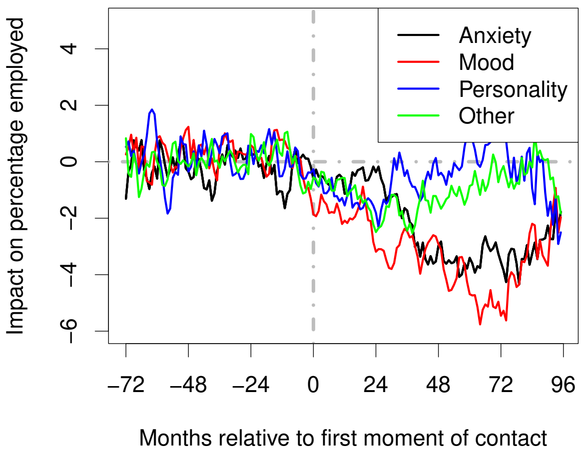
(c) Migration background



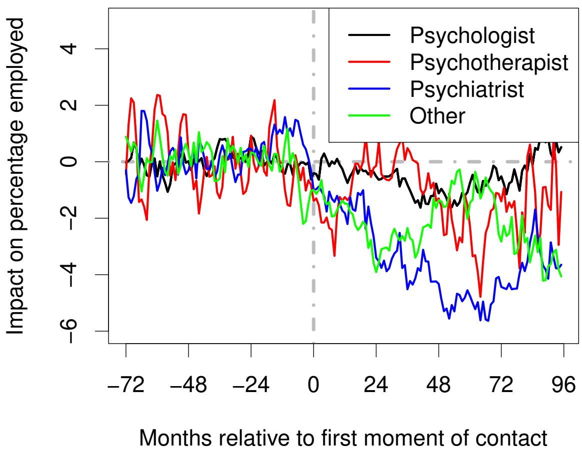
(d) Education level



(e) Main diagnosis



(f) Provider



more in the long run, while the negative effects for Dutch natives fade out over time. Lower-educated individuals on the other hand suffer more in the short term, while higher-educated individuals suffer more in the long term.

Heterogeneity with respect to an individual's main mental health diagnosis and type of provider is shown in panels (e) and (f). The impact of increased waiting time is the largest for anxiety and mood disorders and for treatment by psychiatrists. This is largely in line with the results of the previous section, which showed that individuals being treated for anxiety disorders and by psychiatrists experienced the largest drops in employment.

2.5.6 Regional waiting time as continuous treatment in an event-study

The IV estimation compares individuals in regions with long waiting times to individuals in regions with short waiting times. An alternative estimation approach is a difference-in-differences (DiD) setup in which regional waiting time is a continuous treatment variable. Given that individuals who do not receive treatment are very different from those who do receive treatment, the setup does not use untreated individuals as a control group. Instead, individuals contacting a treatment provider in a region with a short regional waiting time are used as counterfactual for individuals who contact a treatment provider in a region with a long regional waiting time.

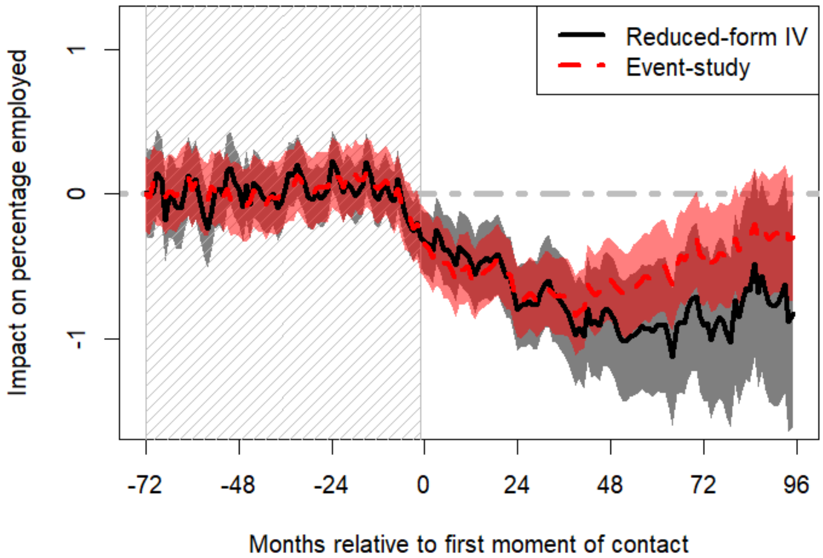
Estimation relies on the so-called strong parallel trends assumption, which implies that the employment trajectories of individuals in a region with a high regional waiting time would otherwise have been the same as the trajectories of individuals in regions with low regional waiting times if they had lived in the latter region(s) (Callaway et al., 2021). Whereas the IV strategy assumes that regional waiting time does not affect employment status prior to the first moment of contact, the DiD strategy does allow for differences prior to the first moment of contact, as long as these differences remain constant over time. As the DiD estimation relies on different assumptions, it can be used as a robustness test. The event-study specification is as follows:

$$E_{it} = \alpha_i + \tau_t + \sum_{l=-71}^{96} \beta_l RW_i + \varepsilon_{it} \tag{2.4}$$

with α_i and τ_t individual and time fixed effects. β_l are the parameters of interest and show the impact over time of regional waiting time. The impact of regional waiting time is normalized to zero at t=72, and the first 71 β_l coefficients are the placebo estimates prior to the first moment of contact.

Given that the event-study estimates correspond to increases in regional waiting time, they should be compared to the reduced form estimates of IV. Figure 2.12 shows these

Figure 2.12: Estimated impact of one additional month of regional waiting time on the probability of employment using reduced form IV (black) and event-study estimation (red)



reduced form IV estimates (a re-scaled version of the estimates in Figure 2.9) and the event-study estimates of the effect of a one-month increase in regional waiting time on the probability of employment. For both estimation strategies, the placebo estimates are similar, thus indicating that the trends of individuals in regions with shorter and longer regional waiting times are parallel and equal. After the first moment of contact, the event-study estimates are similar to the IV estimates, but the event-study estimates appear to fade out over time. This could indicate that time-varying characteristics have an impact on employment. The IV specification controls for some of these time-varying characteristics whereas the event-study does not.

To conclude, increased waiting time for mental health treatment has large effects both on the probability of being employed and the probability of receiving sickness/disability benefits. Heterogeneity of the impact of increased waiting time is significant, with individuals with a migration background and less educated individuals experiencing larger negative effects.

2.6 Unequal access to mental health treatment

Given the negative effects of waiting time and the heterogeneity of these effects, a crucial unanswered question is whether access to mental health is also heterogeneous. In this section, I investigate whether average waiting times differ between groups. The groups considered are the same groups as used in the heterogeneity analysis above; (1) gender, (2) migration background, and (3) education level. In addition, I explore four mechanisms that may drive differences in access to mental healthcare. First, sorting towards regions with short waiting times. Second, differences in pre-treatment labor market status. Third, differences in the propensity to contact a mental healthcare provider, resulting in differences in the severity of mental health problems at the first moment of contact. And fourth, the selection towards specific types of mental healthcare providers. I estimate how individual waiting time depends on individual characteristics using the following regression model:

$$IW_i = \alpha + \beta X_i + \delta Z_i + \varepsilon_i \tag{2.5}$$

with IW_i individual waiting time, X_i indicators for the various groups considered, and Z_i other control variables. In the baseline specification, no control variables (Z_i) are included. I thus estimate the total difference in waiting times between the various groups. To determine which mechanisms cause the observed differences in waiting time, various control variables are sequentially added.

Table 2.4 shows the differences in average waiting times. The average waiting time is 62.3 days. The raw differences in waiting time between the various groups (column (1)) show a small difference based on gender but larger differences based on migration background and education level. First-generation migrants and less educated individuals have to wait more than one week longer on average than their respective counterparts. To control for spatial sorting, column (2) includes regional fixed effects. The resulting estimates are similar, implying that spatial sorting does not explain the differences in waiting times. Differences in pre-treatment employment status do partly explain the observed difference in waiting times based on education level as shown in column (3). When comparing individuals seeking treatment for a similar diagnosis, the difference in waiting time is actually larger (column (4)).[21] Finally, column (5) includes provider fixed effects, hence comparing individuals with similar pre-treatment employment, seeking treatment for a similar mental health diagnosis at the same provider. The small difference based on gender disappears, while differences based on migration background and education

[21]When zooming into individuals with a depression of the same severity, the conclusion remains unchanged as shown in Appendix Table 2.A.6

Table 2.4: Differences in average waiting time by gender, migration background and education level

	Waiting time in days				
	(1)	(2)	(3)	(4)	(5)
Gender[a]:					
Female	1.4**	1.3**	2.0**	1.4**	-0.2
	(0.2)	(0.2)	(0.2)	(0.2)	(0.2)
Migration background[b]:					
1^{st} generation[c]	7.9**	9.3**	7.1**	10.9**	7.1**
	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)
2^{nd} generation[d]	2.4**	3.3**	2.8**	4.4**	3.1**
	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)
Education level[e]:					
Low	12.9**	11.8**	8.0**	10.0**	3.2**
	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)
Middle	6.9**	6.0**	4.1**	5.0**	0.7**
	(0.2)	(0.2)	(0.2)	(0.2)	(0.2)
Demographic controls	X	X	X	X	X
Regional fixed effects		X	X	X	
Pre-treatment employment			X	X	X
Mental health diagnosis				X	X
Provider fixed effects					X
Mean waiting time	62.3	62.3	62.3	62.3	62.3
Sample size	524,707	524,707	524,707	524,707	524,707

Standard errors shown in parenthesis; *significant at a 5% significance level; **significant at a 1% significance level; (a) Baseline gender is male; (b) Baseline migration background is native, (c) Individuals who migrated to the Netherlands, (d) Children of first-generation migrants (e) Baseline education level is high

level decrease as well. However, first-generation migrants and less educated individuals on average still have to wait seven and three days longer respectively.

The differences in waiting times based on migration background and education level can have various explanations. These individuals might be less aware of their options for finding providers with shorter waiting times, or resource constraints might force them to choose the closest provider geographically. Additionally or alternatively, these individuals might be less capable of explaining their problems, for example, due to language barriers. The heterogeneity analysis of Section 2.5 indicated relatively large effects of waiting times for individuals with a migration background and individuals with a lower education level. Hence, the cost of increased waiting time appears to be larger for these groups, and they tend to have longer waiting times. Given that education level and migration background are strongly linked to socioeconomic status, these differences in waiting times might therefore further increase inequality in society.

2.7 Conclusion

The increasing demand for mental health problems and limited treatment capacity for mental health treatments has resulted in long waiting times in many OECD countries. Delays in accessing some other forms of care have been shown to negatively impact employment (Godøy et al., 2022; Williams & Bretteville-Jensen, 2022), yet little is known about the microeconomic impact of waiting times for mental healthcare specifically. Using administrative data for the Netherlands on mental health treatments and labor market outcomes, I estimated the causal impact of waiting times for treatment on employment.

As a benchmark, I first showed that the onset of mental health problems and the subsequent start of treatment is associated with a nine percentage point drop in an individual's employment probability, along with an increased inflow into both sickness/disability insurance and social assistance. Next, I conducted causal analyses on the effects of waiting times using regional waiting times as instruments. I showed that an increase in waiting time of two months (one standard deviation) decreases the probability of employment by four percentage points and increases the probability of receiving sickness/disability benefits by two percentage point for at least eight years. Differential impacts of waiting time on employment are substantial, with less educated individuals and those with a migration background experiencing the largest negative effects. These two groups also have to wait longer on average before receiving treatment. The burden of increased waiting time is therefore especially large for vulnerable groups, potentially increasing inequality in health and labor outcomes.

The obtained estimates of the effects of waiting time for mental health treatment have both similarities and differences with the estimates of waiting time for orthopedic surgeries as examined by Godøy et al. (2022) and for substance abuse treatment as examined by Williams & Bretteville-Jensen (2022). Comparable to the estimates in this chapter, a one-month increase in waiting time increases the probability of receiving disability benefits by one percentage point for orthopedic surgery and it decreases the probability to be employed by three percentage points for substance abuse treatment. Godøy et al. (2022) also find disproportionately large effects on less educated individuals. However, there are marked differences in the extensive-margin effects of waiting time. While Godøy et al. (2022) find limited extensive-margin effects, both the present study and that of Williams & Bretteville-Jensen (2022) actually find employment to mostly be affected on the extensive margin. Another difference lies in the impact of delays on healthcare utilization. Waiting time for orthopedic surgery has a limited impact on the utilization of care while waiting time for mental health treatment and for substance abuse treatment increases the utilization of care substantially. Nevertheless, waiting times for all three types of care have large and negative effects, which underscores the need to ensure timely access to a variety of healthcare services.

In the context of mental health treatment, waiting times can be reduced by either reducing the demand for treatment through prevention or by increasing the supply of treatment. The Dutch government has expressed a strong interest in prevention, as it is seen as a cost-effective intervention as compared to interventions aimed at increasing treatment capacity (Rijksoverheid, 2022). Conversely, however, a back-of-the-envelope calculation does indicate that the costs of increased provision would be dominated by savings on other government expenditures.[22] Reducing the waiting time in the Netherlands by one month for one year would yield a reduction in employment loss of approximately 2,000 individuals, with associated savings of almost €380 million. To achieve such a reduction in waiting time, an additional 50 psychiatrists/psychologists would be needed for one year, with a labor cost of approximately €5 million. These calculations are clearly an oversimplification, but they nevertheless show that the economic gains from reduced waiting times can be substantial. These findings have some notable policy implications. For instance, in the Netherlands, there are many students studying psychology, but only a relatively small share of them work as psychologists after graduation. Given the degree to which the social benefits of increased supply of mental healthcare could exceed the costs, it stands to reason that better compensation for mental health professionals could be in everyone's best interest if higher pay attracts more workers to the field.

Alongside such efforts to reduce waiting times, the differential effects of delayed treatment on vulnerable groups suggest that gains can be made by focusing on individuals with a migration background and those with lower educational attainment. Not only do these groups wait longer for mental healthcare on average, but the effect of waiting times on their employment status is also greater. Increased access to mental health treatment in neighborhoods with low socioeconomic status, along with targeted mental health interventions for (unemployed) vulnerable groups, might alleviate some of the additional burden these individuals face.

[22]See Appendix Section 2.A for a back-of-the-envelope calculation of cost savings.

2.A Appendix

2.A.1 Additional Data

Table 2.A.1: Construction of mental healthcare expenditures and physical healthcare expenditures based on expenditure categories used by Statistics Netherlands

Mental healthcare	Non-mental healthcare	Pharmaceuticals
First-line psychological healthcare	General practitioner	Pharmacy
Mental healthcare	Hospital healthcare	
Basic-mental healthcare	Paramedical healthcare	
Specialist mental healthcare	Nursing without stay	
Geriatric rehabilitation healthcare		

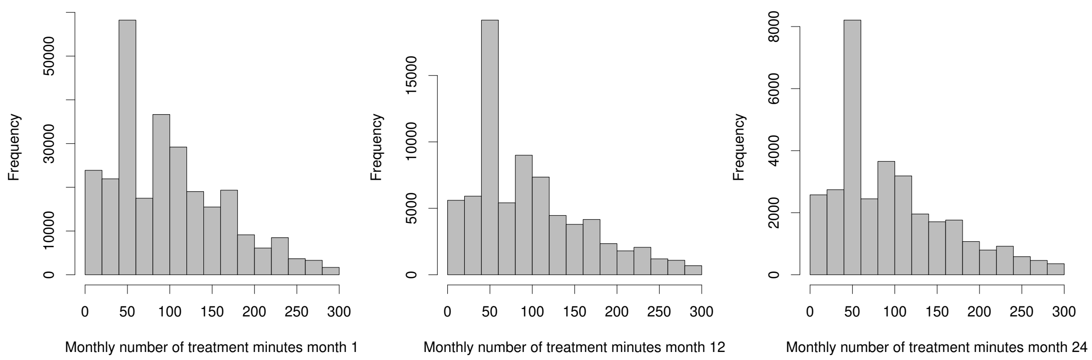
Note: Several expenditure categories, such as healthcare abroad and other costs, as used by Statistics Netherlands are excluded from all categories as it is not clear to which category they belong.

Figure 2.A.1: The distribution of treatment minutes in the 1^{st}, 13^{th} and 25^{th} month of treatment

(a) Month 1

(b) Month 13

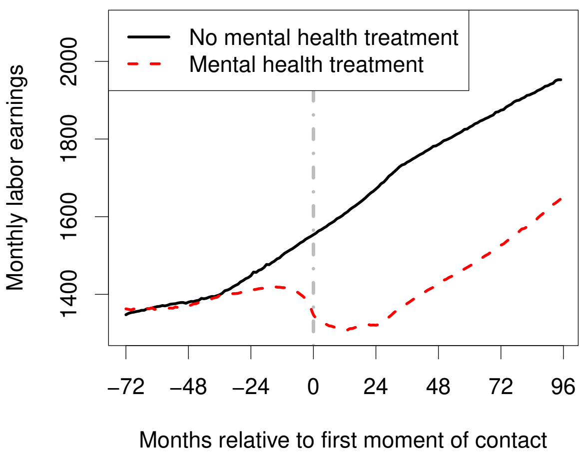
(c) Month 25



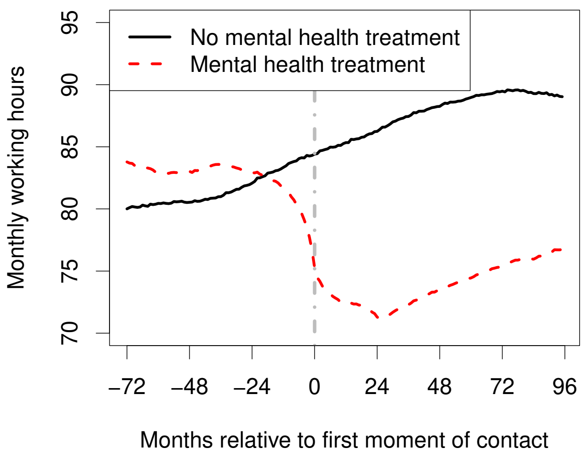
2.A.2 Additional results: onset of mental health problems

Figure 2.A.2: Trends in labor market outcomes for individuals with and without mental health treatment relative to the first moment of contact

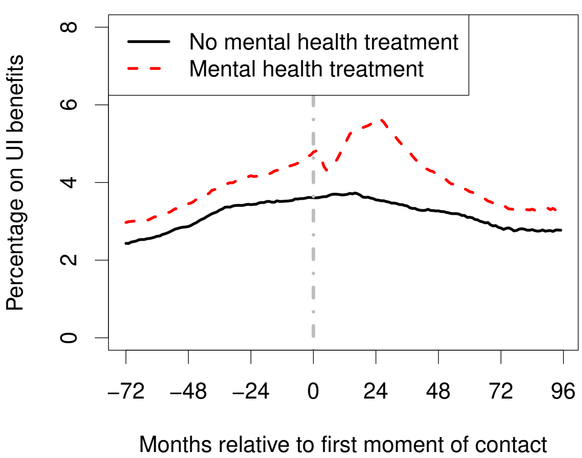
(a) Monthly labor earnings



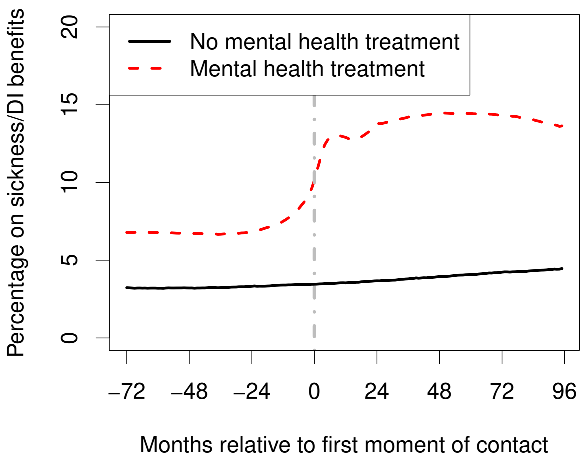
(b) Monthly working hours



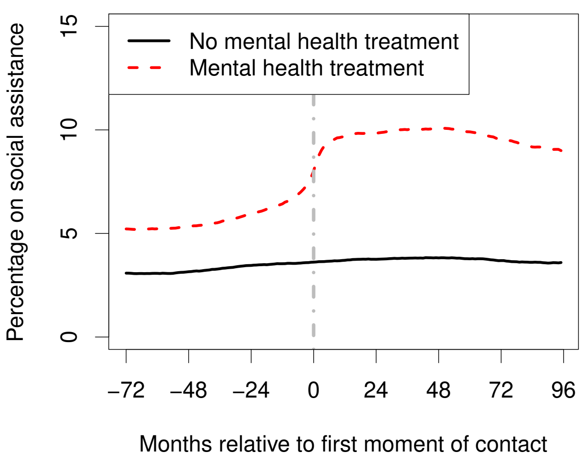
(c) UI benefits



(d) DI benefits



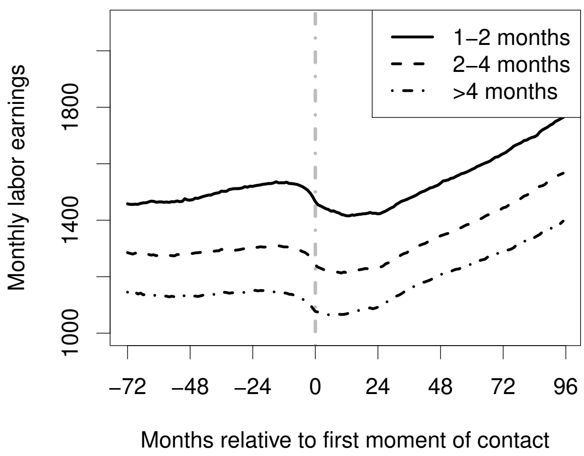
(e) Social assistance



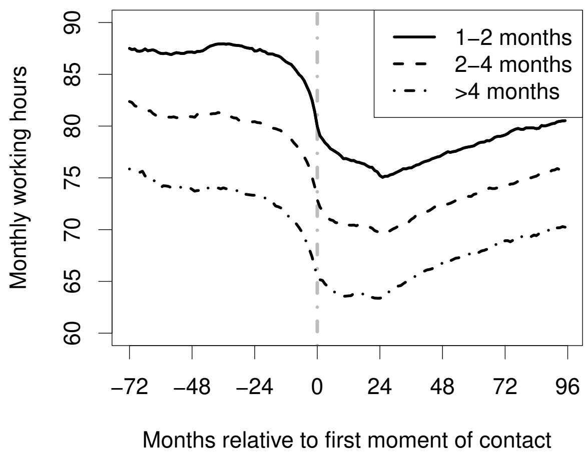
2.A.3 Additional results: impact of waiting time

Figure 2.A.3: Trends in labor market outcomes relative to the first moment of contact for groups based on their individual waiting time

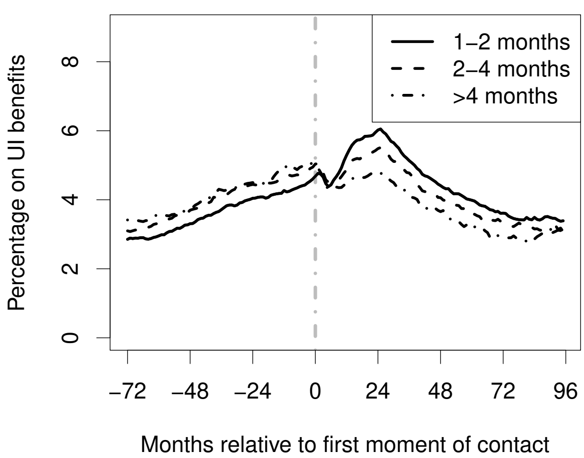
(a) Monthly labor earnings



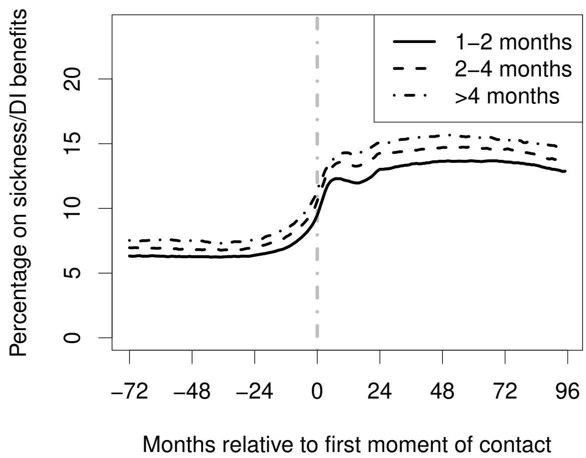
(b) Monthly working hours



(c) UI benefits



(d) DI benefits



(e) Social assistance

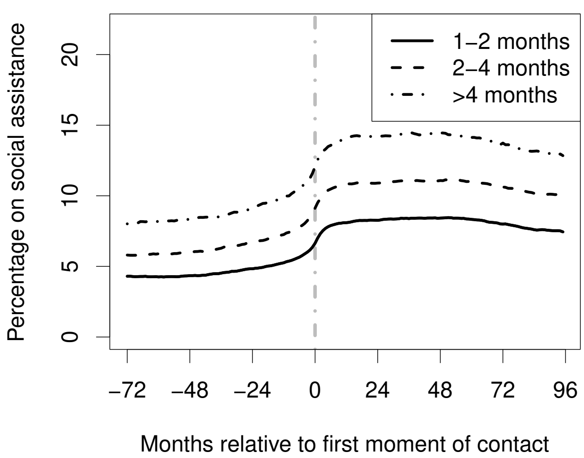


Table 2.A.2: Full set of control variables X_i used in the IV specification

Control variable	Values
Demographics:	
Gender	Male or Female
Age	Dummies for 5-year age groups
Migration status	Native, first or second generation migrant
Education level	Low, middle, high or unknown
Calendar-month of first contact	Year-month fixed effects (60 dummies)
Mental health status:	
Mental health diagnosis	DSM-IV classification
Crisis	Indicator for crisis admission
Zero days until intake	Indicator for days until intake equal to zero
Pre-treatment employment[a]**:**	
Employment	Indicator for being employed
UI	Indicator for receiving UI
Social assistance	Indicator for receiving social assistance
Sickness/DI	Indicator for receiving sickness/DI benefits
Municipality characteristics[b]**:**	
Percentage Caucasian	Dummy for decile
Percentage of Moroccan migrants	Dummy for decile
Percentage of Turkish migrants	Dummy for decile
Average house valuation	Dummy for decile
Percentage of owner-occupied houses	Dummy for decile
Percentage of housing-corporation-occupied houses	Dummy for decile
Percentage of houses build prior to the year 2000	Dummy for decile
Average income per inhabitant	Dummy for decile
Share below 40th percentile in income distribution	Dummy for decile
Share above 20th percentile in income distribution	Dummy for decile
Share with income below social minimum	Dummy for decile
Percentage of individuals receiving UI benefits	Dummy for decile
Percentage of individuals receiving DI benefits	Dummy for decile
Population density	Dummy for decile

(a) Pre-treatment employment outcomes are measured 24 months prior to treatment, or 24 months prior to outcome in case of placebo outcomes; (b) Value of municipality characteristics is the decile in which the municipality falls in the distribution over all municipalities

Figure 2.A.4: Non-parametric estimates of the association between waiting time (grouped by week) and probability to be employed 12 months after the first moment of contact

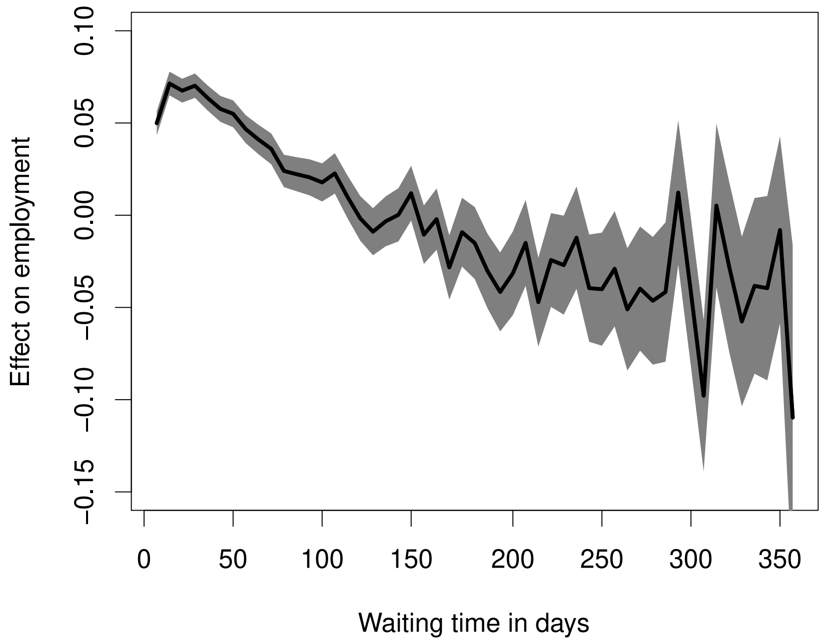


Table 2.A.3: Impact of (lagged) regional unemployment rate on the regional waiting time

	Regional waiting time			
Unemployment rate t	0.26	0.85	-1.14	-1.07
	(0.67)	(2.17)	(2.30)	(2.29)
Unemployment rate t-1		-0.74	5.41	5.50
		(2.21)	(3.65)	(3.71)
Unemployment rate t-2			-3.59	-5.37
			(2.31)	(3.67)
Unemployment rate t-3				1.67
				(2.29)
	Regional waiting time			
Employment rate t	0.18	-1.17	-0.54	-0.39
	(0.12)	(0.55)	(0.55)	(0.55)
Employment rate t-1		0.73	-0.77	-0.75
		(0.55)	(0.79)	(0.79)
Employment rate t-2			0.81	0.48
			(0.55)	(0.79)
Employment rate t-3				0.70
				(0.55)

Standard errors shown in parentheses; *significant at a 10% significance level; **significant at a 5% significance level

Table 2.A.4: Impact of regional waiting time on probability to reach intake/start treatment and on the composition of patients contacting mental healthcare providers

Impact on probability to:	
Reach intake	0.07%
	(0.13)
Start treatment	-0.06%
	(0.15)
Impact on composition of patients	
Demographics:	
Female	0.18%
	(0.20)
Age	-0.12*
	(0.05)
Dutch Native	-0.41%*
	(0.17)
Low education level	0.35%
	(0.18)
Middle education level	-0.34%
	(0.22)
High education level	-0.01%
	(0.21)
Mental health diagnosis:	
Mood disorder	-0.65%**
	(0.18)
Personality disorder	0.61%**
	(0.11)
Anxiety disorder	-0.33%*
	(0.17)
Other disorder	-0.38%*
	(0.19)
Mental health provider:	
Psychologist	-0.49%*
	(0.20)
Psychotherapist	-0.40%**
	(0.13)
Psychiatrist	0.42%**
	(0.16)

Standard errors shown in parentheses; *significant at a 5% significance level; **significant at a 1% significance level

Table 2.A.5: Heterogeneity of first-stage estimates of the impact of regional waiting time on individual waiting time

Gender:	Male	Female		
	0.42**	0.40**		
	(0.01)	(0.01)		
Age:	<35	35-50	>50	
	0.47**	0.41**	0.32**	
	(0.01)	(0.01)	(0.01)	
Migration background:	Native	1^{st} generation	2^{nd} generation	
	0.39**	0.48**	0.50**	
	(0.01)	(0.02)	(0.02)	
Education level:	Low	Middle	High	
	0.48**	0.46**	0.36**	
	(0.02)	(0.01)	(0.01)	
Diagnosis:	Anxiety	Mood	Personality	Other
	0.44**	0.33**	0.62**	0.36**
	(0.01)	(0.01)	(0.03)	(0.01)
Provider:	Psychologist	Psychotherapist	Psychiatrist	Other
	0.47**	0.32**	0.32**	0.39**
	(0.01)	(0.02)	(0.02)	(0.01)

Standard errors shown in parentheses; *significant at a 10% significance level; **significant at a 5% significance level

Figure 2.A.5: Estimated impact and 95% CI of one additional month of waiting time on probability to be employed using IV with regional controls (black) and OLS (red)

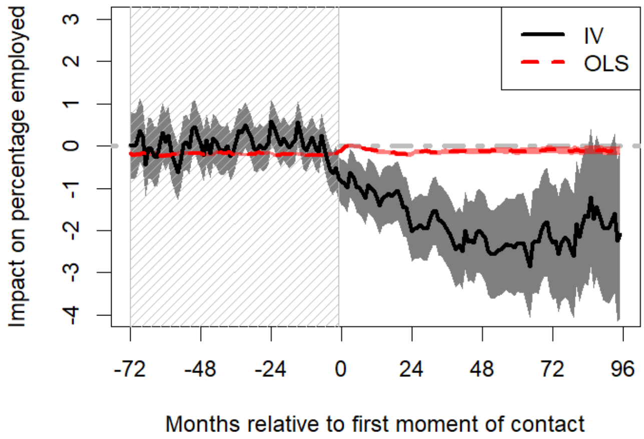


Figure 2.A.6: Estimated impact and 95% CI of one additional month of total waiting time (black) and time until intake (red) on probability to be employed

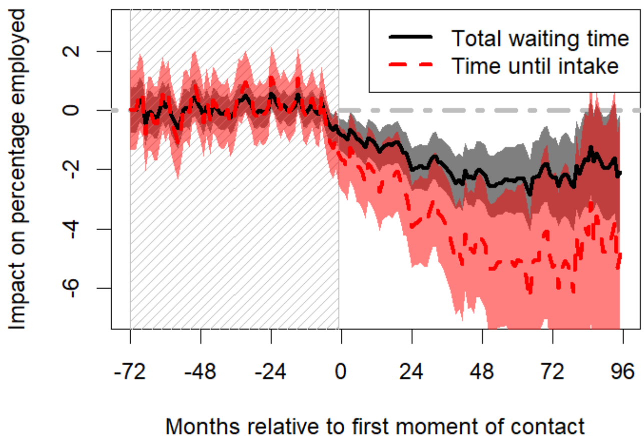


Figure 2.A.7: Estimated impact and 95% CI of one additional month of waiting time on probability to be employed using regional controls (black) and regional fixed effects (red)

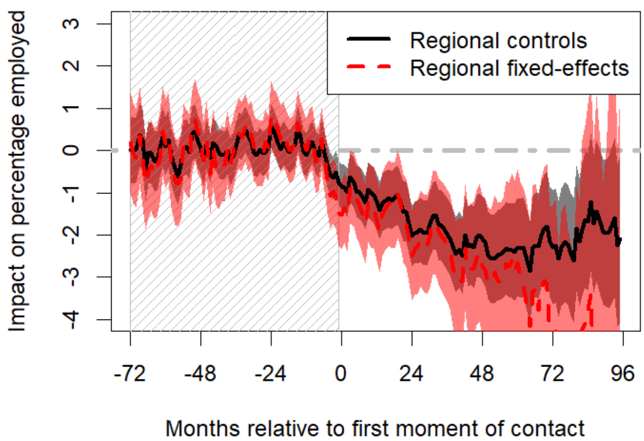
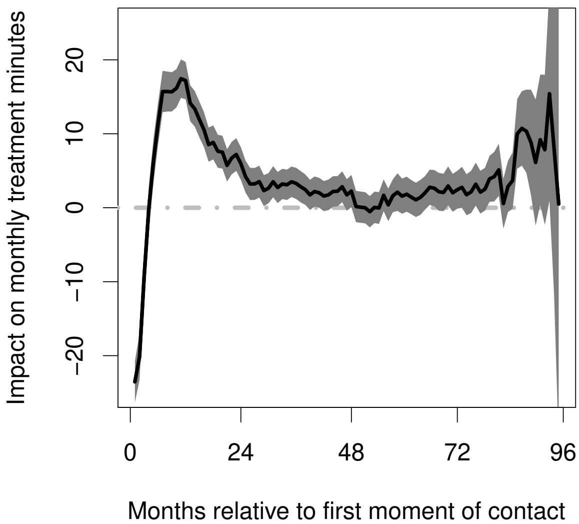
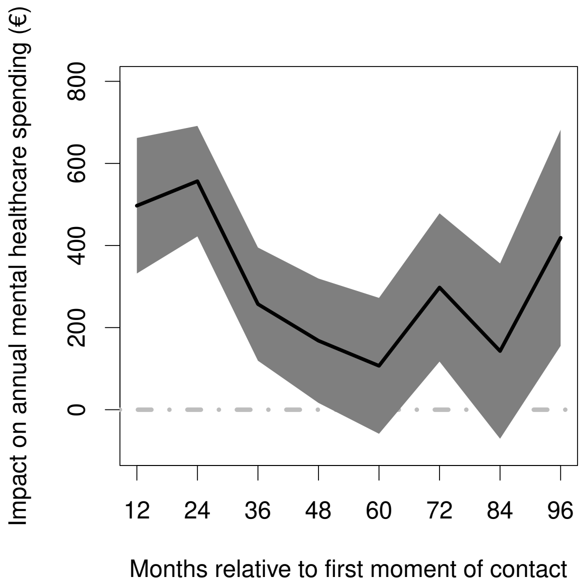


Figure 2.A.8: Estimated impact and 95% CI of one additional month of waiting time on healthcare utilization

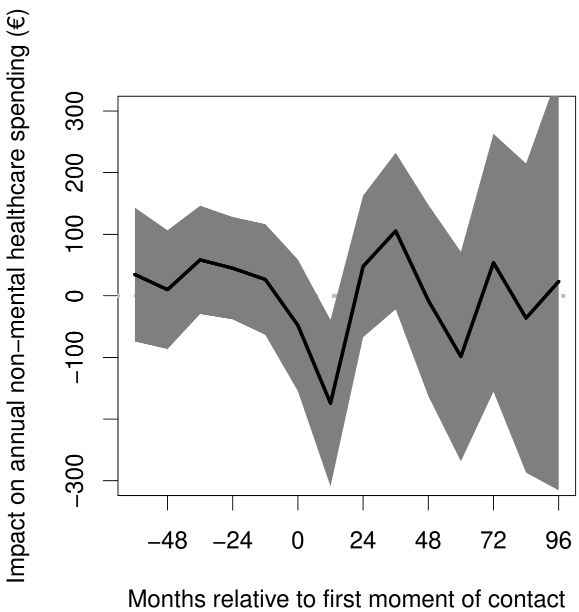
(a) Monthly number of treatment minutes



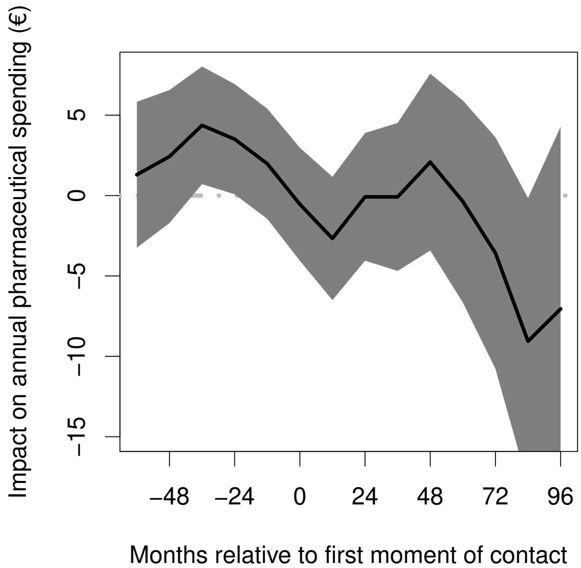
(b) Annual spending on mental healthcare



(c) Annual spending on non-mental healthcare



(d) Annual spending on pharmaceuticals



2.A.4 Additional results: differences in average waiting time

Table 2.A.6: Differences in waiting time for individuals with depression of given severity

	Low severity	Mild severity	High severity
Gender[a]:			
Female	-0.2	0.0	0.3
	(1.5)	(0.8)	(1.4)
Migration background[b]:			
1^{st} generation[c]	19.3**	13.2**	7.8**
	(2.1)	(1.1)	(1.7)
2^{nd} generation[d]	7.8**	5.0**	9.9**
	(2.3)	(1.2)	(2.0)
Education level[e]:			
Low	6.9**	6.7**	5.4**
	(2.2)	(1.2)	(1.9)
Middle	1.9	1.0	-2.1
	(1.7)	(1.0)	(1.7)
Demographic controls	X	X	X
Regional fixed effects			
Pre-treatment employment	X	X	X
Mental health diagnosis	X	X	X
Provider fixed effects	X	X	X
Mean waiting time	53.4	53.7	44.3
Sample size	5,980	19,094	6,746

Standard errors shown in parenthesis; *significant at a 5% level; **significant at a 1% significance level; (a) Baseline gender is male; (b) Baseline migration background is native, (c) Individuals who migrated to the Netherlands, (d) Children of first-generation migrants (e) Baseline education level is high

2.A.5 Back-of-the-envelope cost savings calculation

In the Netherlands, a one-month reduction in waiting time for one year would affect at least 100.000 individuals starting treatment per year. According to the IV estimates on employment, this would lead to a reduction in employment loss of approximately 2.000 individuals for at least eight years. The average cost to society of someone without employment has been estimated to be approximately €24.000,- according to the audit office of the Dutch government. A reduction in employment loss of 2.000 individuals thus translates into a cost savings of almost €380 million.

Chapter **3**

Why are workers with fixed-term contracts more likely to apply for disability insurance than permanent workers?

3.1 Introduction

In most OECD countries, workers with fixed-term contracts suffer from worse health conditions than workers with permanent contracts (OECD, 2010).[2] This negative association is found both in studies based on cross-sectional data and in panel surveys – see Kim et al. (2012), Virtanen et al. (2005) and Benach et al. (2014) for survey studies – and pertains to both mental and physical health problems (Bardasi & Francesconi, 2004). From a public finance perspective, however, evidence on the differences in the enrollment of temporary and permanent workers into disability insurance (DI) is scarce.[3] When workers with fixed-term contracts have a higher risk of applying for DI, a crucial question is whether their contract type causally generates this higher probability of applying for DI or whether there is segmentation in the labor market such that relatively unhealthy workers with weaker labor market prospects end up in fixed-term contracts. The latter has been suggested for the US, where vulnerable workers with low productivity levels

[1]This chapter is based on Koning et al. (2022b). The online appendix of this chapter can be found at www.rogerprudon.com/research

[2]Note that 'flexible work' may refer to various employment constructions in the literature. In this chapter, we focus on fixed-term contracts, which can be considered the most widespread type of flexible employment.

[3]The literature on the impact of nonstandard work arrangements tends to focus on the wage effects of specific types of nonstandard work arrangements, such as contracts with temp work agencies (Drenik et al., 2020; Katz & Krueger, 2019; Goldsmith & Schmieder, 2017).

make up an increasing fraction of Social Security Disability Insurance (SSDI) benefit recipients (Maestas, 2019; Autor & Duggan, 2003; Deshpande & Lockwood, 2022). These workers have either worse health conditions (a 'health disability') or limited prospects in the labor market, which prevent them from working (a 'work disability') (Benítez-Silva et al., 2010).

In this chapter, we study the mechanisms driving the differences in DI applications and awards between workers with fixed-term and permanent contracts.[4] We use large-scale administrative data from the Netherlands on labor market histories, disability application records, and health consumption and health treatments. In the time period under investigation (2010-2015), the prevalence of fixed-term contracts increased, and the probability of both applying for and being awarded DI for this group was higher than for workers with permanent contracts. While the Netherlands has an exceptionally high rate of temporary employment, such contracts are similarly prevalent in other European countries such as Spain, France and Italy.[5]

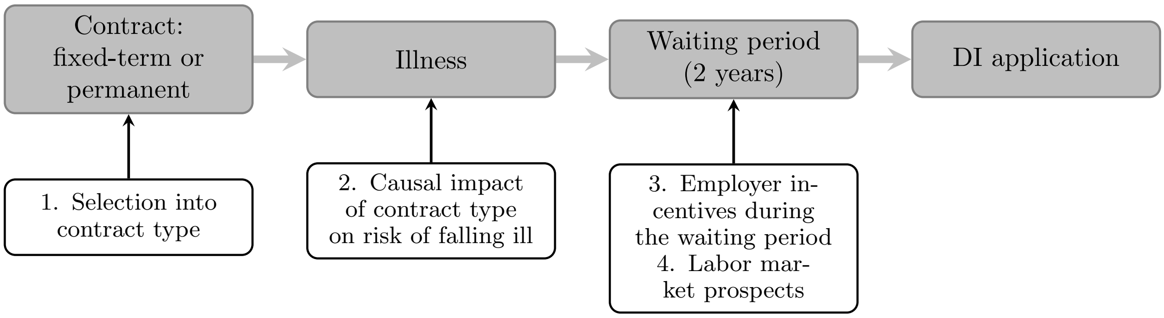
We hypothesize that there are four potential explanations for this higher probability of applying for DI that are relevant at different, consecutive stages before the DI application, as illustrated by Figure 3.1. These include (i) ex-ante compositional differences in the characteristics of workers with fixed-term contracts and of those with permanent contracts, (ii) the causal impact of the contract type on the probability of falling ill, (iii) the impact of differences in the role of the employer during illness (but prior to DI application), and (iv) the role of outside options in the labor market in the decision to apply for DI. By unifying the analysis of each of these explanations, we provide a comprehensive picture of the drivers of the higher probability of DI application among workers with fixed-term contracts. Such an analysis is essential for guiding policy: if contract type has a causal impact, policies to reduce DI applications could focus on discouraging the use of temporary work arrangements. However, if not, structural changes to employer incentives or improvements in the labor market prospects of vulnerable workers – e.g., through retraining – would be required to reduce DI applications.

In brief, we find that the combined effect of all four mechanisms amounts to an increase in the probability of applying for disability benefits of 70-80%. This increase corresponds to a scenario in which workers with fixed-term contracts have the same composition and a similar probability of falling ill as workers with permanent contracts but their employers have fewer incentives during their illness and they have worse outside options in the labor market. The observed difference in the probability of DI application over the entire period

[4]In the main analysis, we focus on the probability of applying for DI. The results using the probability of being awarded DI benefits are reported as a robustness check and yield similar conclusions.

[5]In all OECD countries in 2015, 12% of workers were employed through fixed-term contracts, while the EU average was actually 16%. In the Netherlands, it was 20% (OECD, 2015).

Figure 3.1: Timeline from employment to DI application: Potential mechanisms explaining the gap in DI application probabilities between fixed-term and permanent workers



of analysis is 30%, which is essentially a weighted average over all possible scenarios in which the various mechanisms are turned either on or off. Decomposing the difference in DI application probabilities, we find that the selection of workers into contract types cannot explain the higher probability of fixed-term contract workers applying for DI. Additionally, the risk of falling ill is not substantially greater for workers with fixed-term contracts (after controlling for observables). Rather, employer efforts during illness and the difference in outside labor market options explain most of the difference in the DI application probabilities. In other words, it is not the fact that fixed-term contracts are associated with more frequent illness but that *conditional on being ill*, workers with fixed-term contracts face different support structures and incentives that make them more likely to ultimately apply for DI.

Following the timeline in Figure 3.1, we first regress a dummy for applying for DI on contract type and sequentially add a wide range of controls for demographics, employment, occupation and prior health. Our initial focus is on selection into fixed-term contracts, which stems from specific worker types and specific jobs having a higher or lower a priori disability risk. After controlling for these factors, we find that the monthly probability of applying for DI for workers with a fixed-term contract is approximately 50% higher than the raw difference of 30%. Even with the inclusion of additional sets of control variables (such as the prior health measures), this gap remains almost constant, suggesting that workers with worse health are not more likely to select into fixed-term contracts. These results are confirmed by non-parametrically weighting the DI application probabilities. While we cannot exclude the possibility that permanent and fixed-term workers differ on dimensions still unobservable to us, the robustness of the application risk premium estimate suggests that selection is unlikely to explain the difference.

Second, we consider differences in the likelihood of falling ill between fixed-term and permanent workers. We interpret these differences as the direct effects of contract type on health deterioration due to a higher occupational hazard or increased stress due to the lack of job security.[6] To proxy for the occurrence of falling ill, we define health shocks as substantial increases in medical consumption for physical health (both hospitalizations and medication) and/or the start of mental health treatment. We then find that the risk of mental health shocks is approximately 10% higher for fixed-term workers than for permanent workers. the gap in the probabilities of applying for DI conditional on being ill decreases to approximately 40%. The risk of a physical health shock is approximately 2% lower for fixed-term workers, and the gap in the DI application probabilities conditional on such a shock remains approximately 50%.

We next focus on what happens *after* a worker falls ill. At this stage, the employer has weaker incentives to facilitate the rehabilitation of workers whose contracts will expire anyway (i.e., explanation 3). This divergence is strengthened by the financial incentives inherent in the targeting of continued wage payments and experience rating toward permanent workers only. In light of these considerations, we exploit a policy reform in 2013 that increased the monitoring obligations and financial consequences of the employer if their fixed-term workers entered DI. Using a difference-in-differences (DiD) strategy on the subsample of workers who have experienced a health shock, we find that approximately half of the conditional difference in the DI application probabilities is explained by differences in employer incentives. As such, we also add to earlier findings in the literature that experience ratings affect fatality and injury rates (Koning, 2009; Kyyrä & Tuomala, 2013; Tompa et al., 2012; De Groot & Koning, 2016).

As a final step, we shift our focus to differences in *incentives* for ill fixed-term and permanent workers during their waiting period. We hypothesize that employees with fixed-term contracts are more vulnerable to bad labor market prospects, as their contracts are likely to end during their waiting period. When their outside options are limited, applying for DI becomes a more attractive option. We assess the relevance of this channel by comparing fixed-term workers in occupations with tight and loose labor markets (Autor & Duggan, 2003; Benítez-Silva et al., 2010). Conditional on illness, we then find that the probability of applying for DI among permanent workers does not vary systematically with market tightness. This contrasts with workers with fixed-term contracts, for whom the probability of applying for DI is substantially lower in tight sectors. Specifically, in sectors with tight labor markets, the difference in the DI application probabilities – conditional on being ill – shrinks to only 8%. Combining our four results, we are able to explain more than 80% of the difference in the DI application probability for employees with fixed-term contracts.

[6] Note that these effects may be dampened if fixed-term contracts incorporate probationary periods and workers therefore have incentives to not be absent (Ichino & Riphahn, 2005; Riphahn & Thalmaier, 2001; Engellandt & Riphahn, 2005).

The remainder of this chapter is organized as follows. In Section 3.2, we describe the relevant institutions in the Netherlands and provide descriptive evidence regarding the DI application probabilities and healthcare expenditures. Section 3.3 illustrates the decomposition framework that is used as a guide for the empirical analyses. Section 3.4 lays out our empirical strategy and presents the results for the various mechanisms. Finally, Section 3.5 compares our results with those of other studies, and Section 3.6 concludes.

3.2 Institutional setting and descriptive statistics

In this section, we provide a brief overview of the institutional setting in the Netherlands regarding DI benefits and the distinction between fixed-term and permanent contracts. We provide definitions of the key terminologies in our analysis and present relevant descriptive statistics.

3.2.1 The DI scheme in the Netherlands

Implementation of the DI system in the Netherlands is carried out by the Employee Insurance Agency (UWV).[7] DI benefit applications can be filed after two years of illness; this is referred to as the 'waiting period' (see also Figure 3.1). DI applications consist of a medical assessment by a medical expert and an assessment of remaining earnings capacity by a labor market expert. If the loss in earnings capacity is less than 35% of pre-application wages, the application is rejected. Benefits amount to 70% of the loss in earnings (relative to pre-disability earnings), although further financial incentives exist to encourage making use of one's remaining earnings capacity. For further details, see Koning et al. (2022a). Applicants with an earnings capacity loss between 35% and 80% are awarded partial benefits, and applicants with an earnings capacity loss above 80% are awarded full benefits.

During the waiting period, permanent and fixed-term workers face different levels of employer support and commitment. For *permanent workers*, employers are obliged to continue paying wages during the worker's illness over the entire waiting period. Concurrent with this, they are obliged to actively monitor the health of the employee and provide support to facilitate rehabilitation, for example, by adjusting working conditions. Moreover, any DI benefits awarded are experience rated for permanent contract workers. For workers with *fixed-term contracts*, the employer continues paying their wages during their illness until the date when the contract expires. Next, these workers receive illness

[7]Note that prior to 2006, a range of reforms was implemented to reduce the DI inflow after the number of DI recipients had grown substantially in the 1980s and 1990s. For further details and evaluations of these reforms, please see Koning & Lindeboom (2015), Van Sonsbeek & Gradus (2012), Godard et al. (2019) and Hullegie & Koning (2018).

benefits through social insurance. Until 2013, the employer's (financial) responsibility for the fixed-term worker ended at that point. In cases in which the ill worker entered the DI system, the extra DI benefit costs were not experience rated.

As the share of DI applications from fixed-term contract workers steadily increased after 2006, the government introduced a reform in 2013 that extended the monitoring and financial obligations of the employer to fixed-term workers on sick leave. Since the implementation of that reform, employers have remained financially responsible for their employees on sick leave after their contracts have expired. This means that the employer premiums for sick pay and the DI benefits for fixed-term workers are experience-rated. To alleviate the financial risks that might arise for small employers, the premium is averaged within sectors for employers with fewer than 10 employees.[8] Together with this change, a one-year medical assessment was introduced for individuals on sick leave whose contracts had expired. A similar assessment was already in place for employees with a permanent contract who were on sick leave. Overall, the 2013 reform substantially reduced the difference in employer obligations and incentives for fixed-term and permanent workers.

It should be noted that fixed-term contracts are limited in duration in the Netherlands. An employer is allowed to hire a worker for at most three consecutive fixed-term contracts, with a joint maximum duration of three years. The fourth contract must be permanent.[9] While permanent contracts offer substantial job protection, neither permanent nor fixed-term contracts can be dissolved during illness. Fixed-term workers therefore remain employed during their illness for as long as their contract lasts.

3.2.2 Data sources

Our analysis is based on three administrative datasets that are merged at the individual level. The combination of these three datasets allows us to construct employment and health status trajectories for all employed Dutch individuals. First, we use tax records provided by Statistics Netherlands. These records contain detailed descriptions of all employment contracts in the Netherlands between 2010 and 2015. They include the commencement and end dates of the contracts, individual and firm identifiers, the type of contract (fixed-term or permanent), the industry code, weekly hours and the salary paid.

Second, we use healthcare expenditure data. These data capture total annual individual healthcare expenditures as covered by the basic health insurance system, as well as a breakdown by healthcare type (17 categories). Basic health insurance is mandatory for

[8]For employers with more than 10 and fewer than 100 workers, the DI premium is a weighted average of the individual and the sector-averaged premium. For this group, the weight of the individual premium linearly increases from 0% to 100% with firm size.

[9]The count is reset after a break of at least three months, meaning that lengthy fixed-term employment spells with the same employer are possible only with short breaks.

all Dutch adults; consequently, the data cover the entire Dutch population. We extend these data with even more detailed data from 2011–2016 on mental health treatment trajectories for which the exact start and end dates are available, as well as the number of treatment minutes per month. These data are also provided by Statistics Netherlands.

Third, we use data describing all DI applications between 2010 and 2015, which are provided by the Employee Insurance Agency (UWV). These data contain detailed information on all applications in terms of the health impairments assessment and the subsequent labor market assessment by vocational experts, which determines the remaining earnings potential and the corresponding degree of disability. Both rejected and approved applications are included in our data.[10]

3.2.3 Descriptive statistics

Column (1) of Table 3.1 presents the descriptive statistics for our full sample of employed individuals. The sample contains over 10 million individuals who are employed for at least one month during our observation window (2010–2015). The top panel includes demographics and education as measured in January 2010. The middle panel describes the healthcare measures averaged over the period 2010–2015. Employed individuals have on average €948 in physical healthcare costs and €187 in mental healthcare costs per year. The lower panel includes the employment measures, again averaged over 2010–2015. On average, individuals have a permanent contract for 48% of the months in our observation window, while in 19% of the months, they have a fixed-term contract. In the remaining months, they are not employed.

Columns (2) and (3) of Table 3.1 show separate statistics for workers with a fixed-term contract and those with a permanent contract, respectively. Due to the longitudinal nature of the data, a single individual can change his or her contract type over time. We therefore classify individuals by their contract types as measured in January 2010. Workers with fixed-term contracts are substantially younger, less likely to be Dutch natives, and less educated than permanent workers. They also have lower physical healthcare expenditures but higher mental healthcare expenditures, which might be driven by the age difference. In the lower panel, we see that workers with a fixed-term contract in January 2010 have a permanent contract in 29.4% of the months in 2010–2015. The reverse pattern is much less common: permanent workers spend only 7.6% of the next five years in fixed-term contracts. Permanent workers also have higher hourly wages and work more hours per week. It is important to stress that fixed-term contracts are not merely stepping stones toward permanent contracts for labor market entrants. Fixed-term contracts are

[10]Our observation window is limited by the availability of the DI application data.

Table 3.1: Descriptive statistics for the full sample and selected subsamples

	Employed population[a]	Fixed-term contract in Jan 2010	Permanent contract in Jan 2010	DI applicants
Demographics[c]:				
Age	37.4	32.2	42.5	43.0
Female	46.5%	50.1%	46.3%	52.5%
Dutch native	77.0%	74.9%	84.0%	72.7%
Education unknown	34.9%	17.4%	42.0%	25.1%
Education (if known):				
Low	17.5%	14.1%	15.1%	34.3%
Middle	41.5%	44.5%	37.7%	44.7%
High	41.0%	41.5%	47.2%	21.1%
Annual health measures[d]:				
Mental healthcare expenditures (in €)	187	216	123	1252
Physical healthcare expenditures (in €)	948	846	1032	3709
Mental health treatment (in minutes)	60.0	81.6	43.2	477.6
Employment measures[e]:				
Permanent contract[f]	47.8%	29.4%	81.1%	49.4%
Fixed-term contract[f]	18.7%	50.8%	7.6%	15.7%
Hourly wage	23.5	19.1	26.2	21.2
Monthly number of working hours	77.1	87.5	107.0	75.0
Disability insurance measures[g]:				
DI application probability	0.06%	0.07%	0.05%	100%
DI applications award rate[h]	61.6%	55.1%	63.6%	61.6%
Number of individuals	10,583,956	1,785,327	5,184,711	253,628

(a) All unique individuals who are employed at some point in time between 2010 and 2015. (b) The reference date for the contract type is January 2010. (c) Demographics in January 2010. (d) Health measures are averages computed over the time window 2010—2015. (e) Employment measures are averages computed over the time window 2010—2015. (f) Percentage of months with the corresponding contract type. (g) DI measures computed over the full time-window (2010-2015). (h) Percentage of DI applicants who have been awarded DI benefits.

prevalent throughout the age distribution. Among the individuals who have a fixed-term contract at some point in time, over 25% work under a fixed-term contract for more than three years, and this share is constant across the age distribution (see Appendix Figure 3.A.1).

Finally, column (4) of Table 3.1 shows statistics for the subsample of DI applicants, of whom there are over 250,000 individuals in our observation window. The most striking differences relative to the working population are their higher age, lower level of education, higher healthcare use (across all measures) and slightly lower wage level. Note that these descriptives conceal the substantially higher probability of applying for DI among fixed-term workers: among DI applicants, the share with a fixed-term contract is similar to the share in the employed population, but to a large extent, this is due to the average

age of DI applicants (older individuals are more likely to hold permanent contracts).[11] Nevertheless, individuals with a fixed-term contract are still significantly more likely to apply for DI. Conditional on application, the probability of being awarded DI benefits is slightly greater for individuals with a permanent contract.

3.2.4 DI application probabilities

To derive the DI application probabilities by contract type, we need to account for the two-year waiting period that precedes DI applications. Specifically, the population at risk for applying for DI in month t consists of all individuals employed in month $t - 24$. The at-risk populations for permanent and fixed-term workers between 2010 and 2015 are shown in panel (a) of Figure 3.1. While the at-risk population with a permanent contract decreases steadily, the prevalence of at-risk individuals with fixed-term contracts increases by a similar magnitude. Note also that the number of unemployed individuals increases sharply due to the recession, although this group is an order of magnitude smaller than the population of employed individuals. UI recipients can also file a DI application if they fall ill.

We classify all DI applicants by their contract type 24 months prior to their application date. Most applicants can be classified into one of these three groups (permanent, fixed-term or unemployed), but a small share of approximately 1% cannot.[12] We assign all those without either an employment contract or UI benefits to the fixed-term contract category, as they do not have a responsible employer during the spell of their illness. The number of DI applicants per category between 2010 and 2015 is shown in panel (b) of Figure 3.1. The largest group of DI applicants are those with a permanent contract, but the difference relative to the number from the group of fixed-term workers is considerably smaller than the corresponding difference in the at-risk populations.

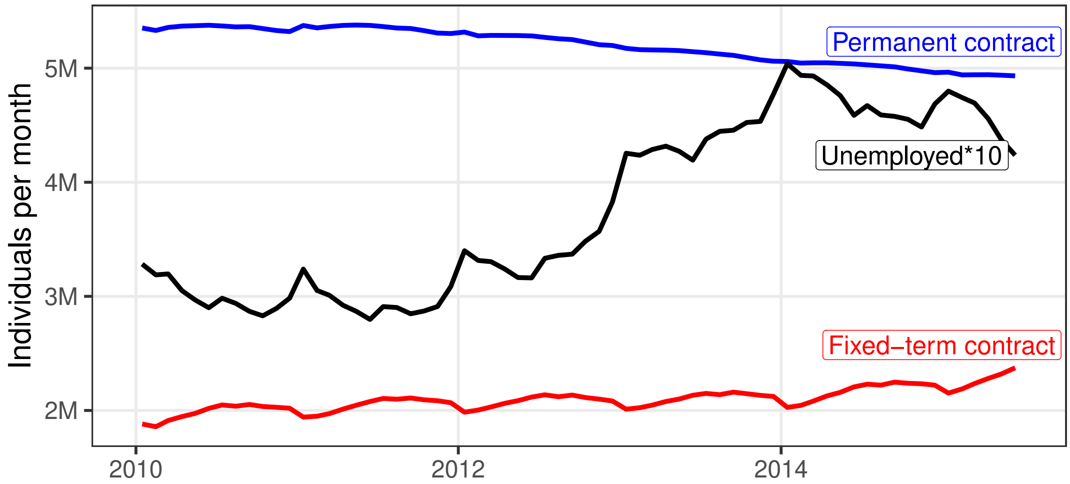
The DI application probability equals the number of DI applicants at time t divided by the size of the at-risk population in month $t - 24$. Panel (c) of Figure 3.1 shows the pronounced difference between the groups. Workers with fixed-term contracts exhibit an application probability approximately 1.5 times as high as those with permanent contracts. The difference is fairly stable until 2013, after which it becomes considerably smaller.

[11] Note that in addition, the sampling in the table makes it infeasible to directly observe relative DI application probabilities because the shares of permanent and fixed-term contracts presented in column (1) represent the number of months, while in column (4), they represent the number of people.

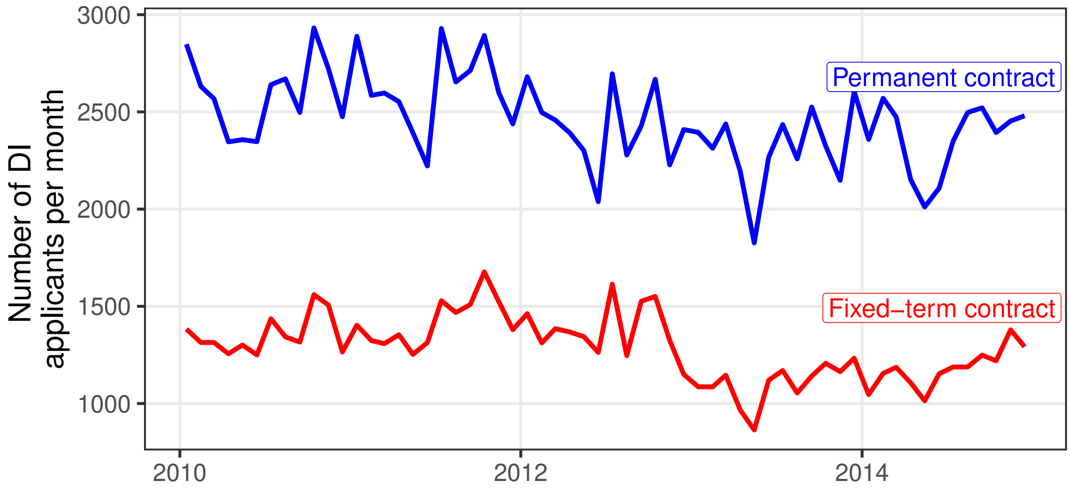
[12] There are several reasons for these applicants remaining unclassified. There are exceptions to the length of the waiting period, meaning that $t - 24$ is not always the relevant month to consider. Furthermore, if a worker falls ill shortly after their contract ends, they are still eligible for DI benefits two years later. In this case, it is difficult to identify the relevant month.

Figure 3.1: Number of workers, DI applicants and DI application probabilities

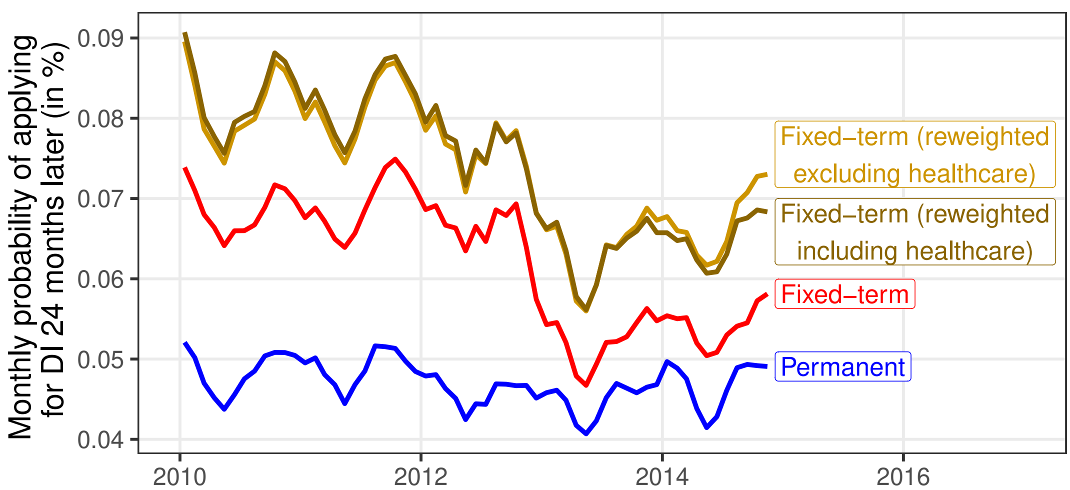
(a) At-risk populations (number of employed individuals)



(b) DI applicants



(c) DI application probabilities (in percentages)



Previewing our analysis in the next section, we briefly explore whether the substantial difference in DI application probabilities between permanent and fixed-term workers stems from the compositional differences shown in Table 3.1. We do so by reweighting the DI application probabilities for fixed-term workers using the distribution of characteristics among the permanent contract workers.[13] We reweight using 48 cells defined by interacting age group, gender and education level, yielding results that are depicted by the yellow line in panel (c) of Figure 3.1. Most notably, conditioning on these demographic characteristics leads to an even larger difference in the DI application probabilities between fixed-term and permanent workers: it is now almost twice as large (before 2013). Additionally, after the reduction in DI application probabilities in 2013, the probability for fixed-term workers remains well above that for permanent workers. The most important explanation is that both the likelihood of securing a permanent contract and the probability of applying for DI increase with age.

Since the probability of applying for DI is clearly correlated with health status, differences in health status – prior to falling ill – between fixed-term and permanent workers might explain the risk premium. Our large number of observations allows for further reweighting using healthcare expenditures. We interact the previously defined cells with five levels of healthcare expenditure (measured in the calendar year prior to month $t-24$) and show the reweighted DI application probabilities as the dark brown line in panel (c) of Figure 3.1. Surprisingly, the results are almost identical to those conditional only on basic demographics. Therefore, once we control for demographics, any remaining health differences are unable to explain the higher probability of applying for DI among fixed-term contract workers.

[13] Reweighting is performed at the monthly level. For month t, we define the share of the population of permanent workers in group j as

$$\alpha_{tj} = \frac{\sum_i^{N_t} I_t(i \in j)}{N_t}.$$

A group j is defined by an interaction of characteristics. Subscript i refers to individuals, and N_t is the total number of permanent contract workers in month t. The reweighted DI application probability for fixed-term workers is the weighted sum of the probability for each group:

$$\tilde{R}_t^T = \sum_j \alpha_{tj} R_{jt}^T$$

3.3 Decomposition framework

To guide the empirical analysis, we now turn to a formal decomposition of the difference in DI application probabilities between fixed-term and permanent workers. Throughout the empirical analysis, we focus on the relative DI risk: the ratio of the probability of applying for DI with a fixed-term contract to that with a permanent contract. We first show that the total relative DI risk can be decomposed into a weighted average of the relative risks under various scenarios. These scenarios are based on the four mechanisms that potentially cause the DI risk gap. We present the main findings here, while the formal proofs and intermediate steps can be found in the Online Appendix. As a starting point, we consider the observed 'raw' relative DI risk:

$$\lambda_{raw} = \frac{P(DI|FT)}{P(DI|P)} \tag{3.1}$$

where FT denotes fixed-term contracts and P denotes permanent contracts. To account for selection, we condition on all observable characteristics. This corresponds to the first step in our decomposition, which results in a reweighted relative DI risk with the weights depending on the distribution of observables for those with a fixed-term contract and for those with a permanent contract.

$$\lambda_{raw} = \alpha_x \cdot \frac{P(DI|FT,x)}{P(DI|P,x)} = \alpha_x \cdot \frac{\pi_{DI}^{FT}}{\pi_{DI}^{P}} = \alpha_x \cdot \lambda_{cond} \tag{3.2}$$

where α_x reflects the impact of selection on the relative risk due to differences in composition. For notational convenience, we define π_{DI}^{FT} and π_{DI}^{P} as the respective risks of applying for DI conditional on having a fixed-term or permanent contract and conditional on all observable characteristics. The corresponding relative DI risk is λ_{cond}.[14]

To analyze differentials stemming from the causal impact of contract type, we next decompose the conditional DI risk into the risk of experiencing a health shock (π_S) and the risk of applying for DI conditional on such a health shock ($\pi_{DI|S}$).[15]

$$\begin{aligned} \lambda_{cond} &= \frac{\pi_S^{FT} \cdot \pi_{DI|S}^{FT}}{\pi_S^{P} \cdot \pi_{DI|S}^{P}} \\ &= \lambda_{shock} \cdot \lambda_{DI|S} \\ &= \lambda_{shock} \cdot \tilde{\lambda}_{DI} \end{aligned} \tag{3.3}$$

[14]Given that all remaining analyses are conditional on observables, we omit the notation regarding conditioning on observables.

[15]Note that our measures of health shocks are imperfect and that we may observe applications for which we observe no prior shock ($\not{S}$). In the Online Appendix, we show that DI applications for which we observe no prior health shock drop out of λ_{cond} under the assumption that the relative risk after a shock is equal to the relative risk when there is no shock ($\lambda_{shock} \cdot \lambda_{DI|S} = \lambda_{no\ shock} \cdot \lambda_{DI|\not{S}}$). We test this assumption and find that it is approximately correct.

Note that from now on, we denote probabilities and relative risks that are *conditional on health shocks* as $\tilde{\pi}$ and $\tilde{\lambda}$, respectively.

We next isolate the third contributor to the DI risk differential, which is due to differences in employer commitment. For this, we further split the conditional probabilities by whether employers are fully responsible for the worker (R) or not ($\cancel{R}$) during the worker's illness. Recall that employers' responsibility for fixed-term workers during illness changed during the time period under investigation. The probability that employers are responsible for fixed-term workers is denoted as $\tilde{\pi}_R^{FT}$ (note that $\pi_R^P = 1$ and hence cancels out; see Online Appendix C.1).

$$\tilde{\lambda}_{DI} \quad = \quad \tilde{\pi}_R^{FT} \cdot \tilde{\lambda}_{DI|R} + \tilde{\pi}_{\cancel{R}}^{FT} \cdot \tilde{\lambda}_{DI|\cancel{R}} \tag{3.4}$$

As the final step in the decomposition analysis, we also split the conditional probabilities with respect to a discrete measure of outside options in the labor market. We denote good labor market prospects with L and bad prospects with $\cancel{L}$. The probability that labor market prospects are good is denoted as $\tilde{\pi}_L$. We assume that labor market prospects are always good for permanent employees ($\tilde{\pi}_L^P = 1$, as they have an ongoing employment contract) but can be good or bad for temporary employees depending on the tightness of the labor market in their sector (see Section 3.4).

$$\tilde{\lambda}_{DI} \quad = \quad \tilde{\pi}_{R,L}^{FT} \cdot \tilde{\lambda}_{DI|L,R} + \tilde{\pi}_{\cancel{L},R}^{FT} \cdot \tilde{\lambda}_{DI|\cancel{L},R} + \tilde{\pi}_{L,\cancel{R}}^{FT} \cdot \tilde{\lambda}_{DI|L,\cancel{R}} + \tilde{\pi}_{\cancel{L}|\cancel{R}}^{FT} \cdot \tilde{\lambda}_{DI|\cancel{L},\cancel{R}} \tag{3.5}$$

By combining (3.2), (3.3) and (3.5), we can summarize the full decomposition framework:

$$\begin{aligned} \lambda_{raw} \quad = \quad & \underbrace{\alpha_x}_{\text{Selection}} \cdot \underbrace{\lambda_{shock}}_{\substack{\text{Causal impact} \\ \text{of falling ill}}} \cdot \\ & \Bigg[\underbrace{\tilde{\pi}_{R,L}^{FT} \cdot \tilde{\lambda}_{DI|L,R}}_{\substack{\text{Responsible employer,} \\ \text{good prospects}}} + \underbrace{\tilde{\pi}_{\cancel{L},R}^{FT} \cdot \tilde{\lambda}_{DI|\cancel{L},R}}_{\substack{\text{Responsible employer,} \\ \text{bad prospects}}} + \\ & \underbrace{\tilde{\pi}_{L,\cancel{R}}^{FT} \cdot \tilde{\lambda}_{DI|L,\cancel{R}}}_{\substack{\text{Not responsible employer,} \\ \text{good prospects}}} + \underbrace{\tilde{\pi}_{\cancel{L}|\cancel{R}}^{FT} \cdot \tilde{\lambda}_{DI|\cancel{L},\cancel{R}}}_{\substack{\text{Not responsible employer,} \\ \text{bad prospects}}} \Bigg] \end{aligned} \tag{3.6}$$

Equation (3.6) shows that the observed 'raw' relative DI application probability can be decomposed into the effect of selection multiplied by the relative risk of experiencing a health shock and the relative risk of applying for DI conditional on experiencing a health shock. The latter term can, in turn, be rewritten as the average of the relative DI risks across four scenarios defined by employer responsibility and labor market prospects, weighted by the relative prevalence of each scenario.

3.4 Empirical analysis

In what follows, we perform regression analyses to decompose the potential contributions of different components to the gap in the DI application probabilities between fixed-term and permanent workers. In line with the sequential nature of this procedure, we first consider the role of selection into contract type in Section 3.4, obtaining supportive evidence for the reweighting results in the previous section and estimating λ_{raw}, λ_{cond} and thereby, indirectly, α_x. Second, in Section 3.4, we estimate the effect of contract type on the probability of falling ill and the subsequent probability of applying for DI (λ_{shock} and $\tilde{\lambda}_{DI}$). Third, we investigate the role of the employer during the two-year waiting period ($\tilde{\lambda}_{DI|R}$ and $\tilde{\lambda}_{DI|\not R}$) in Section 3.4 by exploiting the 2013 reform that established employer responsibility for temporary workers. Finally, we estimate whether the probability of applying for DI depends on labor market prospects in Section 3.4, thereby completing the full decomposition shown in Equation (3.6).

3.4.1 Selection into contract type

Empirical specification

In the first step of our analysis, our aim is to assess the extent to which observable (pre-illness) differences between fixed-term and permanent workers explain the difference in DI application probabilities. To do so, we estimate both the observed relative DI application probability (λ_{raw}) and the relative DI application probability conditional on observables (λ_{cond}) using linear regression models.[16] Compared to the reweighting results shown above, we extend our set of covariates substantially. All employed individuals are included in a monthly panel. Since workers who apply for DI in a given month are no longer part of the at-risk population, we exclude observations of such workers during months that are less than 24 months prior to their application. The key explanatory variables are a set of four employment status dummies E^j_{it}, which equal one if individual i's contract in month t equals j and zero otherwise. Here, j can be (i) a fixed-term contract, (ii) a permanent contract, (iii) unemployment benefits after a fixed-term contract or (iv) unemployment benefits after a permanent contract. We include individuals receiving UI benefits, as such individuals have substantially higher DI application rates. UI benefits might represent an important pathway from employment to DI that would be ignored if we excluded this group.[17] The regression model is as follows:

[16]As a robustness check, we replace the linear specifications with logistic regressions.

[17]To test the robustness of our results, we also estimate a specification in which we simply classify individuals receiving UI benefits after a fixed-term contract as fixed-term contract workers and those receiving benefits after a permanent contract as permanent contract workers. The results are similar.

$$DI_{it} = \sum_{j}^{4} \beta_j E_{it}^{j} + \delta X_{it} + \varepsilon_{it}, \tag{3.1}$$

where DI_{it} as an indicator dummy that is equal to one if individual i applies for DI in period $t + 24$, meaning that he or she fell ill in month t. Differences in the DI application probabilities between the contract types are captured by β. We sequentially add control variables X_{it} to assess the extent to which they explain the differences between contract types. We first add demographic controls, which we then extend with controls for job characteristics and two-year lagged healthcare expenditures. To minimize omitted variable bias, we include all relevant cross-products by using a double LASSO specification following Belloni et al. (2014).[18] Finally, we estimate an upper bound on any potential remaining bias from selection on unobservable characteristics using the methods suggested by Oster (2019).

While our data (technically) allow for the use of individual fixed effects to exploit individual variation in contract type, we choose not to include them. First, we would effectively lose all individuals who never apply for DI, which is the vast majority of our sample. Related, the contract type effects (β) would then be identified solely from individuals who switch between fixed-term and permanent contracts.[19] Furthermore, an individual typically only applies for DI once and then 'leaves' the sample afterwards, limiting the scope for a fixed-effects estimation. We therefore focus on cross-individual variation in contract type and DI risk. Following Abadie et al. (2017), we choose not to adjust our standard errors for clustering, as our sample contains the full Dutch population.

Results

The regression estimates for Equation (3.1) are presented in Table 3.1. As a reference point, we find that a fixed-term contract increases the monthly probability of applying for DI by 0.013 percentage points, which amounts to a relative probability (λ_{raw}) of 1.3. Additional controls are included in the subsequent columns. The inclusion of age, gender, nationality, education level, family composition and population density[20] as controls in column (2) *increases* the relative application probability to 1.68. This is mainly due to age differences: younger individuals are more likely to have fixed-term contracts and are less likely to apply for DI. Controlling for job characteristics (wage, working hours and sector) reduces the difference in the probabilities to 43% (column (3)).

[18] See Online Appendix Section C.2 for a discussion of the double LASSO specification.

[19] To study the effect of switches in more detail, we estimate a robustness specification in which the subsample of individuals who recently switched contract types are included as separate employment contract groups.

[20] Population density is measured at the municipality level and consists of six categories ranging from "nonurban" (less than 500 individuals per km^2) to "very urban" (more than 2,500 individuals per km^2).

Table 3.1: Regression results for monthly DI application risk for temporary and permanent workers

	DI application							
	(1)	(2)	(3)	(4)	(5)	(6)	$(7)^h$	$(8)^i$
Fixed-term contract coefficient	0.013	0.027	0.018	0.019	0.019	0.020	0.020	0.020
Estimated DI application prob.: Fixed-term[a]	0.056	0.068	0.062	0.062	0.062	0.062	0.062	
Estimated DI application prob.: Permanent[a]	0.043	0.040	0.043	0.043	0.043	0.043	0.043	
Relative probability[b]	1.30	1.68	1.43	1.45	1.45	1.45	1.45	
Robustness specifications:								
DI award[d]	1.13	1.60	1.44	1.45	1.45	1.45	1.45	
Logistic regression[c]	1.30	1.68	1.43	1.45	1.45	1.45	1.45	
Merge UI groups[e]	1.39	1.76	1.39	1.41	1.40	1.40	1.40	
Include switcher groups	1.23	1.63	1.38	1.40	1.40	1.40	1.40	
Quarter-of-year dummies		X	X	X	X	X	X	
Demographic controls[f]		X	X	X	X	X	X	
Job controls[g]			X	X	X	X	X	
Lagged health controls				X	X	X	X	
Regional fixed effects					X			
Sector x education controls						X		
All relevant cross-products							X	
Number of control dummies	0	50	133	143	546	703	672	
R^2 baseline regression	0.026	0.060	0.065	0.089	0.090	0.091	0.108	
Observations (ind.*month)	475 million							

The sample in all columns includes the full observation period (January 2010 to June 2015). All regressions include dummies for receiving UI benefits after a fixed-term contract and for receiving UI benefits after a permanent contract (estimates not reported). All coefficients in this table are statistically significant with P values < 0.0001. See Appendix Table 3.A.1 for the specific control variables included. (a) Average predicted DI probability if all individuals had a fixed-term contract (or, correspondingly, a permanent contract). (b) Ratio of estimated DI probability among fixed-term workers and estimated DI probability among permanent workers. (c) Relative probabilities from the logistic regressions are calculated as e^{β}, where β is the fixed-term contract coefficient from the logistic regression. (d) Regression of the relative probability of receiving a DI award. (e) Regression of the relative probability of applying for DI with UI groups merged. (f) Age, gender, nationality and education level. (g) Wage, number of working hours and sector of employment. (h) Cross-products are included based on their predictive power over DI applications and contract type using a double LASSO specification; see Online Appendix Section C.2 for details. (i) Upper bound on selection on unobservable characteristics using Oster (2019) analysis: $\beta^* = \tilde{\beta} - (\beta^0 - \tilde{\beta})\frac{1.3*\tilde{R}-\tilde{R}}{\tilde{R}-R^0}$.

In column (4), we include 10 dummies for the level of lagged healthcare expenditures to control for differences in health status. The effect of lagged health on DI risk is substantial: individuals with lagged healthcare expenditures in the top 10% of the distribution are 40 times more likely to apply for DI than those in the bottom 10% (see Appendix Table 3.A.2 for the healthcare coefficients from the regression in column (5)). The effect of large healthcare expenditures is ten times as large as the effect of contract type. Nevertheless, the difference in DI application probabilities due to contract type remains unchanged. Selection based on health appears limited and does not explain the difference in the DI application probabilities between contract types.

Columns (6) and (7) in Table (3.1) show that our key regression results remain constant with the inclusion of more extensive sets of controls. This holds for the inclusion of 408 municipality dummies, as well as the addition of interactions between sectors (70) and education levels (10) to control for more specific occupational characteristics that may drive the risk of disability. We also consider a specification that allows for interactions between all control variables. To balance the added value of these interactions and the risk of overfitting, we estimate a double LASSO specification (Belloni et al. (2014)). The results, shown in column (7), again yield constant contract-type effects.

To address the concern that there remains selection based on unobserved characteristics, we follow the approach proposed by Oster (2019). The idea is to compute an upper bound for the contract type effect by extrapolating changes in the coefficient estimates as additional covariates are added, weighted by the corresponding change in R-squared. We use the specification in column (3) as our baseline and the specification in column (7) as our extended model. Given the stability of the coefficient estimates when moving from column (3) to column (7) and the strong (relative) increase in R^2, the effect of selection on unobservable characteristics is limited: the calculated upper bound on the fixed-term contract coefficient is 0.020 (compared to our estimate of 0.019 in column (7)).[21]

Table 3.1 also shows the results from four robustness specifications, and these results are very similar to the baseline results. First, we examine DI awards instead of DI applications. Given the high level of similarity between the results, we conclude that contract type has little impact on the probability that a DI application is approved. Second, we show that the marginal contract type effects are almost identical to the average effects with a logit specification. Presumably, this reflects the fact that our large sample size allows for a sufficiently flexible specification of the linear model. Third, we reclassify

[21]Note that we use the restricted estimator and the R_{max} value proposed by Oster (2019); $R_{max} = 1.3\tilde{R}$. Given our large sample size, it is not computationally feasible to estimate the unrestricted estimator. Using larger R_{max} values, e.g., $R_{max} = 2\tilde{R}$, does not alter our conclusion.

individuals who apply for DI benefits while receiving UI benefits.[22] This slightly increases the absolute DI application probabilities (the application probability for UI beneficiaries is high) and slightly decreases the resulting relative probability. The decrease in the relative probability is because the DI application probability is high for UI beneficiaries regardless of their contract type prior to entering UI. Last, we reclassify individuals whose contract type changed in the last six months. More precisely, we add two employment statuses: (1) switches from fixed-term to permanent contracts and (2) switches from permanent to fixed-term contracts. The DI application probabilities of the switcher groups are very similar to those of the nonswitcher groups.[23] The inclusion of the switcher groups does not alter the relative probabilities. The gap in DI application probabilities is thus neither caused nor masked by contract-type effects that carry on after contract-type switches. This indicates that anticipation of contract-type switches, or strategic behavior in the reporting of sickness, does not seem to cause the gap in application behavior.

The parametric estimates strengthen our confidence in the findings from the reweighting exercise shown in Figure 3.1. Despite stark compositional differences between workers with fixed-term and permanent contracts, the higher DI application probability among fixed-term workers is barely explained by these differences. Demographic differences conceal the fact that the difference is actually slightly larger than raw numbers suggest, while differences in prior health have a negligible additional effect. The extrapolation based on Oster (2019) suggests that additional unobserved characteristics are unlikely to change the results substantially. We conclude that selection of relatively unhealthy workers into fixed-term contracts does not explain the high DI application probability among fixed-term workers.

3.4.2 Impact on the probability of falling ill

In the previous sections, we uncovered selection into contract types by worker and job type and showed that the gap in DI application probabilities is still present after correcting for compositional differences. We now decompose this gap into the impact of contract type on the relative risk of falling ill (λ_{shock}) and the subsequent relative DI application probability conditional on being ill ($\lambda_{DI|S}$). Contract type may impact the probability of falling ill through, for example, differences in occupational hazards, fewer prevention activities or increased stress due to a lack of job security.

[22] Individuals who receive UI benefits when their last contract was temporary are included in the temporary contract category. Individuals who receive UI benefits when their last contract was a permanent contract are included in the permanent contract category.

[23] Individuals who switch from fixed-term to permanent contracts have DI application probabilities similar to those of individuals with permanent contracts, and likewise, individuals who switch from permanent to fixed-term contracts have DI application probabilities similar to those of individuals with a fixed-term contract.

Empirical specification

We regress an indicator for falling ill on contract type. In all model variants, we add the same set of controls as in the previous analyses and derive the implied relative risk of falling ill (λ_{shock}) using these regression results. Subsequently, we assess how the difference in the probability of falling ill may contribute to the difference in the DI application probabilities by considering the probability of applying for DI within the subsample of individuals who have fallen ill ($\lambda_{DI|S}$).

Administrative data in the Netherlands do not provide information on employee sick leave. We therefore proxy for this event by identifying increases in healthcare expenditures. We define a negative health shock using various thresholds. In the baseline model, we define a mental health shock as beginning a mental health treatment trajectory. A physical health shock is defined as an increase in annual healthcare expenditures from below the median for the population (approximately €150) to above the 90th percentile (approximately €2,400). Accordingly, the at-risk population is defined as the working population without mental health treatment or with healthcare expenditures below the median. This results in samples of workers who are sufficiently 'healthy' such that they can, according to our definition, experience a negative health shock. Our regression model for experiencing a negative health shock is as follows:

$$H_{it} = \sum_{j}^{4} \beta_j^S E_{it}^j + \delta^S X_{it} + \varepsilon_{it}^S \,, \tag{3.2}$$

where H_{it} is an indicator for individual i experiencing a negative health shock in period t. Again, our interest lies in the estimate of β^S, which in this case captures the association between the type of contract and the likelihood of experiencing a negative health shock. We sequentially add the same rich set of control variables as in Subsection 3.4, with the exception of healthcare expenditures. We thus reduce the scope for omitted variable bias, strengthening the claim that β^S captures causal effects. Note that by defining our outcome as a *change* in health status, we essentially control for baseline health and thus relax the assumptions required for a causal interpretation of β^S.

We next estimate whether the probability of applying for DI *c*onditional on a health shock differs by contract type. Together with the results of Equation (3.2), this allows us to decompose the DI application differential into a part that is due to differences in the probability of a health shock and a part that is due to differences in the DI application probabilities *conditional on falling ill*. Specifically, we re-estimate Equation (3.1) for the subsample of employees who experience a negative health shock. We denote the month (for mental health shocks) or year (for physical health shocks) in which the shock occurs as t_i^* and include only observations from the six months before and six months after t_i^*.[24] Similar to the outcome variable in our earlier analysis, the outcome variable is an indicator equal to one if a DI application is filed at time $t + 24$ and zero otherwise.

Results

Tables 3.2 and 3.3 show the regression estimates for the risk of experiencing a mental or physical health shock, respectively.[25] Additional controls are included in columns (2)-(6). The tables also present the implied conditional relative DI application probabilities for both mental and physical health shocks. Any differences between the unconditional and conditional relative DI application probabilities follow from the differences in the probability of experiencing a negative health shock.

Without controlling for any characteristics, employees with fixed-term contracts have a 40% higher risk of experiencing a mental health shock (Table 3.2, column (1)). For a physical health shock, their risk is 18% lower (Table 3.3, column (1)). Both of these observed differences decrease markedly when we control for demographics, as shown in column (2). When adding job controls, regional fixed effects and interactions between sector and education level, there is little impact on the relative risks. With the most extensive sets of controls, which includes all relevant cross products, we find that fixed-term contracts increase the risk of a mental health shock by 8%, while for physical health problems, the risk is almost identical between the two contract types.[26] When computing an Oster (2019) upper bound on the contract type effect, which takes into account selection based on unobservable characteristics (column (7)), the estimated contract type coefficients become even smaller. These results are robust to alternative health shock definitions that apply a higher cutoff for time spent in mental health treatment or physical healthcare expenditures.

We next consider the probability of applying for DI conditional on experiencing a health shock ($\lambda_{DI|S}$). For comparison, Tables 3.2 and 3.3 show both the unconditional probability of applying for DI (as also reported in Table 3.1) and the probability conditional on experiencing a health shock. Both for workers with fixed-term contracts and those with permanent contracts, the estimated probability of applying for DI after a mental health shock is approximately 10 times as large as the unconditional probability. This indicates that the mental health shocks that we identify are indeed strong predictors of

[24]The exact timing of the health shock is difficult to observe because (i) the health expenditure data are annual and (ii) there may be a wait time for certain types of healthcare such that expenditures increase with some delay. To address these issues, we also include the six months prior to the health shock.

[25]One might expect the 2013 reform to affect the probability of falling ill for temporary workers, but we find that results based on the post-2013 period only are very similar.

[26]Additionally, the evolution of health after a health shock is similar for both contract types.

Table 3.2: Regression results for a negative mental health shock (> 0 minutes of treatment) and a subsequent DI application

	Mental health shock							DI application	
	(1)	(2)	(3)	(4)[e]	(5)	(6)[f]	(7)[g]	Unconditional	Conditional
Fixed-term contract coefficient	0.041	0.015	0.010	0.009	0.009	0.007	0.005	0.019	0.192
Estimated health shock prob.: Fixed-term[a]	0.142	0.125	0.122	0.121	0.121	0.121		0.062	0.639
Estimated health shock prob.: Permanent[a]	0.101	0.109	0.112	0.112	0.112	0.114		0.043	0.446
Relative risk[b]	1.40	1.14	1.09	1.08	1.08	1.06		1.45	1.43
Robustness specifications:									
Larger mental health shock (> 200 minutes)	1.56	1.19	1.10	1.08	1.08	1.06		1.45	1.41
Number of observations	286 million							475 million	4 million
Quarter-of-year dummies		X	X	X	X	X	X	X	X
Demographic controls[c]		X	X	X	X	X	X	X	X
Job controls[d]			X	X	X	X	X	X	X
Regional fixed-effects				X					
Sector x education controls					X				
All relevant cross-products						X	X		

All coefficients in this table are statistically significant with P values < 0.0001. See Appendix Table 3.A.1 for the specific control variables included. (a) Average predicted monthly probability of experiencing a health shock if all individuals had a fixed-term contract (or, correspondingly, a permanent contract). Probability is reported as a percentage; i.e., 0.1 means a 0.1% risk of experiencing a shock in any given month. (b) Ratio of the estimated probability for fixed-term workers and the estimated probability for permanent workers. (c) Age, gender, nationality and education level. (d) Wage, number of working hours and sector of employment. (e) Based on a random subsample of 10 million observations due to computational load. (f) Cross-products are included based on their predictive power over DI applications and contract type using a double LASSO specification; see Online Appendix Section C.2 for details. (g) Upper bound for selection on unobservable characteristics using Oster (2019) analysis: $\beta^* = \tilde{\beta} - (\beta^0 - \tilde{\beta})\frac{1.3*\tilde{R}-\tilde{R}}{\tilde{R}-R^0}$

Table 3.3: Regression results for a negative physical health shock ($> 90^{th}$ percentile) and a subsequent DI application

	Physical health shock							DI application	
	(1)	(2)	(3)	$(4)^e$	(5)	$(6)^f$	$(7)^g$	Unconditional	Conditional
Fixed-term contract coefficient	-0.052	-0.004	-0.013	-0.012	-0.013	-0.012	-0.009	0.019	0.043
Estimated health shock prob.:Fixed-term[a]	0.238	0.521	0.519	0.520	0.520	0.520		0.062	0.133
Estimated health shock prob.: Permanent[a]	0.291	0.532	0.537	0.536	0.536	0.536		0.043	0.090
Relative risk[b]	0.82	0.98	0.97	0.97	0.97	0.97		1.45	1.48
Robustness specifications:									
Larger physical health shock ($> 99^{th}$ percentile)	0.56	1.02	0.96	0.96	0.97	0.97		1.45	1.58
Number of observations	192 million							475 million	11 million
Quarter-of-year dummies		X	X	X	X	X	X	X	X
Demographic controls[c]		X	X	X	X	X	X	X	X
Job controls[d]			X	X	X	X	X	X	X
Regional fixed-effects				X					
Sector x education controls					X				
All relevant cross-products						X	X		

All coefficients in this table are statistically significant with P values < 0.0001. See Appendix Table 3.A.1 for the specific control variables included. (a) Average predicted monthly probability of experiencing a health shock if all individuals had a fixed-term contract (or, correspondingly, a permanent contract). Probability is reported as a percentage; i.e., 0.1 means a 0.1% risk of experiencing a shock in any given month. (b) Ratio of the estimated probability for fixed-term workers and the estimated probability for permanent workers. (c) Age, gender, nationality and education level. (d) Wage, number of working hours and sector of employment. (e) Based on a random subsample of 10 million observations due to computational load. (f) Cross-products are included based on their predictive power over DI applications and contract type using a double LASSO specification; see Online Appendix Section C.2 for details. (g) Upper bound for selection on unobservable characteristics using Oster (2019) analysis: $\beta^* = \tilde{\beta} - (\beta^0 - \tilde{\beta})\frac{1.3*\tilde{R}-\tilde{R}}{\tilde{R}-R^0}$

later DI application. Comparing the results between the two contract types, we find that the probability of applying for DI is 43% higher for fixed-term workers, which is close to the unconditional difference (45%).

After a physical health shock, the conditional DI application probability is approximately twice as high as the unconditional probability. For workers who have experienced a physical health shock, we find a large difference in the DI application probabilities (48%) between fixed-term workers and permanent workers, and this difference is even higher than the unconditional difference (45%). Therefore, any differences in the risk of falling ill between contract types are confined to mental shocks. However, the increased risk of mental health shocks does not explain the high probability of applying for DI among fixed-term workers. Instead, it seems that the divergence starts to occur only *after* the onset of illness.

3.4.3 Employer incentives during the waiting period

After falling ill, employees face a two-year waiting period during which employer incentives and obligations to support reintegration differ by contract type. Those with permanent contracts receive support from their employer, who is obliged to monitor their progress and actively facilitate their rehabilitation. Employers also face financial consequences if their employees enter DI through experience rating: DI contributions depend on the inflow of their employees into DI during the previous 10 years. Until 2013, for fixed-term workers, these employer responsibilities ended when their contract expired.

As discussed in Section 3.2, for fixed-term workers, the role of the employer during the waiting period prior to DI application became much more similar to that for permanent workers beginning in 2013. Using the notation we introduced in Section 3.3, the relative DI application probability prior to 2013 corresponds to $\tilde{\lambda}_{DI|\not{R}}$. From 2013 onward, the relative DI application probability corresponds to $\tilde{\lambda}_{DI|R}$.

Empirical specification

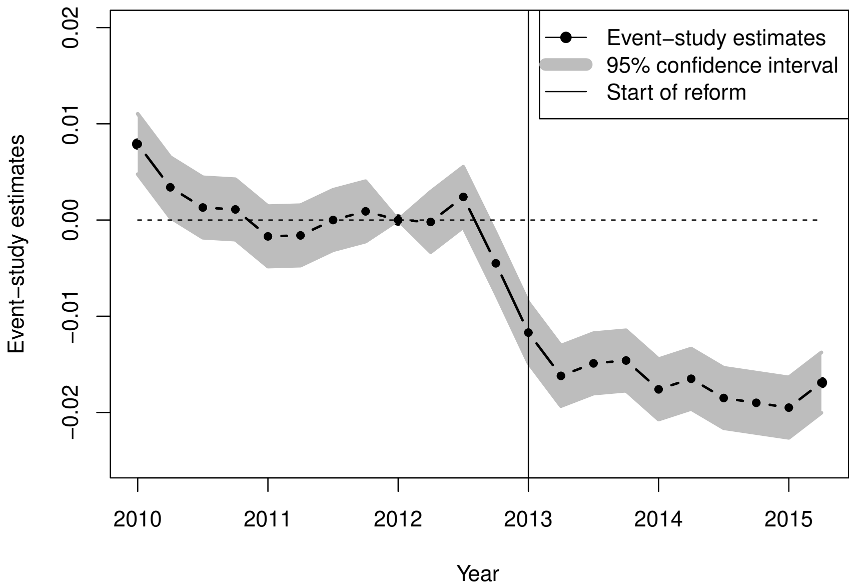
To identify the importance of employer incentives during the waiting period, we exploit the 2013 reform using a difference-in-differences (DiD) strategy. Since permanent workers were unaffected by the reform, we use these workers as the control group. We derive our specification by adding interaction terms between the contract type dummies and a dummy for post-2013 observations to the specification in Equation (3.1):

$$DI_{it} = \sum_{j}^{4} \beta_j E_{it}^{j} + I_{t>2013} \sum_{j}^{4} \delta_j E_{it}^{j} + \gamma X_{it} + \varepsilon_{it}, \tag{3.3}$$

Since X_{it} includes quarter-of-year dummies, this specification corresponds to a conventional DiD model with fixed-term workers as the treatment group and permanent workers (for whom nothing changed after 2013) as the control group. The model enables the estimation of $\tilde{\lambda}_{DI|R}$ and $\tilde{\lambda}_{DI|\not{R}}$. We estimate the DiD model only on samples of workers who have experienced a negative health shock. The reason is that the 2013 reform primarily aimed to increase employer commitment to their fixed-term workers *during their illness*.

To assess the validity of the parallel trends assumption, we first perform an event-study analysis; i.e., we interact the contract-type dummies with the quarter-of-the-year dummies in the regression.[27] For the event-study analysis, we use the DI applications of *all workers* (thus, the probability of applying for DI is not conditional on experiencing a health shock) as this enables us to use a substantially longer time window.[28] Figure 3.1 shows the event-study estimates for fixed-term contracts, in which the first quarter of 2012 is used as the baseline. Prior to the 2013 reform, the difference in the DI application probabilities between permanent and temporary workers is fairly constant over time, implying that the parallel trends assumption holds. Only in the last quarter of 2012 the DI risk for temporary workers decreases substantially relative to that for permanent workers. This is likely due to anticipation among employers given that their inflow into

Figure 3.1: Event-study estimates for the 2013 reform for fixed-term contracts



[27]The regression model is:

$$DI_{it} = \sum_{j}^{4} \beta_j E_{it}^j + \sum_{Q=2010Q1}^{2015Q2} \sum_{j}^{4} \delta_{jQ} E_{it}^j + \gamma X_{it} + \varepsilon_{it}$$

.

[28]Only a short period of time around health shocks is used in the conditional DI regressions.

DI during the years prior to 2013 was used to determine their DI insurance contributions in 2013. The gap in DI risk decreases further in the first two quarters of 2013 and remains constant afterwards.

Results

To obtain a single estimate that can be incorporated into the decomposition, we focus on our simple DiD estimation (specification (3.3)). Table 3.4 shows the estimated conditional DI application probabilities from before and after 2013. In contrast to our earlier decomposition stages, in this part of the decomposition, we explain a substantial part of the observed difference in application probabilities. For both mental and physical health shocks, the gap in application probabilities is almost halved: the relative probability drops from 1.59 to 1.34 for those with mental health shocks and from 1.63 to 1.27 for those with physical health shocks. Accordingly, pre-2013 differences in employer incentives and obligations during the waiting period explain a substantial part of the gap in

Table 3.4: Estimation results: Difference-in-differences model for DI applications conditional on experiencing a health shock

Panel A: DI application conditional on experiencing a mental health shock (> 0 minutes)		
	Pre 2013	Post 2013
Estimated DI application prob.: Fixed-term[a]	0.725	0.591
Estimated DI application prob.: Permanent[a]	0.455	0.442
Relative risk[b]	1.60	1.34
Robustness specifications:		
DI award	1.55	1.45
Larger mental health shock (> 200 minutes)	1.57	1.34
Number of observations	1.5 million	2.7 million

Panel B: DI application conditional on experiencing a physical health shock (> 90^{th} percentile)		
	Pre-2013	Post-2013
Estimated DI application prob.: Fixed-term[a]	0.161	0.100
Estimated DI application prob.: Permanent[a]	0.099	0.079
Relative risk[b]	1.63	1.27
Robustness specifications:		
DI award	1.53	1.41
Larger physical health shock (> 99^{th} percentile)	1.71	1.36
Number of observations	6.2 million	4.7 million

Regressions contain the same control variables as the regression in Table 3.1, column (4): quarter-of-year dummies, demographics and job controls. See Appendix Table 3.A.1 for the specific control variables included. (a) Average estimated monthly probability of applying for DI if all individuals had a fixed-term contract (or, correspondingly, a permanent contract). (b) Ratio of estimated DI application probabilities for fixed-term and permanent workers.

the DI application probabilities.[29,30] At this point, it should be emphasized once more that the 2013 reform did not fully offset the initial differences in employer incentives and obligations that existed up to that point. The estimated reform effect therefore provides a lower bound on the total potential importance of employer incentives and obligations. Further analyses, as discussed in Appendix Section 3.A, suggest that the majority of the reform effect can be attributed to increased monitoring, while the effect of experience rating appears to be limited.[31]

The reform affected the probability of being awarded DI benefits to a lesser extent, as the relative risks of receiving a DI award decrease to 1.45 and 1.41 for mental and physical health shocks, respectively. Therefore, conditional on applying for DI, the probability of being awarded DI benefits increased for fixed-term workers. One potential explanation is that increased monitoring affected mostly fixed-term workers who would have been denied DI benefits if they had applied. For individuals hit by a larger mental or physical health shock, the estimated results are similar.

3.4.4 Labor market prospects of ill employees

Having considered the risk of falling ill and the role of the employer during illness, we now turn to the final stage during which differences in the probability of applying for DI might arise: the decision to apply for DI after two years of illness. Workers with fixed-term contracts who have been sick for up to two years, during which time their contract ended, face very different labor market prospects than ill workers with a permanent contract. As has been argued in the literature, such differences in outside options may well explain the higher propensity to apply for DI benefits among vulnerable groups in the labor market, such as those with fixed-term contracts (Autor & Duggan, 2003).

Empirical specification

To assess the importance of the labor market prospects of ill employees, we consider sector-level labor market tightness as a proxy for the labor market prospects of ill workers. Labor market prospects are by definition worse for ill workers whose contracts have expired, but

[29] Note that the drop in the conditional DI application probability for individuals with a permanent contract who experienced a physical health shock – as reported in Table 3.4 – reflects the effect of right censoring. Individuals who experience a health shock prior to 2013 are often also in the risk sample after 2013 due to uncertainty about the timing of the shock, but the reverse does not hold.

[30] Prinz & Ravesteijn (2020) reach a similar conclusion regarding the impact of the reform on temporary agency workers, who form a (small) subset of all fixed-term workers.

[31] The 2013 reform simultaneously introduced extra financial incentives and monitoring obligations for employers. We can extend our DiD model and exploit the fact that the incentive effects were proportional to firm size. In doing so, we can disentangle the importance of the various elements – as shown in Appendix Section 3.A. Increased monitoring and the one-year assessment account for approximately 80% of the total effect, whereas the introduction of experience rating accounts for approximately 20%.

in tight labor markets, the difference in prospects relative to workers with a permanent contract is likely to be smaller. We categorize 70 sectors as “tight” or “loose” based on the percentage of vacancies relative to the number of filled jobs (see Online Appendix Section C.3 for the categorization of each sector and the distribution of contract types across sectors). Approximately 15% of all employment contracts are classified as belonging to a tight labor market.[32]

The tight and loose labor market categories are incorporated into the DiD specification from the previous subsection. Accordingly, we allow the 2013 reform to have different treatment effects in tight and loose labor markets. This requires labor market tightness to evolve similarly over time in tight and loose labor markets. In Online Appendix Section C.4, we do indeed find similar overall trends. The DiD analyses provide an estimate of how the probability of applying for DI after experiencing an illness differs by the degree of labor market tightness before and after the 2013 reform.

Results

Table 3.5 shows the average estimated DI application probability conditional on experiencing a health shock, and health shocks are now also stratified by the type of labor market (“tight” or “loose”). As a result, we obtain estimated probabilities for sets of workers with a health shock, the same contract type, and the same degree of labor market tightness, measured before and after 2013.

First, after the reform, the estimated conditional probability of applying for DI among permanent workers who experienced a mental or physical health shock is fairly similar to the corresponding value before the reform. In each case, labor market tightness seems more or less irrelevant to the DI application decision. This lends credence to the idea that the existence of an employment contract – and a corresponding employer commitment – renders alternative labor market opportunities unimportant in the decision of whether to apply for DI.

Also in line with expectations, the DI application probability is generally higher for fixed-term workers than for permanent workers, and this gap shrinks after 2013 regardless of the degree of labor market tightness. More strikingly, the gap between permanent and fixed-term workers is substantially smaller in tight sectors than in loose sectors. This is the case both before and after 2013. For example, before 2013, for those with mental health shocks, the relative DI application probability is 1.70 in loose sectors and only 1.38 in tight sectors. Once we consider tight sectors in the post-2013 period, the relative probability of applying for DI among fixed-term workers decreases to only 1.10 for mental

[32] The distribution of contract types in tight sectors is comparable to the distribution in loose labor markets; see Online Appendix Section C.4.

Table 3.5: Regression results for DI applications conditional on a negative health shock and stratified by tight vs. loose labor market sectors

Panel A: Mental health shock (> 0 minutes)				
	Pre-2013		Post-2013	
Labor market tightness[a]	Loose	Tight	Loose	Tight
Estimated DI prob.: fixed-term[b]	0.763	0.646	0.611	0.494
Estimated DI prob.: permanent[b]	0.448	0.467	0.432	0.451
Relative risk[c]	1.70	1.38	1.42	1.10
Robustness specifications:				
DI award	1.67	1.28	1.54	1.16
Larger mental health shock (> 200 minutes)	1.67	1.33	1.41	1.08
Number of observations	1,326,962	186,101	2,347,014	319,721
Panel A: Physical health shock (> 90^{th} percentile)				
	Pre-2013		Post-2013	
Labor market tightness[a]	Loose	Tight	Loose	Tight
Estimated DI prob.: fixed-term[b]	0.177	0.138	0.110	0.070
Estimated DI prob.: permanent[b]	0.097	0.094	0.076	0.072
Relative risk[c]	1.82	1.47	1.45	0.98
Robustness specifications:				
DI award	1.75	1.16	1.61	0.92
Larger physical health shock (> 99^{th} percentile)	1.93	1.43	1.55	1.00
Number of observations	5,551,666	607,829	4,241,347	456,050

Regressions contain the same control variables as in the regression reported in Table 3.1, column (4): quarter-of-year dummies, demographics and job controls. See Appendix Table 3.A.1 for the specific control variables included. (a) Average predicted probability of applying for DI if all individuals worked in a sector with a tight/loose labor market before/after 2013. (b) Average predicted probability of applying for DI if all individuals had a fixed-term contract (or, correspondingly, a permanent contract). (c) Ratio of the estimated probability for fixed-term workers and the estimated probability for permanent workers.

health shocks and to 0.98 for physical health shocks.[33] These findings suggest that poor labor market prospects contribute substantially to the increased propensity to apply for DI among fixed-term workers. In Appendix Table 3.A.3, we further show that the relative application probabilities in the various scenarios are similar if we do not condition on health shocks. Hence, employer incentives during the waiting period and labor market prospects seem to be equally important for the sample of individuals for whom we do not observe a health shock.

Considering DI awards instead of applications, we find that the impact of labor market

[33]It should be noted that this exercise yields a lower bound on the impact of labor market prospects: even in sectors with tight labor markets, prospects are by definition slightly better for individuals with a permanent contract than for individuals whose fixed-term contract has expired.

tightness is slightly larger. This indicates that labor prospects are an important determinant for all fixed-term workers, regardless of the severity of their medical condition. Labor market prospects appear to play an equally important (or even more important) role for workers who are likely to be awarded DI benefits as for workers who are less likely to be awarded DI benefits. This is confirmed by the fact that when we consider workers who are hit by a larger mental or physical health shock, the results change little.

3.5 Findings in perspective

We have studied four potential mechanisms that may explain the increased likelihood of applying for DI observed among individuals with fixed-term contracts. In what follows, we summarize the findings and decompose the raw relative DI application probability into its various sources using the framework explained in Section 3.3. We relate these findings to existing evidence in the literature. The implied contributions to the observed DI application differential are depicted in Figure 3.1. Factors below the (dashed) 0% line decrease the relative probability of applying for DI among fixed-term workers, while those above increase the relative probability.

As a reference point, the observed relative DI application probability (λ_{raw}) equals 1.30, which implies that over the time period that we consider, individuals with a fixed-term contract are 30% more likely to apply for DI benefits than permanent workers. In the figure, this percentage corresponds to the sum of the positive and negative contributing factors. Once we correct for compositional differences, the relative probability increases to 1.48 (λ_{cond}). This is illustrated in Figure 3.1 by the dark green blocks below the (dashed) 0% line.

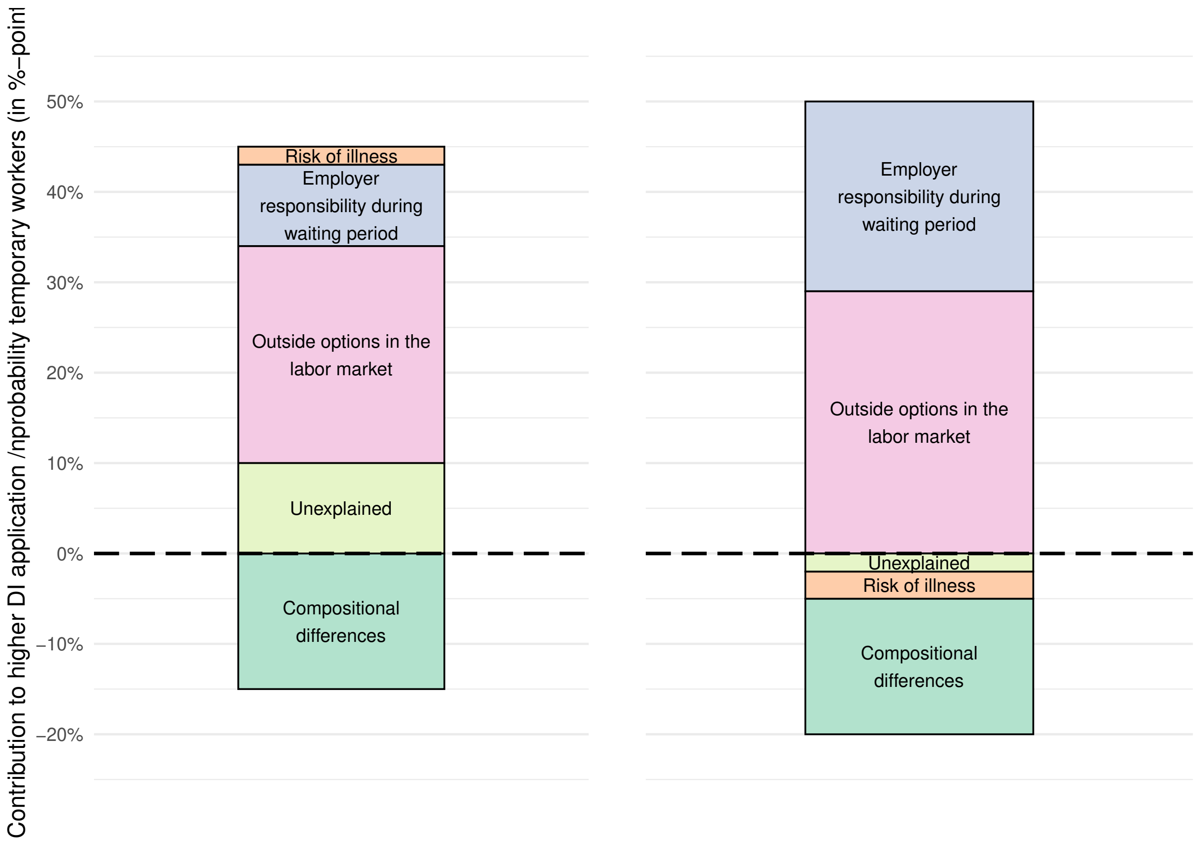
When considering the probability of falling ill, we distinguish between mental and physical health shocks. The risk of experiencing a mental health shock is 10% higher for fixed-term workers, while the probability of filling a DI application conditional on experiencing such a health shock is still 43% higher for fixed-term workers. Contract type has a negligible effect on the risk of experiencing a physical health shock, and consequentially, the difference in the DI application probabilities remains fairly constant when conditioning on whether an individual has experienced a physical health shock.

The role of the employer *after* the onset of illness is substantial: especially for physical health conditions, the gap in the DI application probabilities decreases by a large amount after the reform in 2013, which equalized monitoring by and the financial incentives of employers during the two-year waiting period. While somewhat smaller for mental health, employer incentives still explain a substantial share of the higher DI application probability among fixed-term workers.

Last, workers with fixed-term contracts appear to be more susceptible to labor market

Figure 3.1: Decomposition of relative DI application probability

(a) Mental health (b) Physical health



conditions than permanent workers in terms of their decision to file a DI application. The gap in the application probabilities is considerably lower in tight labor markets. Once we shut down all four mechanisms by comparing workers (i) with similar characteristics, (ii) who have experienced a health shock, (iii) after 2013, (iv) in a tight labor market, the gap in the DI application probabilities almost disappears. As illustrated in Figure 3.1 by the blue and purple blocks, it is mainly employer responsibilities during illness and differences in labor market prospects, respectively, that increase the relative DI application probability. The remaining unexplained difference (light green) is small.

Our conclusions change little when we consider DI awards instead of DI applications. Individuals with fixed-term contracts are substantially more likely to be awarded DI benefits, but the differential is not explained by sorting into contract type or by the causal impact of contract type on the probability of falling ill. However, both employer incentives and the labor market prospects of sick employees explain a large portion of the difference in DI award risks.

One important takeaway from our analysis is that employers do not offer fixed-term contracts specifically to workers with health conditions. This contrasts with previous research on selection into contract type, which finds that individuals in ill health are less likely to obtain a permanent contract (Wagenaar et al., 2012). Taking a broader perspective, there is also evidence from the US that the additional employer responsibilities stipulated in the Americans with Disabilities Act (ADA) lowered the chances of disabled workers being hired (Acemoglu & Angrist, 2001). In our setting, however, the a priori health conditions of workers are probably largely unobserved by employers. Given that our health proxies are strongly predictive of DI applications, one could therefore argue that employers have limited ability to select on the more severe health issues that could lead to a DI application.

In line with other research, we do find that the use of fixed-term employment contracts increases the prevalence of mental health problems (Kim et al., 2012; Virtanen et al., 2005; Benach et al., 2014). Nevertheless, the causal effect is relatively small compared with the large *associations* found both in this chapter and in previous papers. We find that the raw difference in the likelihood of a mental health shock between temporary and permanent workers is 40%, but the gap decreases to only 10% after controlling for a wide range of pre-illness job and worker characteristics. The difference in the prevalence of physical health problems even completely disappears after controlling for these observables. What this suggests is a role for selection – particularly on age – in explaining differences in the likelihood of experiencing a health shock as well as in the use of different contract types. Arguably, this type of selection may also explain the positive associations found in other papers. Our results are probably most comparable to those of Caroli & Godard (2016), who find a strong association between job insecurity and a wide range of health outcomes. When controlling for endogeneity with instrumental variables, however, only the prevalence of mental health problems appears to be affected.

Turning to the waiting period that precedes potential DI applications, we find that differences in employer incentives explain almost half of the gap in the DI application probabilities. This confirms earlier work on the effects of employer experience rating, which suggests reductions in DI inflows of 7 to 24% (Prinz & Ravesteijn, 2020; Koning, 2009, 2016; Hawkins & Simola, 2020). Our estimated effect of the 2013 reform is similar in magnitude, but it should be emphasized that in addition to experience rating, the formal monitoring obligations of employers were increased as well. Additionally, note that experience rating was already in place for workers with permanent contracts in 2013, whereas most other papers evaluate the introduction of experience rating in a context in which no experience rating exists.

Finally, our results shed new light on the concept of a 'work disability' to explain changes in DI inflow risks. In the literature, a work disability is commonly defined as the extra inflow into DI schemes induced by unfavorable business cycle conditions (Autor

& Duggan, 2003; Autor, 2011; Benítez-Silva et al., 2010). This presumes that changes in DI inflows are driven by economic conditions and not by changes in health conditions and that marginal applicants are predominantly low-productivity workers who are also more likely to flow into UI. Consistent with this interpretation, our analysis shows that DI application probabilities are higher for low-productivity workers, who are also more likely to have fixed-term contracts, than for permanent contract workers. However, health conditions are certainly not irrelevant. Specifically, the higher probability of applying for DI among fixed-term workers originates from the interaction between health conditions and economic conditions. Health conditions are more likely to lead to DI applications by fixed-term workers since they experience less employer commitment during their illness and have less favorable outside options.

3.6 Conclusion

The aim of this chapter is to explain and decompose the large difference in DI application probabilities between workers with fixed-term contracts and those with permanent contracts. Using rich Dutch administrative data from various sources, we show that the compositional differences between fixed-term and permanent workers cannot explain the observed differential. Additionally, the risk of falling ill is not substantially higher for fixed-term workers. We observe that most of the gap in the DI application probabilities arises after the onset of illness, and we show that the role of the employer during the waiting period is crucial. Increased employer responsibility for ill fixed-term workers in 2013 substantially reduced the difference in DI application probabilities. Finally, we find that opportunities in the labor market also matter for the decision to apply for DI. Conditional on illness, the probability of applying for DI increases among fixed-term workers if labor market prospects worsen. Among permanent workers, we find no such link. Jointly, these factors explain more than 80% of the difference in the probability of applying for DI between fixed-term and permanent workers.

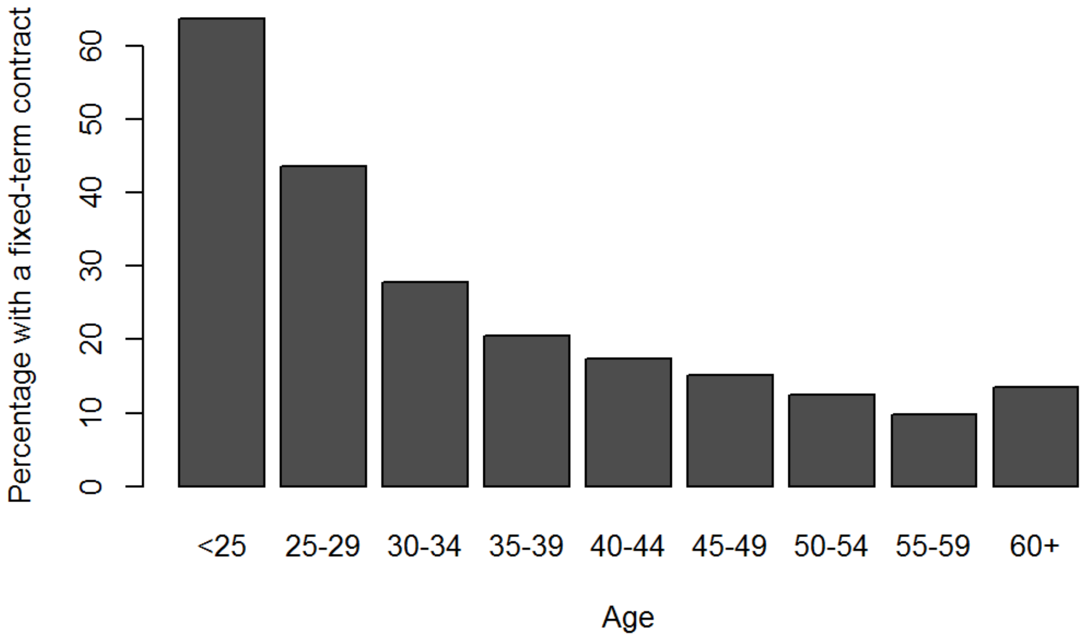
From a policy perspective, a key takeaway from our analysis is that the higher DI application probability observed among fixed-term workers emerges during the waiting period that precedes DI applications. In this period, employers play a crucial role in that they may or may not, e.g., implement work accommodations. Depending on the contract type and the corresponding employer incentives and obligations, different workers with similar health conditions face DI application probabilities that vary substantially. This provides a novel perspective on the concept of ‘work disabilities’: the economic context and corresponding contract settings do matter, but this is relevant only at the onset of the health problem. While this calls for sufficient commitment from employers of workers with fixed-term contracts, we are aware that the options for how to do so are more limited than for permanent workers. Increased obligations and incentives for employers may have a negative effect on overall employment and could trigger substitution toward the UI scheme and social assistance. Public employment offices could therefore play a more active role in supporting fixed-term workers during their waiting period. Our findings suggest that this is especially relevant in slack labor markets where job opportunities are limited.

3.A Appendix

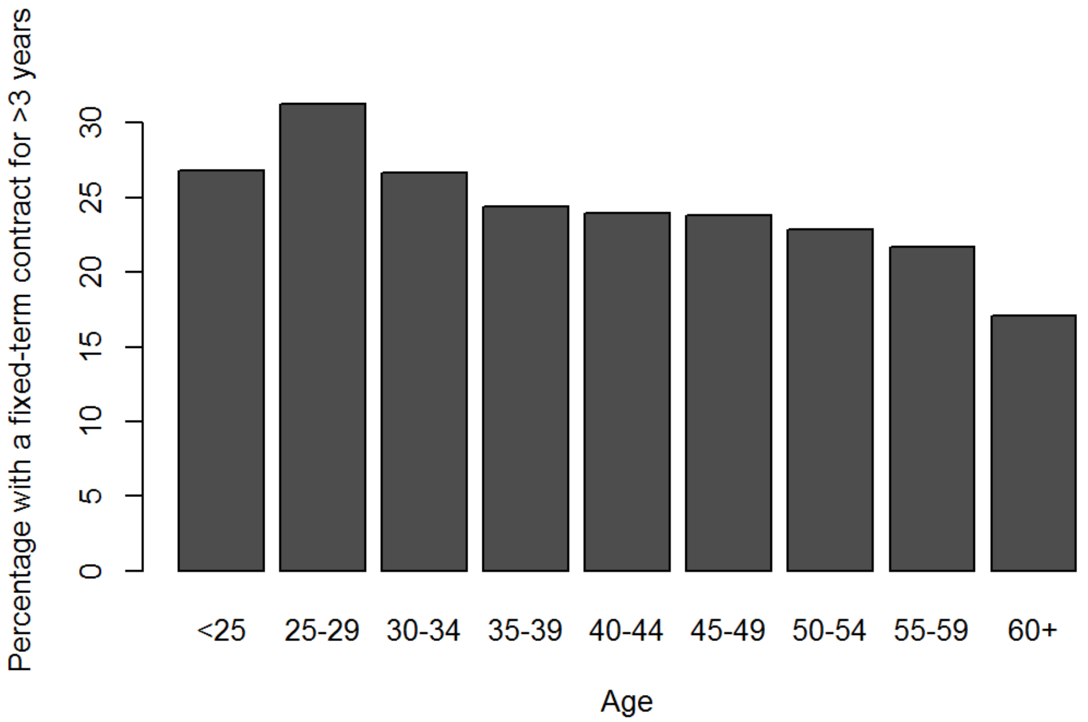
3.A.1 Distribution of fixed-term contracts over the age distribution

Figure 3.A.1: Distribution of fixed-term contracts over age categories

(a) Distribution of fixed-term contracts over age categories



(b) Share per age category of individuals with a fixed-term contract for more than three years, conditional on having a fixed-term contract once



3.A.2 Full set of control variables

Table 3.A.1: Full set of control variables for regressions

Control variable	Values
Quarter-of-the-year controls	22 quarter-of-year dummies
Demographic controls:	
Gender	Male or Female
Age	$\leq$ 24, five year age groups from 25-59, $\geq$ 60
Education	Education level split into 10 categories
Nationality	Dutch or Non-dutch
Family composition	Single / non-single and and 0/1/2/3+ kids
Population density of municipality	<500, 500–1000, 1000–1500, 1500–2500, >2500
Job controls:	
Monthly wage	$\leq$ 1000, 500 euro brackets up to 5000, $\geq$ 5000
Weekly number of working hours	10 hour brackets from 0–40, $\geq$ 40
Sector of employment	70 sector dummies
Health controls:	
Health cost last year	10 dummies based on cost deciles
Health cost current year	10 dummies based on cost deciles
Health cost next year	10 dummies based on cost deciles

3.A.3 Additional empirical results

Table 3.A.2: Healthcare cost coefficients DI application probability regression (Table 3.1, column (5))

Healthcare cost decile	Coefficient
10%–20%	0.0038
20%–30%	0.0075
30%–40%	0.0119
40%–50%	0.0165
50%–60%	0.0223
60%–70%	0.0285
70%–80%	0.0434
80%–90%	0.0686
90%–100%	0.1279
Missing	0.0153

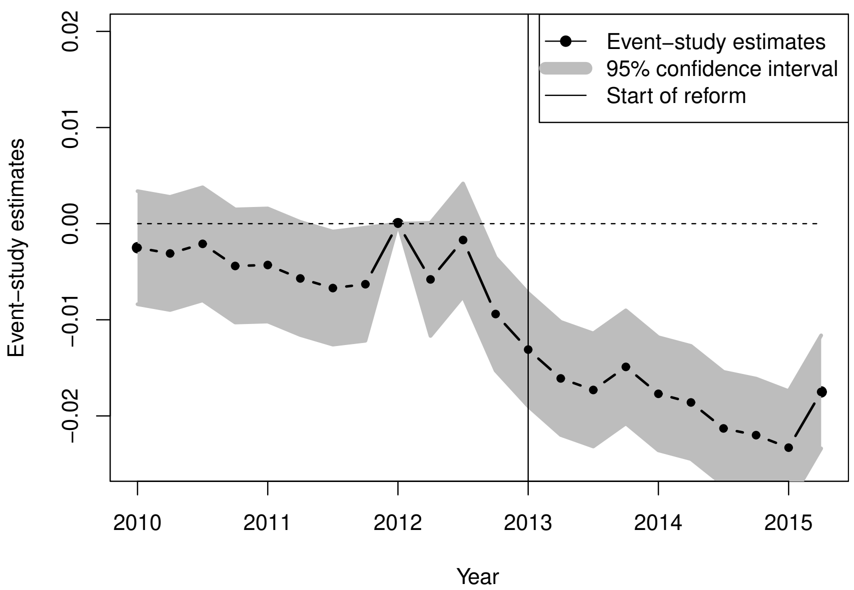
All coefficients in this table are statistically significant with P-values < 0.0001. The number of individual-year observations equals 475 million.

3.A.4 The 2013 reform: monitoring and experience rating

The disability reform of 2013 encompassed two major changes to the DI system. First, the reform increased monitoring and introduced an assessment after one year of illness for all workers on sick leave with a temporary contract. Second, the reform introduced experience rating for the same group of workers, making employers financially responsible for all their previous employees that have entered DI in the last two years – so also those employees no longer employed at the claims assessment. The impact of experience rating varies by firm size. Small firms with less than 10 employees pay a sector-level premium, whereas firms with more than 100 employees pay an individual premium that is fully based on their lagged DI inflow. Firms with 10 to 100 employees pay a weighted average of the sector-level premium and an individual premium (the individual weight increases from 0 to 100%). By exploiting the differential experience rating effects across observed firm size, we intend to disentangle the total effect of the reform into the effect of increased monitoring – which we assume equal across firm size – and the effect of experience rating.

Specifically, the effect of increased monitoring is estimated by comparing employees with permanent contracts to employees with temporary contracts at small firms.[34] To assess the validity of this approach, we first conduct an event-study that assesses the parallel trend prior to the reform and also to evaluate the dynamic reform effects. Figure

Figure 3.A.2: event-study analysis on the effect of increased monitoring

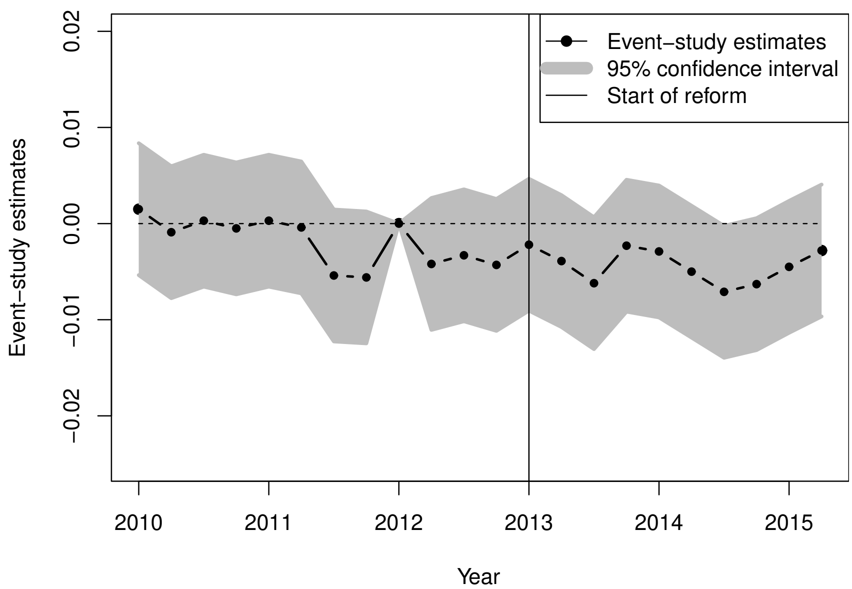


[34]To make the treatment and control group more comparable, we could also compare employees with permanent contracts at small firms to employees with temporary contracts at small firms. This yields very similar results.

3.A.2 shows the quarterly estimates in which the first quarter of 2012 is used as baseline. The general picture is comparable to the event-study performed in Section 3.4. That is, prior to the 2013 reform the gap in DI application probability is constant over time, lending credence to the parallel trends assumption. The gap decreases slightly in the last quarter of 2012, and continues to decrease in 2013 and 2014. Note also that the magnitude of the estimated DiD effects are very similar to the DiD effects estimated in Section 3.4. The effect of increased monitoring is approximately –0.014%-points, almost equal to the full effect of the reform (the full effect of the reform equals –0.016%-points).[35] This suggests limited effects of the introduction of experience rating for temporary workers.

To estimate the effect of experience rating, we compare employees with a temporary contract at small firms to employees with a temporary contract at large firms. For all of these employees, the reform increased monitoring and introduced the one-year assessment. However, experience rating was only introduced for employees at large firms and not for employees at small firms.[36] Figure 3.A.3 shows the quarterly event-study estimates. Once again, prior to the reform the parallel trends assumption seems to hold. However, also after the reform, the event-study estimates are small and insignificantly different from zero. This indicates that introducing experience rating, on top of increased monitoring, has a limited effect. When using a single post 2013 dummy, we find a borderline significant effect of –0.002%. This implies that experience rating explains at most 10% of the total

Figure 3.A.3: event-study analysis on the effect of experience rating



[35]The total effect is obtained by estimating a standard difference-ind-difference regression; we interact employment status with a post-2013 dummy.

[36]Note that we estimate the effect of introducing experience rating conditional on increased monitoring.

Table 3.A.3: Regression results for DI application not conditional on a negative health shock and stratified by tight and loose labor market sectors

	Pre 2013		Post 2013	
Labor market tightness[a]	Loose	Tight	Loose	Tight
Estimated DI application prob.: Fixed-term[b]	0.076	0.063	0.056	0.042
Estimated DI application prob.: Permanent[b]	0.043	0.042	0.041	0.040
Relative risk[c]	1.77	1.50	1.36	1.06
Number of observations	231 million	29 million	191 million	24 million

Regressions contain the same control variables as the regression of Table 3.1, column (4): quarter-of-year dummies, demographic and job controls. See Appendix Table 3.A.1 for specification of control variables. (a) Average predicted probability of applying for DI if all individuals would work in a sector with tight/loose labor market tightness before/after 2013; (b) Average predicted probability of applying for DI if all individuals would have a fixed-term contract (and similar for permanent contracts); (c) Ratio of estimated probability fixed-term and estimated probability permanent.

effect of the reform.

Contrasting to most of the literature, our results point at small effects of experience rating. One potential explanation for this is that the reform introduced experience rating for employees with temporary contracts at large firms only, while experience rating was already in place for employees with permanent contracts at these firms. These large firms might already have implemented return-to-work activities without discriminating between employees with permanent and temporary contracts. In addition, disentangling the impacts of the elements of the reform requires the additional assumption to hold that the effect of increased monitoring and increased financial incentives are independent. This assumption is necessary to draw the conclusion that monitoring *without* experience rating would have been almost equally effective. Given these limitations, we focus on the aggregate impact of the reform in the main analysis.

3.A.5 The impact of employer incentives and labor market prospects on the full sample

In the theoretical decomposition of the difference in DI application probability between fixed-term and permanent workers, we make the equal relative risk assumption. This assumption implies that the relative risk of a health shock, multiplied by the conditional relative DI application probability, equals the relative risk of no health shock multiplied by the relative DI application probability conditional on no shock. Given that our estimates of the relative risk of a health shock are approximately one, this assumption implies that the relative DI application probability conditional on a health shock equals the relative DI application probability conditional on no health shock. For completeness, Table 3.A.3

shows the estimated relative risks for the full sample. This sample thus contains both individuals with and without a health shock. Given that the majority of the sample is not hit by a health shock, the estimated probabilities of applying for DI are lower. However, the obtained relative risks in the various scenario's are very similar to the ones which we obtain when conditioning on health shock and the equal relative risk assumption is thus likely to hold. This also implies that the impact of employer incentives and labor market prospects are similar for individuals who are hit by a health shock, and individuals who are not hit by a health shock.

Chapter **4**

Sick or unemployed? Examining transitions into sickness insurance at unemployment benefit exhaustion.

4.1 Introduction

It is well-known that exit rates from unemployment insurance (UI) benefits increase considerably at benefit exhaustion. To explain the existence of these spikes, workers' moral hazard in search behavior, (hyperbolic) discounting, and the existence of storable job offers have been put forward (Marinescu & Skandalis, 2021; Card et al., 2007; Boone & Van Ours, 2012; Kyyra et al., 2019; Paserman, 2008). Since the work of Card et al. (2007), the research on spikes has also been broadened to workers that leave the unemployment registers without work at the end of UI benefit entitlement. This changes the perspective on the welfare consequences of changes in UI benefit length. Reductions in the benefit entitlement period may reduce moral hazard in job search but also discourage workers to register as unemployed and stay part of the labor force (Howell & Azizoglu, 2011).

The existence of spikes with out-of-the-labor-market exits raises the deeper question of whether this concerns workers with a high value of leisure or workers unable to work, typically due to health conditions. If the spike consists of workers with a high value of leisure, this points to moral hazard effects. However, if the spike consists of workers with serious health conditions, this provides an argument for the option to apply for adjoining sickness or disability benefits. In many developed countries such concurrent schemes exist for unemployed workers, but the boundaries between them are not always clear-cut and provide room for benefit substitution (Henningsen, 2008).

[1]This chapter is based on Koning & Prudon (2023).

Using large-scale administrative data from the Netherlands, this chapter documents a spike at benefit exhaustion consisting of workers that start receiving sickness insurance (SI) benefits that in turn may eventually lead to the receipt of disability (DI) benefits. We aim to investigate what drives the spike in exits into SI. To do so, we assess which types of workers apply for SI benefits at the spike, compared to workers applying for SI in earlier months. In particular, we try to distinguish between workers with mild or severe health conditions who are likely to be unable to work due to their illness, and workers with no or very limited health conditions who apply to extend the duration of benefit receipt. We refer to these two scenarios as a catch-up of initial non-take-up or that of moral hazard.

In our setting, there are several explanations of why both of these types of workers might apply for SI at the end of their UI entitlement. For workers with mild or severe health conditions, the spike can be explained by behavioral failures, such as status-quo bias and lack of information. Status-quo bias implies that workers have a tendency to remain on their status-quo option, UI in this case, instead of actively and immediately switching to SI. Once UI expires, these workers no longer have the option of remaining in their status-quo option and hence call in sick. Additionally, workers may initially simply be unaware of the option to apply for SI benefits. Knowing that they receive letters at the end of UI benefit receipt informing them of the entitlement conditions for SI, there may be a push towards increased exits into this scheme at that time. If these workers have health conditions which limit or prohibit them from working, they should be on SI. Hence, these scenarios refer to a case where the spike in inflow into SI is actually a catch-up of the initial non-take-up of SI.

The other type of workers, those with limited or no health conditions, may choose to enter SI to increase their social security wealth from UI and SI benefits. Due to a phase-in period, which will be discussed in detail in Section 4.2, timing an SI application at the month of UI benefit exhaustion maximizes the total duration of benefit receipt. This gives an explanation for the spike of applications that presupposes fully rational behavior of workers with limited health conditions, as well as the assumed ability to time applications during UI entitlement. Concurrent with this explanation, workers with no or mild health conditions might also apply for SI because of the same behavioral biases as workers with more severe health conditions; status quo bias and lack of information. There is also a third potential behavioral explanation for why these workers apply once their UI entitlement ends. Present-biased workers will procrastinate their search effort, which renders them more likely to reach the moment of benefit exhaustion (Paserman, 2008). At this point, they may find out that they cannot find a job, and apply for SI benefits to compensate for the loss of UI benefits. If the health conditions of these workers are not severe enough to limit them from working, they should not be on SI and their inflow can thus be characterized as moral hazard.

With various explanations for the spike pointing at specific worker groups as 'marginal' respondents, the impact on the targeting of SI benefits is largely an empirical matter. If the spike mainly stems from moral hazard effects, the marginal SI applicants should on average have less severe health conditions. Concurrently, we would expect that these workers have a higher likelihood of being screened out of SI during the two main external screening moments. Alternatively, if the spike mainly stems from workers who should have entered SI already prior to the end of their UI entitlement, we would expect them to be similar in terms of health conditions, compared to workers who entered SI earlier. As a result, external screening moments should also similarly impact them.

Throughout our analysis, we estimate flexible regression models on a sample of approximately one million UI spells which started between 2013 and 2015. The models include effects for the elapsed time on UI – representing genuine duration effects and selection effects due to unobservables –, calendar time, and the residual time left until UI benefit exhaustion. Given that the UI benefit length depends on the employment history of individual workers, we separate the effect of reaching the end of the UI entitlement period under the assumption that workers have common elapsed time effects. We first estimate regressions for the monthly exit rate from UI into sickness benefits, allowing us to infer the relative size of the spike compared to the inflow into SI before benefit exhaustion. To investigate the persistence of the spike, we also estimate models for the incidence of continued receipt of sickness benefits over longer stretches of time – and ultimately leading to the receipt of disability insurance (DI) benefits. As a second step, we characterize and compare workers applying for sickness benefits just before and at the spike of UI benefit exhaustion. The comparisons are made using detailed information on (mental) health, demographics, and labor market outcomes prior to unemployment.

We find a spike in exits out of UI into SI: UI recipients are 30% more likely to apply for SI benefits when they reach the month of UI benefit exhaustion, compared to prior months. In relative terms, this spike is larger than the spike for exits into employment (about 10%). In terms of the continued receipt of SI benefits, the relative risk increases to about 60% and returns to about 40% after (medical) assessments by the social benefits administration that occur after about 12 months of sickness benefit receipt. The initial spike of the inflow in SI benefits is not mirrored by any differences in the health outcomes, suggesting that delayed applications of workers with similar health conditions are driving the spike at benefit exhaustion. Differences in the exit rates out of SI during the first four months of sickness result in less severe health conditions for the remaining spike cohort still receiving SI on average.

The spike effect for exits into employment we find is comparable to e.g. Card et al. (2007) for Austria and Roed & Westlie (2012) for Norway. The spikes of exits out of the labor force found by Card et al. (2007) and Kyyra et al. (2019) are larger than the spike into sickness benefits we find, but this is probably due to the eligibility constraints for SI that are relevant in our context. We expand on the existing literature by explicitly examining the (long-term) targeting effects of benefit schemes. Our study comes closest to Roed & Westlie (2012) and Henningsen (2008), who analyze the importance of incentives for benefit shifting in Norway and Denmark at the moment of benefit exhaustion.

Our findings also add to a considerably broader literature on screening and targeting of social insurance. Theoretical work in this field addresses the implications of classification errors arising from imperfect information about the individual's true disability status (Parsons, 1991; Kleven & Kopczuk, 2011). Empirical studies typically look at the impact of changes in the rigor of the screening process on DI applications and inflows, but usually with limited attention to targeting (Autor & Duggan, 2014; Autor et al., 2015; Markussen et al., 2017; Liebert, 2019; Haller et al., 2020).[2] More recently, Deshpande & Li (2019) and Godard et al. (2022) study the targeting effects due to self-screening and screening into DI. Both find deterrence effects of increased application costs, which predominantly impacted potential applicants with less severe conditions. Finally, Deshpande & Lockwood (2022) investigate the welfare effects of DI benefit receipt for workers with less-severe conditions but weak labor market positions. They argue that DI benefits are desirable for this group, even though their health conditions are relatively mild. We add to this literature by examining how both the health conditions and labor market position evolve of those who remain on SI.

This chapter proceeds as follows. The institutional setting is described in Section 4.2. Section 4.3 presents a theoretical framework for the decision to delay the SI application. Section 4.4 discusses the administrative data and provides the first descriptive evidence of spikes in exit rates. Sections 4.5 and 4.6 lay out the empirical strategy and present the findings. Finally, Section 4.7 concludes.

4.2 Institutional setting

We study unemployment insurance (UI) and sickness/disability insurance (SI/DI) in the Netherlands focusing on the time period 2013-2015. Below we provide a brief overview of the eligibility criteria and generosity of these systems. UI is mandatory for all workers in the Netherlands. The Dutch Employee Insurance Agency (UWV) is responsible for

[2] Related to this work, various studies show evidence on the presence of substitution effects between UI and sickness or DI schemes, see e.g. Koning & Van Vuuren (2010), Autor & Duggan (2003), Borghans et al. (2014) Hofmann (2014) and Berg et al. (2019). Again, these studies underline the importance of benefit shifting, but not specifically at the moment of benefit exhaustion.

collecting UI premiums and benefit payments, together with case management for ongoing eligibility conditions during UI receipt. UI applicants need to have worked at least 26 weeks of the previous 36 weeks and at least four of the last five calendar years. Below a certain earnings cap, UI benefits are equal to 75% of the pre-UI wage earnings in the first three months and 70% of the pre-UI wage earnings in the subsequent months of benefit receipt. The maximum benefit entitlement period is determined by the employment history of workers: each additional year of employment increases the UI entitlement period by one month, with a cap of 24 months of benefit receipt at maximum.[3]

When UI benefits are exhausted, workers can apply for adjoining welfare benefits that are administered by local municipality offices. Benefits amount to 50% of the statutory minimum wage for individuals in households with two welfare recipients and 70% of the statutory minimum wage in single households. Eligibility for welfare benefits is only for workers with no or insufficient income from partners and without substantial financial assets. In effect, UI benefit exhaustion may imply a substantial drop in income.

In case of sickness, workers with UI benefits can apply for sickness insurance (SI) benefits that replace 70% of the pre-application earnings. Eligibility to the SI scheme is limited to unemployed workers. Workers who fall sick while being employed (either temporarily or permanently) are entitled to continued wage payments during the sickness period. Sick workers with a temporary contract are entitled to SI benefits once their contract expires. For both unemployed and temporary workers – known as so-called “safety netters” – the Employee Insurance Agency is responsible for the administration of premiums and benefits. Public SI benefits are not relevant for workers with permanent contracts for whom the employers’ responsibilities cover the full two-year waiting period that precedes applications to public Disability Insurance (DI) benefits (Koning & Lindeboom, 2015).

SI applications are initiated by individual workers, who are informed about their eligibility to SI in letters at the start of the UI entitlement and two months before the moment of benefit exhaustion. Workers call in to apply for SI to the front offices of the Dutch Employee Insurance Agency, whereas case managers from back offices call these workers after a few days[4] and register the conditions of the workers. In principle, there is no doctor’s certificate needed for an application at that stage. SI applications are awarded in almost all cases, but case managers may forward cases to medical assessors if they doubt the plausibility and severity of illness. Workers may then be summoned for a meeting with the case manager. Depending on the assessed severity and durability of conditions, another phone call is planned at the moment of expected recovery or at the

[3]To determine the employment history of workers, elapsed employment histories up to January 1*st* of 1998 are set equal to the age at that moment minus 18.

[4]If the case manager is not able to get into contact with the worker, the application process is frozen for the time being.

end of the first month of illness at the latest. From then on, continued cases are taken over by a team consisting of the case manager, a medical assessor, and a vocational expert.

Workers reporting sick are subject to a 'phase-in' period where there is continued receipt of UI benefits. The duration of the phase-in period is equal to the remaining UI entitlement duration, with a maximum of 13 weeks. After the phase-in period, workers start receiving SI benefits. Workers can apply for disability insurance (DI) benefits two years after reporting sick, regardless of the duration of the phase-in period. The remaining UI benefit entitlement stays valid after SI or DI benefit receipt ends.

Due to the phase-in setup, there can be a financial advantage for workers to delay SI applications up until the end of UI entitlement. As long as workers face a positive probability of not being awarded full DI benefits and if their (expected) duration of SI benefit receipt exceeds their remaining UI entitlement period, a one-month delay in SI applications implies a one-month increase in the total period of benefit receipt. With a phase-in period of 13 weeks, workers may gain 13 weeks of benefit payments at most by applying for SI benefits in the very last week of UI entitlement. So as long as workers have the ability to time applications, an optimal strategy for workers would therefore be to delay applications. Obviously, this may be more difficult for workers with severe health conditions who can no longer meet the ongoing entitlement conditions for UI.

The first intensive screening by the Employee Insurance Agency takes place about one year after the start of SI benefit receipt; this is the so-called 'first-year-assessment'. This assessment is similar to the screening for DI claims that is conducted two years after the onset of illness: medical assessors determine the conditions and limitations of workers and vocational experts determine the "most-earning jobs" that are still feasible so as to estimate the workers' residual earnings capacity. For the first-year assessment, only workers with an assessed loss of at least 35% of their old earnings capacity continue receiving SI benefits; those whose loss in earnings capacity is deemed to be smaller may apply for any remaining entitlement periods in UI or adjoining welfare benefits (if relevant). After the DI assessment after two years of SI benefit receipt, SI benefit receipt will end and the worker will either receive full DI benefits (when the loss of earnings capacity exceeds 80%), partial benefits (when the loss of earnings capacity is between 35% and 80%) or no DI benefits at all (when the loss of earnings capacity is less than 35%).

4.3 Theoretical framework

To better understand the various mechanisms potentially driving the increased inflow into SI at the end of UI entitlement, we propose a simple theoretical framework in which an unemployed worker can choose to apply for SI during the UI benefit period. The following subsections discuss the setup of the model and its predictions.

4.3.1 Model setup

In each period (month), an unemployed worker with some level of health h decides whether to apply for sickness insurance (SI) or whether to remain on UI. When entering SI, workers remain on SI and potentially also on DI for a combined total of n months. If they are still eligible for UI after these n months, the choice problem is repeated. Once the elapsed time on UI, t, exceeds the total UI entitlement, T, workers can no longer choose to be on UI or SI but will receive social assistance (SA). For most workers, the benefit income from SA is lower than for UI and SI (and can amount to zero with partner income or with assets). To simplify the model, we assume that exits into employment are independent of the decision whether to stay on UI or apply for SI. Exits into employment are therefore not incorporated into the stylized model. Hence, the money utility in period t of respective valuation options can be written as:

$$U_t \;=\; \begin{cases} \sum_{j=1}^{n} \gamma^{j-1}\, b^{SI}(h) \;+\; \gamma^n\, U_{t+n} & \text{if on SI and } t \leq T \\ b^{UI}(h) \;+\; \gamma\, U_{t+1} & \text{if not on SI and } t \leq T \\ b^{SA} + \gamma\, U_{t+1} & \text{if } t > T, \end{cases}$$

with SI as an indicator for receiving SI, b^{SI} the benefit value of SI and potentially DI, b^{UI} the benefit of UI and b^{SA} the value of social assistance. The subsequent period's utility, U_{t+1}, is discounted by a factor γ. The total value of SI and DI benefits, $\sum_{j=1}^{n} \gamma^{j-1} b^{SI}$, depends on the level of health h. Workers who are in worse health have a lower likelihood of being screened out during the two-year waiting period and a higher probability of being granted DI benefits after two years of sickness ($\frac{\partial n}{\partial h} < 0$). Consequently, the total valuation of SI and DI receipt decreases as health increases.

If a worker decides to remain on UI, he/she has to adhere to ongoing UI job search requirements and has meetings with caseworkers. The value of unemployment benefits in one period therefore consists of both the (monetary) benefits (mb^{UI}) and the money disutility of the ongoing job search requirements, $\phi(h)$:

$$b^{UI}(h) \;=\; mb^{UI} - \phi(h)$$

Adhering to the ongoing job search requirements is more difficult for workers who are in poor health. Hence, the instantaneous money utility value of UI increases as health improves ($\frac{\partial b^{UI}}{\partial h} > 0$). In our model framework, we assume that the total duration of SI benefit receipt is unaffected by delaying the SI application. The implications of this are discussed in Appendix Section 4.A. Additionally, we assume that the expected duration of SI receipt, n, exceeds the remaining UI entitlement. If this is not the case, workers do not gain from delaying their SI application.

4.3.2 Model predictions

In our model, workers face the decision to apply for SI and the choice to delay their SI application to a later month.[5] When not taking into account the option to delay the SI application, a worker will compare the money utility of entering SI to the utility of remaining on UI when deciding whether to apply for SI. This yields the following comparison:

$$\underbrace{\sum_{j=1}^{n} \gamma^{j-1}\, b^{SI}(h)}_{\text{Net present utility value of SI}} \quad > \quad \underbrace{\sum_{j=1}^{T-t} \gamma^{j-1}\, b^{UI}(h) + \sum_{j=T-t+1}^{n} \gamma^{j-1}\, b^{SA}}_{\text{Net present utility value of UI and SA}} \tag{4.1}$$

A worker chooses to apply for SI if the net present money utility value of doing so exceeds the net present money utility value of receiving UI and SA benefits for the same duration. As health deteriorates, the left-hand side of this equation increases due to a longer expected duration of SI benefit receipt, while the right-hand side decreases due to the increased cost of adhering to the UI job search requirements. The likelihood to apply for SI thus increases as health deteriorates.

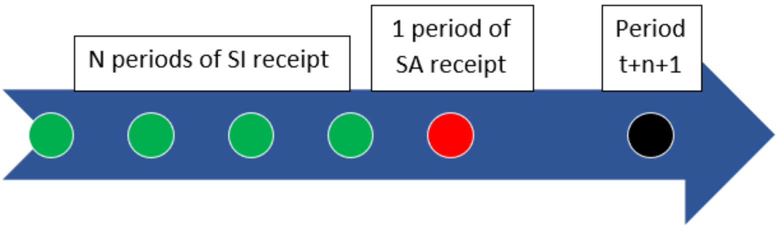
If it is optimal for a worker to apply for SI, he/she chooses to apply immediately or delay the application. This decision problem depends on the duration of the remaining UI entitlement period. The incentives to delay are strongest for workers who have less than 13 weeks of UI entitlement remaining, as delay translates one-to-one into increased UI benefit duration in this case. We illustrate the incentives in the case of 13 or fewer weeks of remaining UI entitlement in this section. Appendix Section 4.A also illustrates the incentives to delay in case the remaining UI entitlement period exceeds 13 weeks. Figure 4.1 shows how a one-month delay affects the outcome for a worker who has less than 13 weeks of remaining UI entitlement, and for whom the expected duration of SI receipt exceeds their remaining UI entitlement.

Figure 4.1 shows that delaying the SI application by one month implies that workers gain one month of UI benefit receipt. This comes at the cost of receiving SI benefits one

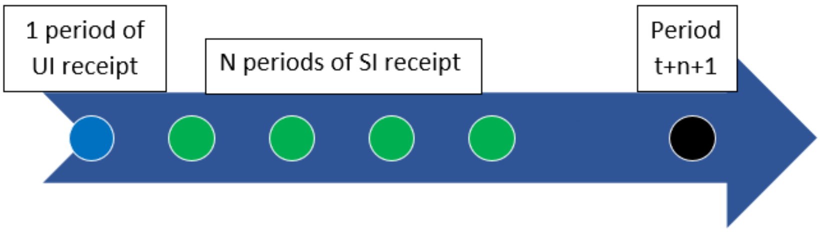
[5] See Appendix Section 4.A for a detailed analysis of the theoretical model.

Figure 4.1: Choice problem for delaying SI application by one month

(a) Outcome without delay



(b) Outcome with a one-month delay



month later and losing one month of SA receipt. Algebraically, this comparison of option values is shown below:

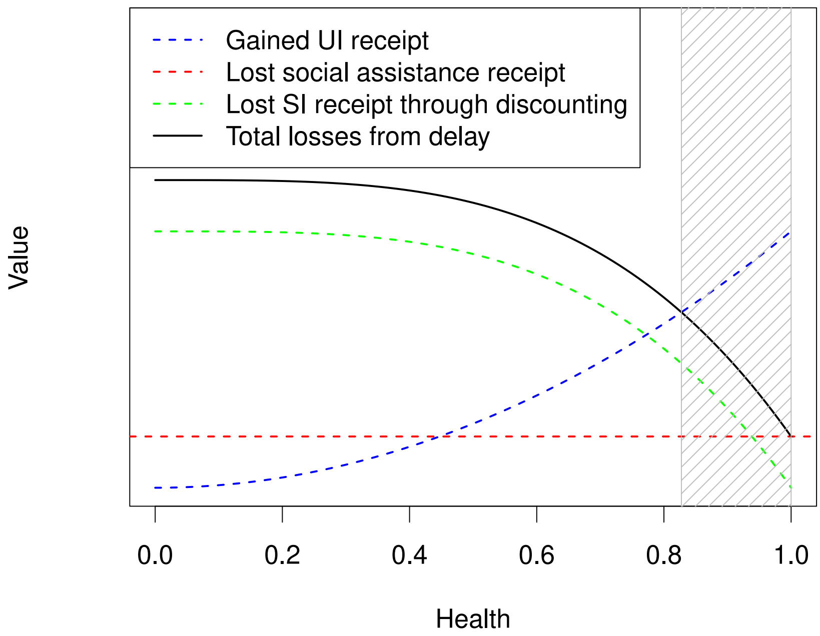
$$\underbrace{b^{UI}(h)}_{\text{Gained UI utility}} > \underbrace{(1-\gamma)\sum_{j=1}^{n}\gamma^{j-1}\, b^{SI}(h)}_{\text{Lost valuation of SI}} + \underbrace{\gamma^n\, b^{SA}}_{\text{Lost SA receipt}} \tag{4.2}$$

The worker will thus choose to delay the UI application if one additional month of UI receipt exceeds the loss incurred through discounting by receiving SI benefits one month later and one month less of SA after SI benefits are terminated. Workers who are not yet within the last 13 weeks of their UI entitlement have smaller incentives to delay their SI application. The maximum benefit they can gain remains the same (13 weeks), but to attain this benefit they need to delay for more months and hence incur the costs of delay for more periods. See Appendix Section 4.A for an illustration.

4.3.3 Comparative statics

The optimal outcome of the theoretical model depends on a worker's health, h, their remaining UI entitlement period, $T-t$, and his/her individual valuation of UI receipt, SI receipt and social assistance. To start with, Figure 4.2 shows the impact of health on the decision to delay the application.

Figure 4.2: Graphical illustration of the optimality condition for delay of SI



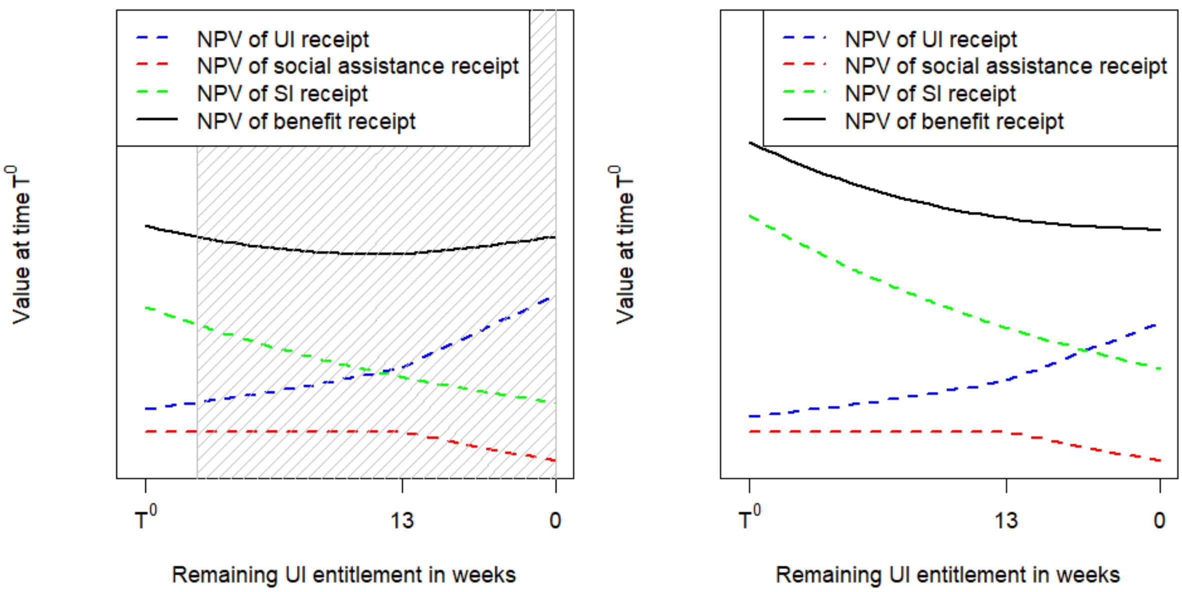
The figure shows that the gained valuation from one additional month of UI receipt increases as health improves. Adhering to the job search requirements is less burdensome for workers in good health. Concurrent with this, healthy workers will on average receive SI for fewer months and therefore have a lower valuation of SI benefit receipt. The loss of delaying this receipt is therefore lower with better health, so delaying is more likely to occur for this reason as well.[6] So the gains from delay increase as health improves, while the costs decrease. Delay is therefore only optimal if the health level is sufficiently good, which is illustrated in the figure by the support with the grey-shaded area. Given the potential gains and losses of delaying an SI application, it might be optimal for workers to delay their application for multiple months, or to apply immediately. The optimal application month is that in which the net present value of the sum of all benefits (UI, SI and social assistance) is maximized. To illustrate the various forces at work in the model, Figure 4.3 shows how the total net present value of benefit receipt (in black) is affected by delay for workers who are in good health (left panel) or in bad health (right panel).

For workers who are in good health, applying for SI benefits is only optimal if their value of SI benefit is relatively high even though their expected duration of SI receipt is relatively short. This could for example be the case for workers with a high disutility from working. For these workers that do decide to apply for SI despite their limited health problems, delaying their application is likely to be optimal. As shown in the left panel of Figure 4.3, the cost of delaying SI receipt is moderate in this case. At the same time,

[6]Note that the valuation of SA is independent of the level of health.

Figure 4.3: Net present money utility value of benefit receipt as a function of the month of SI application

(a) Good health (b) Bad health



adhering to the job search requirements is not costly for these workers, given that they have no health problems. These workers therefore value UI benefits highly and gain a lot from receiving UI for a longer period of time. In the example of the left panel of Figure 4.3 it becomes optimal for them to delay their SI application if they are within the last 26 weeks of their UI entitlement. Once the remaining UI entitlement period is less than 26 weeks, the maximum net present value of benefit receipt is attained by applying for SI in the last week of UI entitlement.

For a worker who is in bad health, the expected duration of SI and potential DI benefits is long, which results in a high valuation of SI receipt. The cost of receiving SI benefits later is therefore relatively high. At the same time, the job search requirements of UI are costly resulting in a low valuation of UI. The gains in UI receipt do not offset the losses of SI and SA. As illustrated in the right panel of Figure 4.3, the net present value of benefit receipt is always maximized by applying for SI benefits immediately, regardless of the remaining UI entitlement period. Any delay would result in a lower net present value. Note that in the extreme case in which the probability to be granted full DI benefits is one and the worker is expected to remain on DI until retirement, there are no incentives to delay the SI application at all.

As discussed in the introduction of this chapter, there might also be behavioral failures at play which result in a delay of the SI application. First of all, workers might be prone to status-quo bias. This implies that the cost of switching to SI is relatively high, preventing workers from applying for SI. At the moment of UI benefit exhaustion, however, all workers have to make a switch, either to SI or to SA. The increased cost due to the status-quo bias applies to both exit options, and hence effects will cancel out each other. This may therefore result in an increased inflow into SI in the last month of UI benefit entitlement. The second potential behavioral failure is present bias. The impact of present bias is ambiguous in our model. For workers with a high valuation of UI relative to SI, the gains of delay are in the present while the costs lie in the future. However, for workers with a relatively high valuation of SI, delay translates to costs in the present and costs in the future. Present bias might thus incentivize workers to delay their SI application, but it might also deter them from doing so. The last behavioral failure concerns a lack of information. If workers are unaware of their eligibility for SI, the choice set of these workers is limited. Once they receive the letter informing them of their eligibility, they re-optimize while including SI in their choice set, and might decide to apply.

4.4 Data

We use administrative data sources which are merged at the level of the individual. The first data set contains information from the Dutch Employee Insurance Agency on the UI receipt for all Dutch inhabitants between January 2013 and December 2015.[7] For each UI spell, we observe the exact starting and ending date, together with the maximum entitlement period. With this data set, we construct samples at risk of applying for SI benefits while on UI. The second data set contains all SI spells over the same time period, containing the starting day, the total duration of sickness, and the outcomes of the first-year and disability claim assessments (if relevant). Third, we use data from Statistics Netherlands to construct monthly time series on employment, working hours, labor income, and the receipt of various social assistance benefits. This allows us to track destinations and labor market outcomes of workers leaving UI or SI. Finally, we have annual information on healthcare costs and monthly information on the number of minutes of mental healthcare treatments. Given the mandatory setup and broad coverage of insurance, it contains information about almost all healthcare treatments for all Dutch citizens. For mental health, we have detailed data on the number of minutes patients are treated each day and the corresponding mental health diagnosis. For non-mental healthcare as well as for mental healthcare, the data contain the yearly expenditures on categories of healthcare. We use these categories to construct total expenditures on

[7] We also have UI records starting before January 2013, but there was no use of first-year assessments earlier than that date. To be able to interpret to our results consistently, we therefore excluded these observations.

mental healthcare and total expenditures on non-mental healthcare.[8]

We observe approximately 1.4 million UI spells. For each unemployed worker, we only use the first UI spell we observe in the data, resulting in a total sample of approximately one million workers. Given that we are interested in effects towards the end of the UI entitlement period, Table 4.1 shows descriptives of individuals with three and one month of remaining UI entitlement. The first two rows of Table 4.1 report exit probabilities into sickness insurance and into employment. The comparison of monthly exit rates into SI

Table 4.1: Descriptive statistics of the sample of workers reporting sick or starting employment three months and one month before UI benefit exhaustion

Sample of UI recipients		
Exit probabilities	3 months	1 month
To sickness	1.77%	2.40%**
To employment	8.92%	10.36%**
Total number of observations (ind. * month)	7,084,192	

Cohorts exiting UI	Sickness report		Employment	
	3 months	1 month	3 months	1 month
Healthcare utilization^d:				
Probability to receive mental healthcare treatment	7.13%	8.21%	2.52%	2.12%**
Number of mental health treatment minutes	11.8	18.0**	3.5	2.5**
Duration of treatment at end of UI in months	6.0	6.2	2.7	2.9
Residual duration of treatment at end of UI	10.5	11.3	5.5	5.2
Annual healthcare cost	3683	3329	1062	995
Demographics^d:				
Male	48.5%	47.4%	52.1%	52.9%
Age	33.4	34.1**	28.7	29.5**
Native	63.2%	57.6%**	67.2%	63.5%**
Low education level	23.1%	24.4%	13.7%	15.8%**
Middle education level	56.0%	53.1%	54.6%	53.0%**
High education level	17.8%	18.5%	27.1%	26.2%
Pre-UI employment outcomes:				
Monthly number of working hours	83.8	82.4	105.5	103.5**
Fixed-term contract	49.9%	48.5%	67.2%	64.1%**
Permanent contract	23.3%	23.2%	21.5%	22.8%**
Monthly labor earnings	1691	1761	2018	2042
Number of workers	3969	3911	14258	12521

Significant difference between cohort exiting three months and one month before maximum UI entitlement at: * a 10% significance level; ** a 5% significance level

[8] See Appendix Table 4.A.1 for categorizations of types of expenditures.

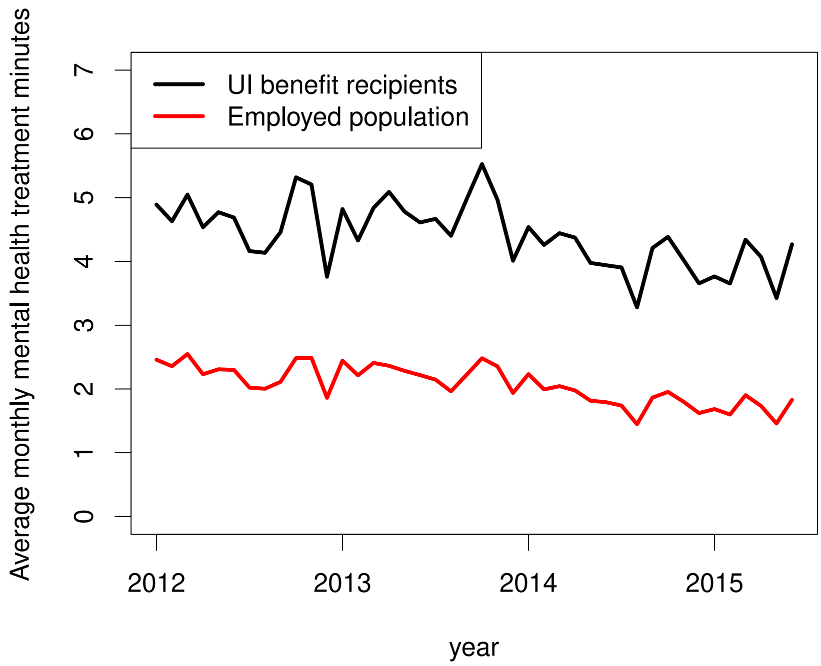
does suggest the presence of a spike at benefit exhaustion, with an increase from 1.77% to 2.40%. It should be noted that the baseline probability to enter SI, and, after the two-year sickness period, DI is already high compared to inflow rates of the employed population. The probability to apply for DI for someone on UI is approximately two percentage points each year, which is about five times the average of employed workers. This elevated level is comparable to the overall application rates in the 1980s in the Netherlands (Koning & Lindeboom, 2015). For exits into employment, we observe slightly higher exit rates at benefit exhaustion as well. In relative terms, this increase is however less pronounced. The increases in exit rates, both towards SI and employment, may reflect the impact of selection and underlines the need for a more formal analysis that decomposes such effects.

The bottom panel of Table 4.1 zooms into the groups which exit UI three and one months before their maximum UI entitlement. Columns (1) and (2) concerns the sample of individuals exiting UI into SI, while columns (3) and (4) consider the sample of individuals exiting UI into employment. It is not surprising to see that the strongest differences in averages follow from the comparison of those exiting into SI and into employment. The use of healthcare is approximately three times as high among those who report being sick. The group is also older, less likely to be native, and lower-educated.

Comparing the cohorts exiting UI three and one month before their maximum entitlement, we find that workers reporting sick in their last month of UI entitlement use more mental healthcare, but spend less on healthcare in total. Differences in terms of demographics and pre-UI employment outcomes are limited between the cohorts entering SI. The only significant differences are the average age and the fraction with a migration background. Differences between the two cohorts flowing into employment are more pronounced. While healthcare utilization is similar for the cohorts, workers finding a job in the last month of their UI entitlement tend to be older and have lower education levels. In accordance with this, they are less likely to have a fixed-term contract and more likely to have a permanent contract prior to entering UI. Additionally, they worked fewer hours per month but had similar monthly labor earnings.

While there are significant differences in healthcare utilization between various subgroups of UI recipients, it is important to note that healthcare utilization of the full sample of UI recipients is at a considerably higher level than for the employed population. To illustrate this, Figure 4.1 shows the average number of mental healthcare treatment minutes received by workers on UI and workers with an employment contract. The large difference in treatment minutes is fully driven by the extensive margin as recipients of UI are almost twice as likely to receive mental healthcare treatment, as compared to employed workers. This provides suggestive evidence that a subsample of UI recipients would actually be eligible for SI benefits, but does not apply for these benefits. One potential explanation

Figure 4.1: Mental health treatment minutes received by workers on UI, and workers with an employment contract between 2012 and June 2015

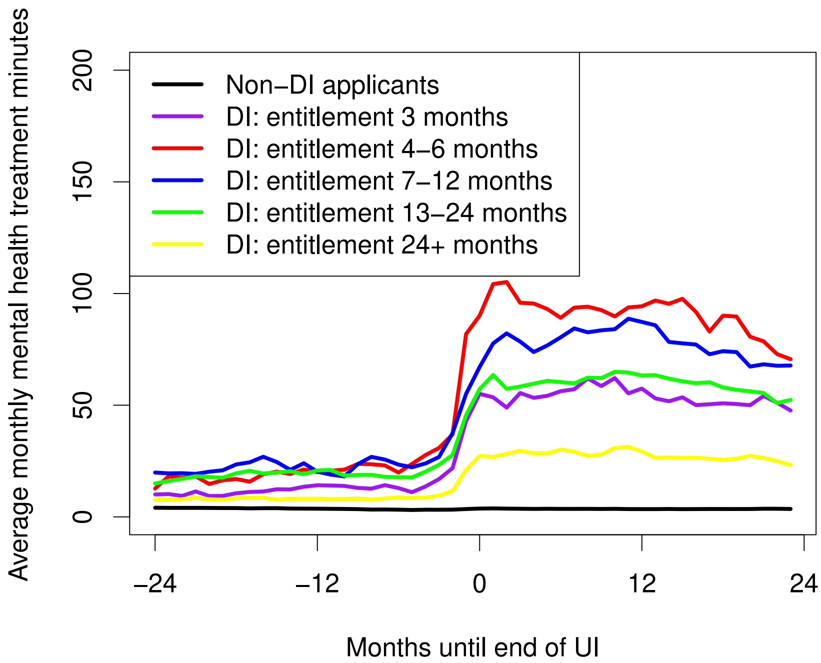


for the presence of a spike into SI at UI benefit exhaustion may therefore be that there is a catch-up effect of the initial non-take-up of SI.

In our analyses, it is key that we use health measures that proxy the risk of flowing into SI and DI. To validate this, Figure 4.2 shows the average number of mental health treatment minutes received by unemployed workers that eventually apply for disability benefits, together with those who do not apply. Averages of mental health treatment minutes are centered around the end of the UI spell. Workers not applying for DI benefits are rarely treated and there is no increase in treatment around the moment UI ends. For workers that eventually apply for DI benefits, however, treatment of mental health problems is much more likely. Specifically, the prevalence and intensity of treatment increase strongly in the months prior to the end of UI. This indicates that many of these workers are experiencing mental health issues in the months prior to reporting sick. The figure also shows marked level differences with respect to the maximum period of UI entitlement. Groups with longer entitlement tend to use less mental healthcare. This reflects an age effect: groups with longer UI entitlement are older, while younger groups are more likely to suffer from mental health problems.

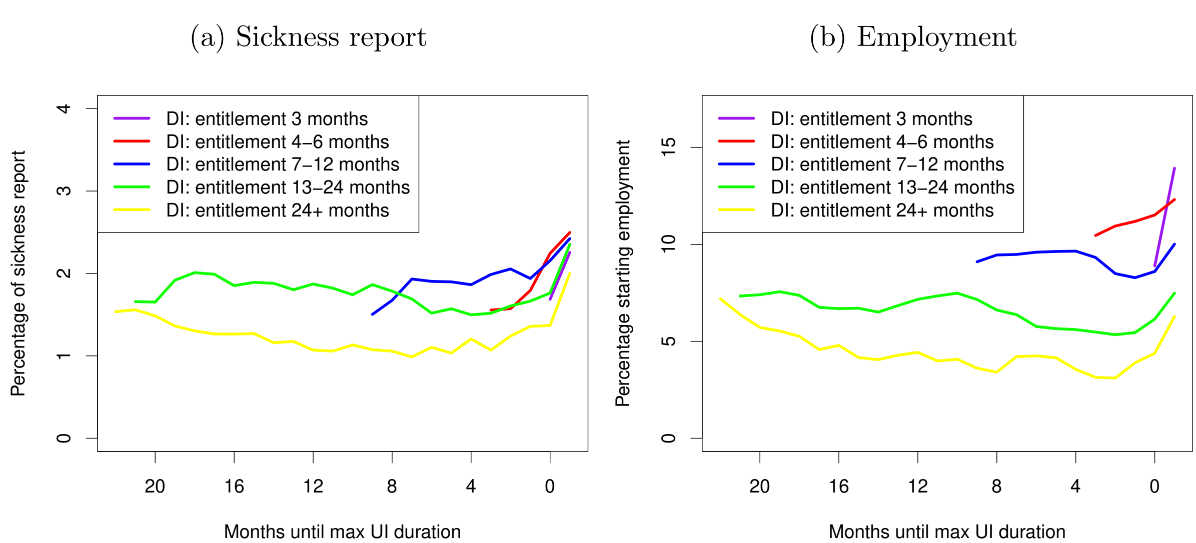
To eyeball the potential presence of spikes at UI benefit exhaustion, Figure 4.3 shows monthly exit probabilities out of UI into SI or into employment. Exit probabilities are calculated separately for groups that are stratified on their maximum UI duration. As the left-hand panel shows, there is a spike in the SI risk towards the maximum UI du-

Figure 4.2: Evolution of mental health treatment relative to the end of UI ("t=0"), stratified w.r.t. the maximum DI entitlement period



ration. The probability of entering into SI is considerably higher in the last month of UI entitlement. The maximum entitlement groups differ in the level of their baseline risk of reporting ill, but all show a spike in the last month of UI entitlement. Given this strong increase in the last month, we define the spike cohort as those flowing into SI in the last month of UI entitlement. We define the pre-spike cohort as those flowing into SI two months earlier. The right-hand panel shows the monthly probabilities to exit

Figure 4.3: Probability exit into SI (left) or to start employment (right) for workers with UI benefits relative to their maximum UI duration ("t=0")



into employment. In contrast to the risk into SI, the spikes at maximum UI entitlement into employment are much less pronounced. For all groups, there is a small increase in the probability to start employment in the last months of UI entitlement, but in relative terms, the increase is small compared to the spike in sickness reporting.

4.5 Methodology

To understand what drives the spike in SI applications at UI benefit exhaustion, we follow a three-step approach. Similar to the existing literature, we first estimate the absolute size of spikes from UI, both into SI and into employment. In this context, "marginal" SI recipients are defined as the extra influx of UI recipients that enter at benefit exhaustion. A novel aspect of our approach is that we investigate the persistence of the spike during the period of SI benefit receipt, ultimately up to and including the moment of DI application after 24 months of SI benefits. These analyses allow us to study the possibility that the spike fades out, both in substance and for specific groups that are initially most responsive to benefit exhaustion. We therefore redefine the outcome as entering SI *a*nd receiving SI benefits for a certain number of months, up to 24 months at maximum.

Second, we examine compositional differences between cohorts leaving UI for SI benefits just before and at the spike. We label these groups as pre-spike and spike cohorts. Assuming monotonicity at the spike, this also provides insight into the population of marginal SI recipients. Finally, we examine the impact of external screening – both for the one-year-assessment and for the application for DI benefits – on cohorts entering SI just before and at the moment of UI benefit exhaustion.

In the first step, we examine the likelihood of reporting sick close to the maximum UI duration. We estimate the probability to report sick in month s since the start of UI receipt, conditional on the elapsed time on UI. We denote t as the observed duration on UI. Conditional on UI receipt ($s \leq t$), the linear probability model for reporting sick in month s of UI benefit receipt is specified as:

$$SI_{i,s,\tau} = \alpha + \gamma_s + \delta_\tau + \psi_{T_i} + \phi_{T_i - s} + \varepsilon_{is} \qquad \text{for} \quad s \leq t\,, \tag{4.1}$$

with s as the elapsed duration of UI receipt, τ as calendar time and T_i as the maximum period of UI benefit entitlement for worker i ($i = 1, ..N$). For a worker who calls in for SI benefits the second month of the UI duration, $SI_{i2,\tau}$ equals one. The parameter α is the intercept, whereas γ_s, δ_τ, and ψ_{T_i} represent fixed effects for calendar time, duration, and maximum UI entitlement, respectively. Vector ϕ represents our parameters of interest, capturing changes in the risk of reporting sick close to the maximum UI duration. In our baseline regressions, we normalize ϕ_3 to zero. This implies that the spike is measured relative to the third month prior to maximum UI duration. In the absence of worker observables and random effects for worker unobservables, duration fixed-effects represent both genuine duration dependence effects and sorting effects due to heterogeneity in worker characteristics. As discussed below, we estimate the impact of sorting effects in a later stage by including individual characteristics as additional control variables.

Given that the elapsed duration in UI plus the residual entitlement of UI equals the maximum period of UI benefit entitlement, simultaneous estimation of δ, ϕ and ψ is only possible by imposing restrictions on some of the parameters. In our baseline specification, we therefore group workers with maximum entitlement periods of 4-6 months, 7-12 months, 13-24 months, and 24+ months. Estimation of ϕ and ψ is thus based upon variation in duration on UI and residual UI entitlement for workers with similar, but not identical maximum UI entitlement periods. As one of our robustness tests, we will additionally use a more narrow grouping of maximum UI entitlement, yielding similar results.

To analyze the persistence of the spike, we also estimate Equation (4.1) with SI risk dummies that are equal to one for workers entering into SI at time s *a*nd subsequently receiving SI for at least m months ($1 \leq m \leq 24$), and zero otherwise. In this way, we examine the extent to which the increased likelihood of calling in sick carries on over time. While the absolute risk level declines with respect to m, our primary interest then lies in the *r*elative risk of entering and staying in the SI scheme at the spike at of benefit exhaustion, as compared to the risk just before the spike.[9] Confidence intervals for the relative risks are obtained with the Delta method.

In the second step of our analysis, we study whether there are compositional differences driving the spike. For this, we employ a model specification that is similar to that of Equation (4.1), but considers worker characteristics of samples conditional on SI exits at elapsed time s:

$$X_{i,s,\tau} = \alpha + \gamma_s + \delta_\tau + \psi_{T_i} + \phi_{T_i - s} + \varepsilon_{is} \qquad \text{for} \quad SI_{is,\tau} = 1, \tag{4.2}$$

with $X_{i,s,\tau}$ indicating a characteristic of the worker and γ_s, δ_τ and ψ as duration, calendar-time, and maximum-duration fixed-effects. Again, the parameter of interest is ϕ, which indicates residual length-of-benefit effects on worker characteristics. For X we use measures of healthcare usage, demographics, and pre-UI labor market status. Health outcomes include the number of mental healthcare treatment minutes received per month,

[9] In the final month of the UI entitlement period, workers can only call in sick until a specific day that follows from the day they started receiving UI benefits. Hence, workers on average can only call in sick for 15 days, which reduces the probability to call in sick in that month. We therefore effectively exclude the final month from our regressions.

the probability of receiving mental healthcare treatment, and annual spending on non-mental healthcare. The demographics we examine include age, gender, nationality, and education level. We also use the number of monthly working hours, the monthly wage, and the type of contract (permanent or fixed-term) as pre-UI labor market outcomes.

Similar to Equation (4.1), we estimate Equation (4.2) on outcome variables that incorporate the persistence of the spike. Specifically, we consider the characteristics of UI recipients that leave UI for SI with a given residual entitlement length and continue receiving SI for at least m months ($1 \leq m \leq 24$). Again, the idea is that relative spike effects in the composition of workers entering SI may fade out as the time in SI proceeds. If spikes into SI consist of workers with only mild health conditions, one would expect these to have more opportunities to exit into employment and to be screened out in the first-year assessment. The fraction of workers with mild health conditions in the spike will then decrease with respect to m.

To test whether the observed compositional differences explain the spike into SI, we re-estimate Equation (4.1) while sequentially including demographic, pre-UI labor market status and healthcare utilization measures as control variables. A Blinder-Oaxaca decomposition is used to assess the explanatory power of the compositional differences.

Finally, we investigate whether workers entering SI at benefit exhaustion are affected differently by external screening from the one-year assessment and the DI assessment after 24 months of SI benefit receipt. For this, we re-estimate Equation (4.2) on the award decision outcomes of the sub-sample of observations at the moments of external screening. We then test whether the spike cohort is more or less likely to have an assessed degree of disability above 35%. We next add various controls to test the extent to which differences in composition explain differential effects of external screening.

As an alternative to our baseline specification, we will conduct various robustness tests that impose alternative restrictions on parameters. Specifically, we extend the pre-spike cohort and model elapsed and residual duration effects using polynomials instead of piece-wise fixed effects to be able to allow for a fully flexible specification of maximum entitlement effects. The exact specifications can be found in Appendix Section 4.A. These robustness tests are performed for each step of our analysis.

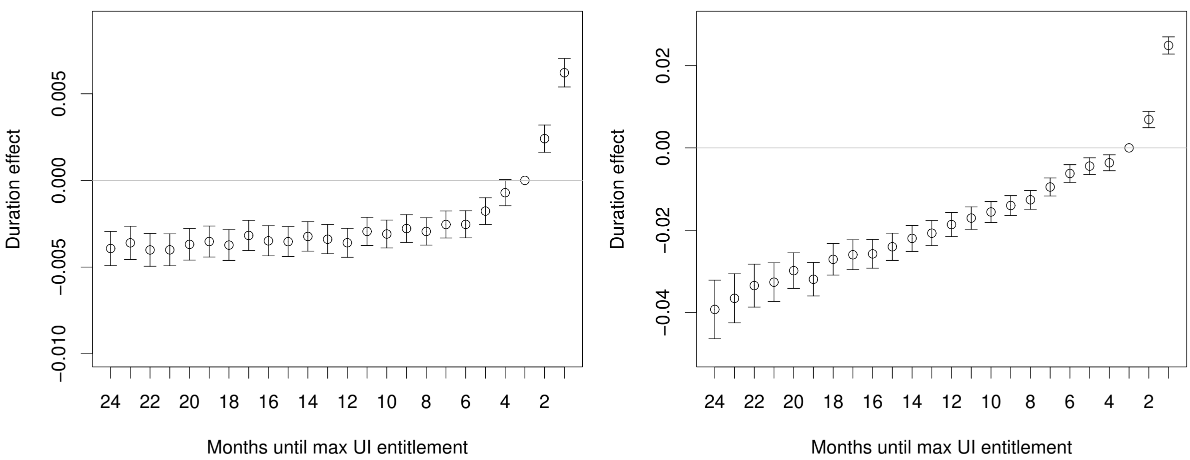
4.6 Results

4.6.1 Spikes in outflow

Figure 4.1 shows the estimated duration effects for exits into SI (left-side panel) and exits into employment (right-side panel). Both panels show a significant increase in exits as workers approach their maximum UI duration. The spike towards employment is larger

Figure 4.1: Duration effects for exits into SI (left) and into employment (right) with and without control variables

(a) Exits into SI (b) Exits into employment



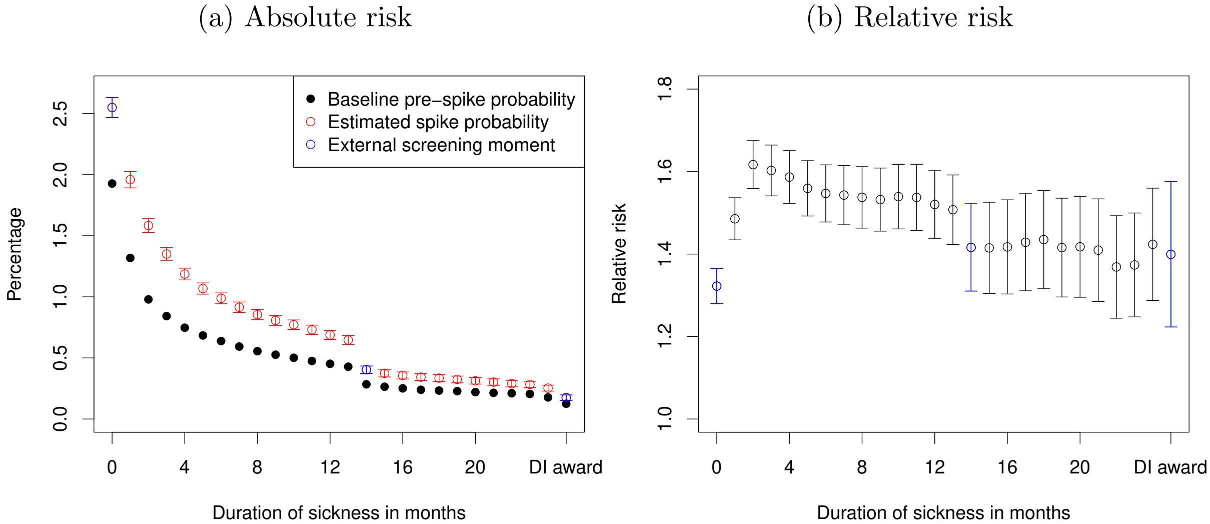
in absolute size but smaller in relative magnitude. The average monthly probability of entering the SI scheme is approximately 1%, while the average probability of finding employment is approximately 9%. In effect, the probability to report sickness increases by almost one-third, while the probability of finding employment increases by approximately one-fifth. The latter effect is comparable to De Groot & Van der Klaauw (2019), who use older data on UI spells from the Netherlands to study exits into employment.

Since behaviors at UI benefit exhaustion may differ across groups with different UI entitlement lengths, we additionally also re-estimate the duration effects, stratified by maximum UI groups. The confidence intervals are substantially larger for these (smaller) samples, but point estimates are very similar to those for the pooled sample (See Appendix Figure 4.A.3). This suggests that the spike is relevant for all groups. Furthermore, the estimated spike is also robust to a range of alternative specifications, as shown in Appendix Section 4.A.

We next examine the persistence of the spikes into the SI scheme in Figure 4.2. As explained earlier, we redefine the outcomes to risk indicators for SI receipt for at least m months ($m \leq 24$). The left panel shows the obtained probabilities of continued SI receipt for at least m months, while the right panel shows the relative risk, i.e. the probability of the spike-cohort as a ratio of the probability of the pre-spike cohort. The absolute risks clearly show a decreasing pattern for pre-spike and spike cohorts, with the majority of all workers exiting SI in the first four months. The next substantial decrease in SI receipt occurs at the one-year-assessment. The maximum entitlement period ends after 24 months when the DI assessment takes place (if relevant).

When comparing the spike and pre-spike cohorts in Figure 4.2, we infer that the

Figure 4.2: Estimated absolute risk and relative risk (ratio of absolute risk spike-cohort and pre-spike cohort) of reaching specified SI benefit durations for worker exits before and at the spike



exit rates out of SI benefits are substantially higher for the pre-spike cohort in the first three months of SI receipt. While unemployed workers in the spike cohort have a 30% higher probability of starting SI benefit receipt, this relative risk increases to 60% for the continued SI receipt for at least three months. In the remainder of the first year of SI benefits, the relative risk stabilizes, indicating that exit rates evolve similarly in both cohorts in these months. The first-year assessment decreases the relative risk, reflecting the fact that relatively more workers in the spike cohort are screened out; we discuss the impact of this assessment in more detail in Section 4.6.2.3. After the first-year assessment, the relative risk remains stable at approximately 1.4, which is still slightly exceeding the initial relative risk of 1.3. The DI assessment after two years does not affect the relative risk. The relative risk remains above the initial relative risk of 1.3, implying that workers in the spike cohort have a higher likelihood of continued SI receipt for two years and, eventually, also a higher likelihood of eventually being awarded DI benefits.

4.6.2 Compositional effects

Our next aim is to characterize (marginal) applicants at the spike of exits out of UI into the SI scheme. As shown in Figures 4.3 to 4.8, we track averages of characteristics of survival cohorts of SI benefit recipients just before and at the spike. The left-hand panels of these figures show the averages of characteristics of the pre-spike and spike cohorts, whereas the right-hand panel shows the relative averages of the spike cohort compared to the average of the pre-spike cohort. As a reference point, the first estimate in both panels reveals (relative) averages for the cohorts at risk of entering SI (those receiving

UI), irrespective of exits into SI. The second estimate, which is printed in blue, concerns samples of pre-spike and spike cohorts that report being sick and are at the start of SI benefit receipt. The subsequent estimates concern pre-spike and spike cohorts conditional on being sick for at least m months for $1 \leq m \leq 24$. The estimated compositional differences are robust to a range of alternative specifications, as shown in Appendix Section 4.A.

4.6.2.1 Demographics and labor market performance

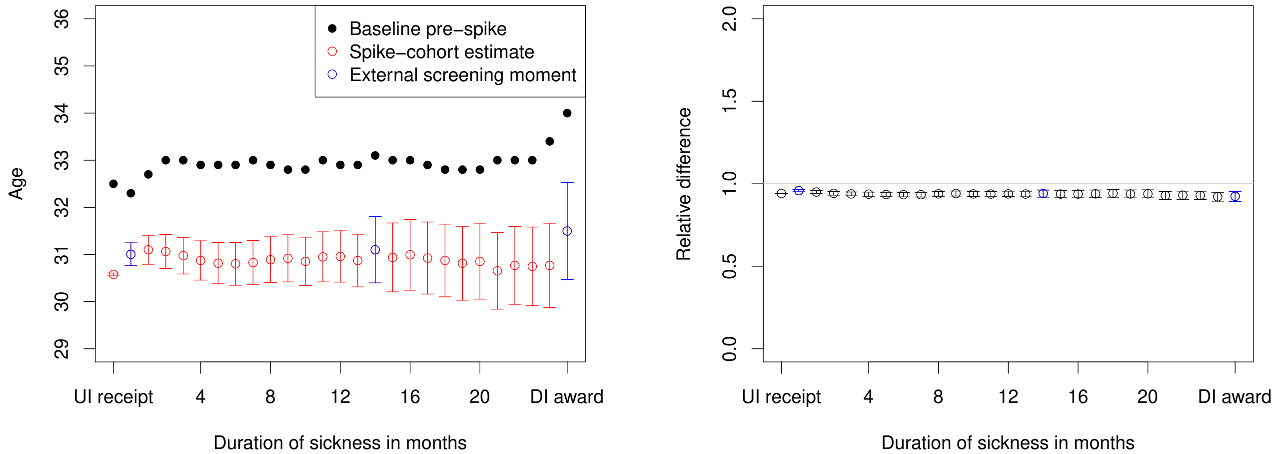
We first consider differences in demographic characteristics between cohorts before and at UI benefit exhaustion. Figure 4.3 shows differences in age (panel (a)), gender (panel (b)), and migration background (panel (c)). We find no gender differentials, but a significant and persistent age differential between cohorts that is about two years. A mechanical effect of tenure on maximum UI entitlement largely drives this age differential. Specifically, when comparing two workers with equal elapsed UI durations but a one-month difference in residual UI benefit entitlement, there is a one-month difference in total entitlement. This difference is mirrored by an age differential since older workers tend to have longer employment histories and longer UI entitlement.[10] The estimated age differential of somewhat less than two years reflects this effect. It is important to note that this difference is already present for the samples at risk of entering SI and remains similar when we condition on SI receipt for at least m months, with $1 \leq m \leq 24$. This means that even though there is a mechanical age difference between the two cohorts, there is no selection on age for workers reporting to be sick after the start of benefit receipt.

Zooming into the migration background of both cohorts (panel (c))), we do find evidence that (marginal) applicants at the spike are different. As a reference point, it is important to stress that workers receiving UI benefits in the last month of UI entitlement are already more likely to have a migration background. This reflects the fact that exit rates out of UI into employment are significantly lower for workers with a migration background. The over-representation of workers with a migration background in the spike widens when we condition on exits into SI, implying that they are more likely to enter SI in the spike. The relative difference increases slightly during the first three months of SI receipt and then stabilizes. One potential explanation for this could be that workers with a migration background are initially less aware of their entitlement to SI, and are therefore more responsive to notifications of this possibility as they approach their maximum UI entitlement. Alternatively, they may have worse labor market prospects, which explains a higher likelihood of (continued) receipt of SI.

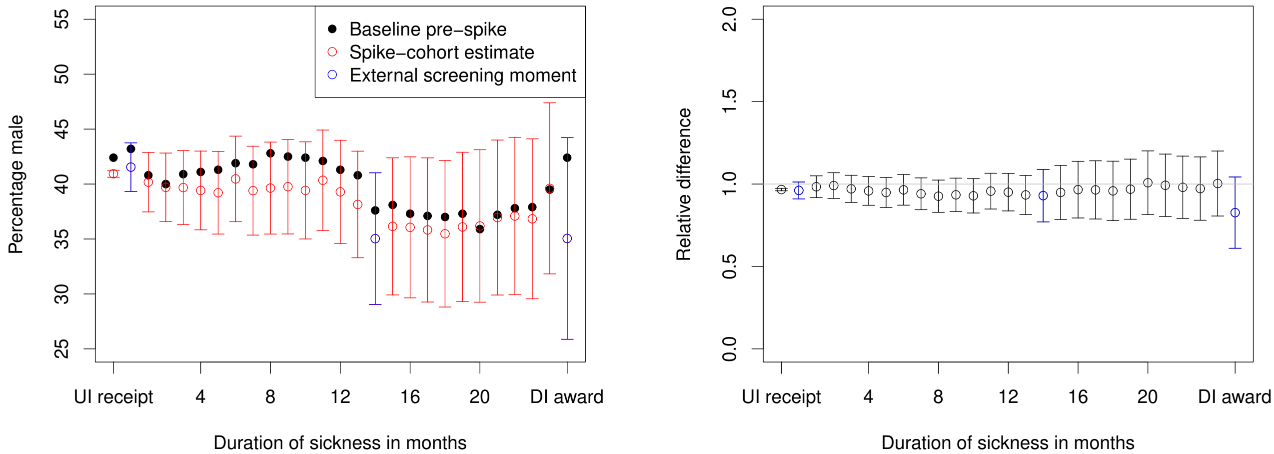
[10] When we exclude the controls for the duration on UI, the estimated difference between the pre-spike and spike cohort is negligible.

Figure 4.3: Estimated absolute (left) and relative (right) demographics of pre-spike and spike cohorts (95% confidence intervals)

Panel a: Age (in years)



Panel b: Gender (0/1 = Female/Male)



Panel c: Native vs. migration background

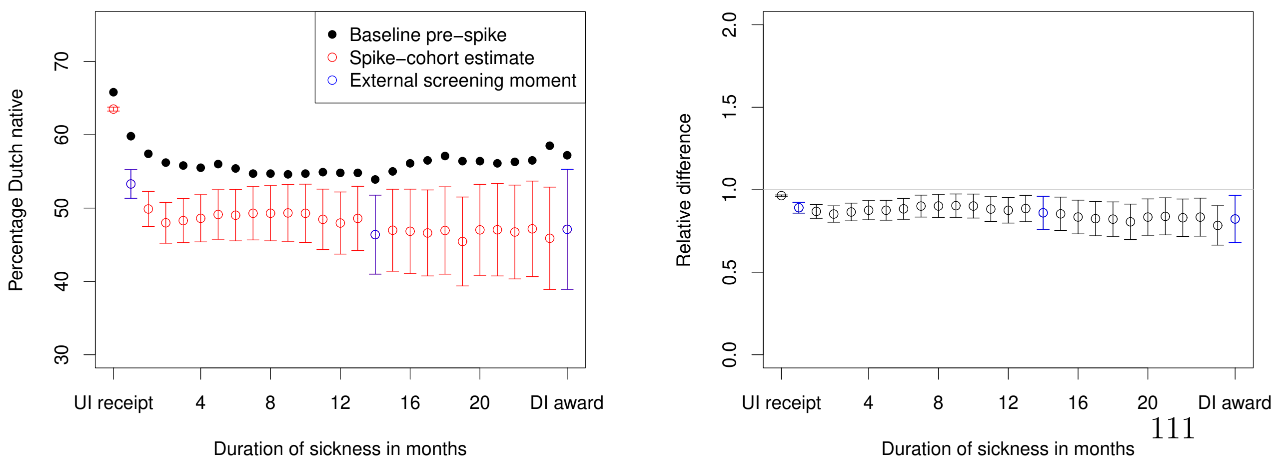
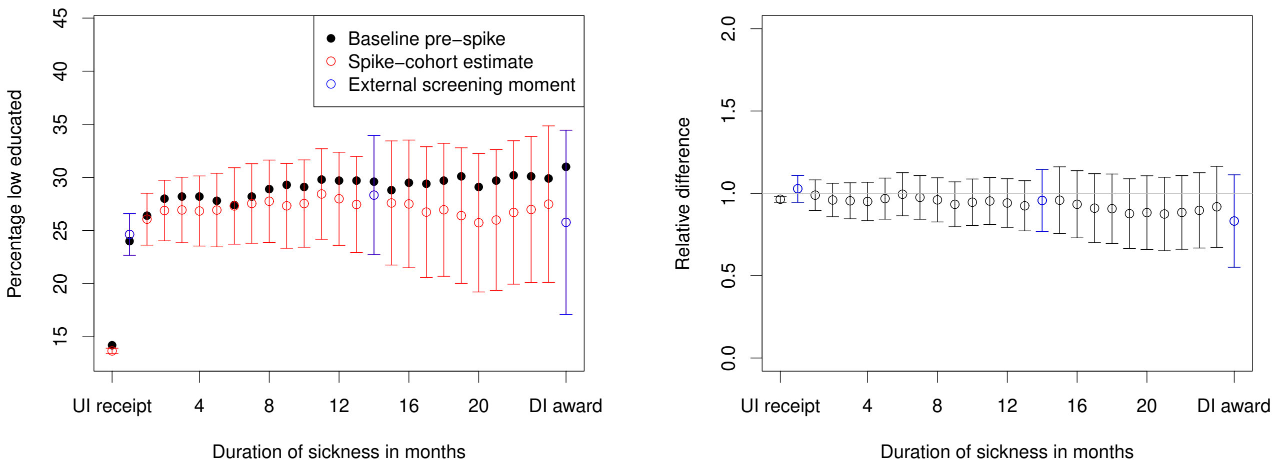
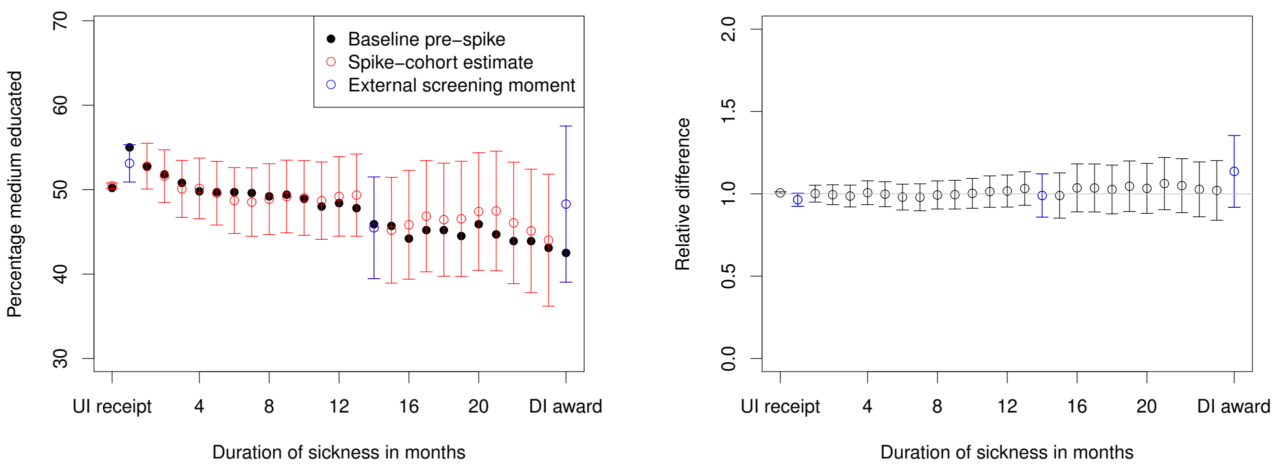


Figure 4.4: Estimated absolute (left) and relative (right) education level fractions of pre-spike and spike cohorts (including 95% confidence intervals)

Panel a: Fraction with low education level



Panel b: Fraction with medium education level



Panel c: Fraction with high education level

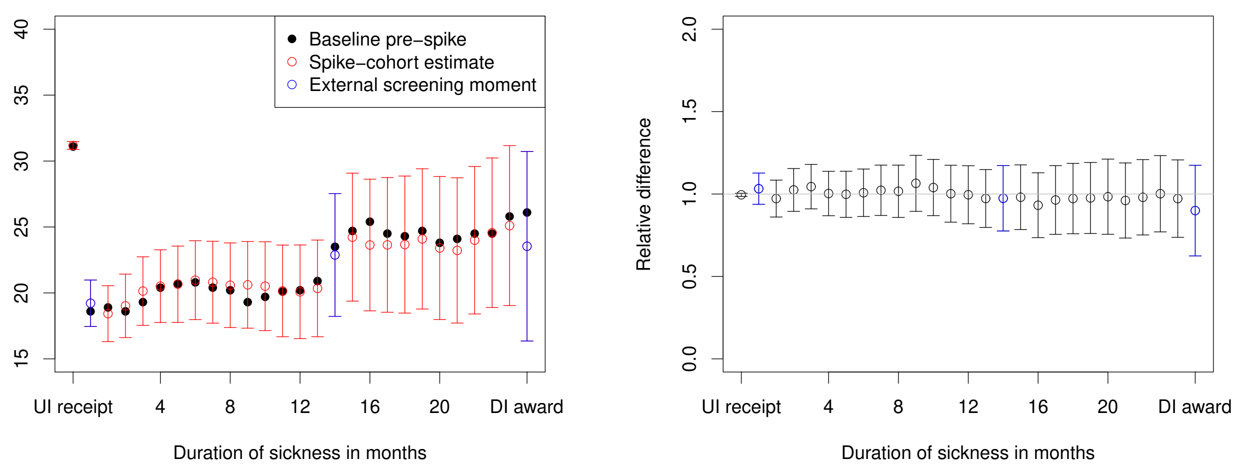


Figure 4.4 shows differences in the education-level fractions. For both spike and pre-spike cohorts, we see that the fraction of highly educated workers increases after 12 months of SI receipt; this mirrors the effect of external screening at the first-year-assessment. With higher pre-sickness earnings, highly educated workers typically have a higher chance of decreases in their assessed earnings capacity that exceed the 35%-threshold (Koning & Lindeboom, 2015). For all education levels, we find no significant differences between the cohorts.

To shed more light on the potential differences in the pre-UI labor market status, Figure 4.5 compares monthly working hours, labor earnings, and the fraction of fixed-term contracts in pre-spike and spike cohorts. Note that all these variables are measured just before the start of UI receipt. The spike cohort has fewer monthly working hours prior to entering UI and is more likely to be employed through a fixed-term contract. Corresponding to this, their pre-UI monthly labor earnings are lower. These differences in pre-UI employment could again reflect differences in labor market prospects.

The previous analyses indicate that the pre-spike and spike cohorts are fairly comparable in terms of demographic characteristics. The main exception to this concerns migration background. While approximately 45% of those who report being sick in the pre-spike cohort have a migration background, this percentage increases to 55% in the spike cohort. Assuming that the average migration background of the non-marginal spike applicants is similar to that of the pre-spike cohort, we infer that 80% of the marginal applicants have a migration background (Almond & Doyle, 2011; Godard et al., 2022). Similarly, the marginal applicants appear to have a weaker pre-UI labor market status.

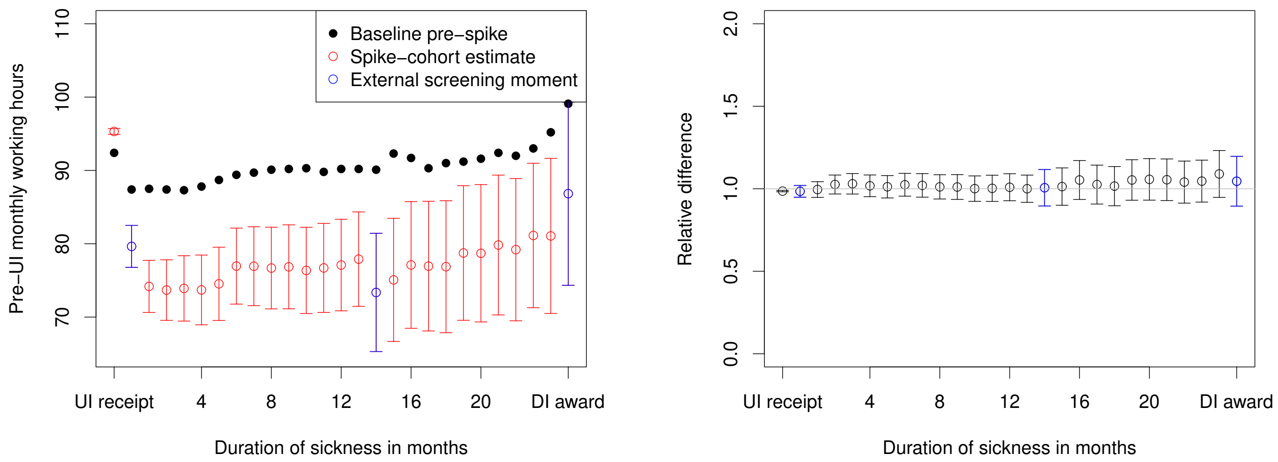
4.6.2.2 Healthcare utilization

Next, we examine differences in healthcare utilization. The resulting comparisons for the incidence (panel (a)) and the number of minutes (panel (b)) of mental health care are shown in Figure 4.6. Both panels show marked increases in mental health treatment utilization over time. This mirrors the joint effect of selection and a genuine worsening of mental health conditions of the pre-spike and spike cohorts. There is also evidence of screening effects at the one-year-assessment.

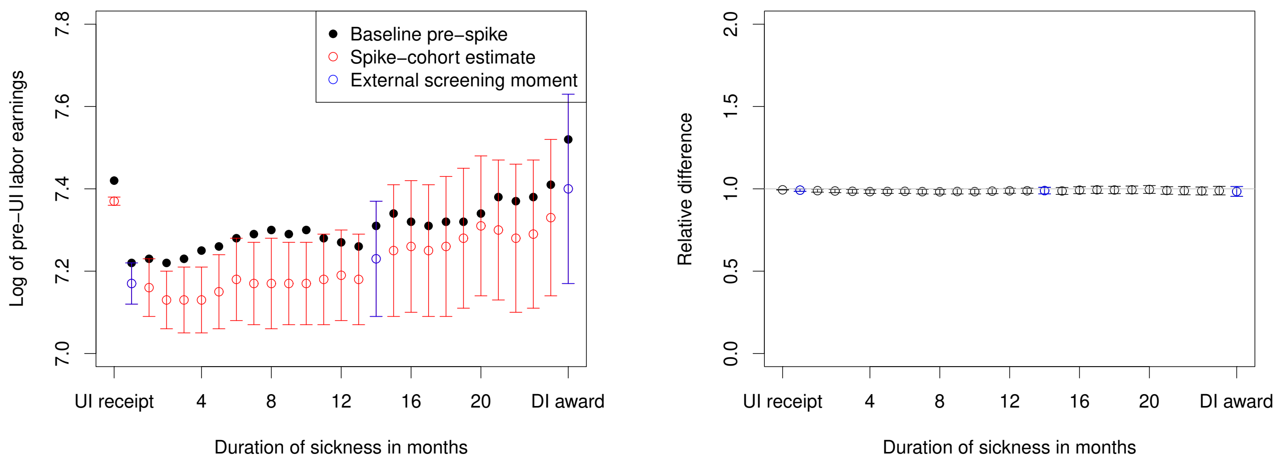
At the start of SI receipt, we see that the probability to receive mental health treatments and the number of mental health treatment minutes are virtually equal for the pre-spike and spike cohort; this is indicated by the second dot (in blue) in the respective panels. To test whether the absence of differences in average treatment does not mask differences in the distribution of treatment, we additionally examine the probability to exceed certain treatment thresholds (i.e. the 25th, 50th, 75th, and 95th percentile of the

Figure 4.5: Estimated absolute (left) and relative (right) pre-UI job characteristics of pre-spike and spike cohorts (95% confidence intervals)

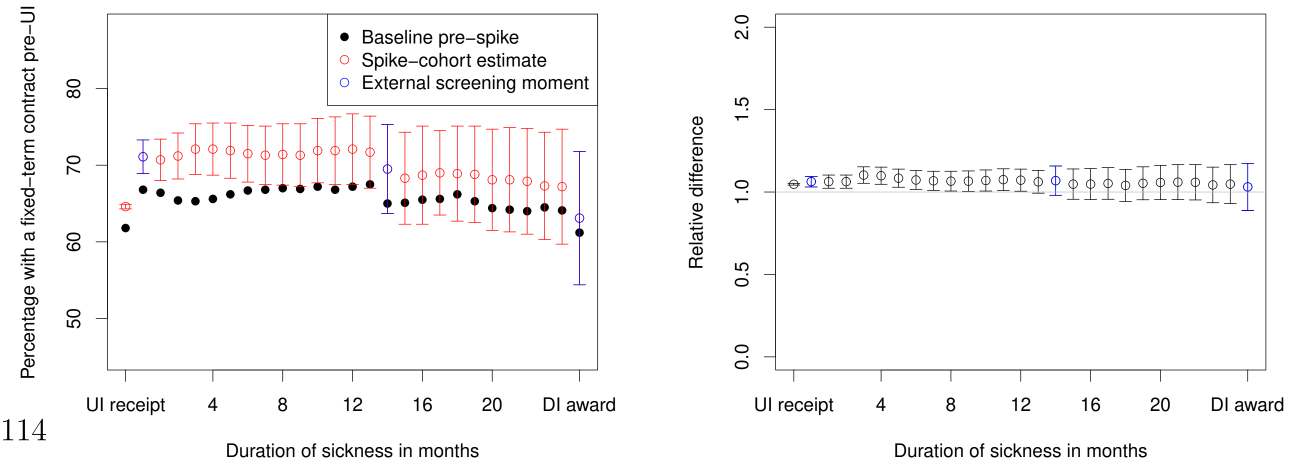
Panel a: Working hours per month



Panel b: Log of labor earnings per month



Panel c: Fraction with fixed-term contract



treatment distribution).[11] Using these alternative measures, we again find no significant difference in healthcare utilization for the cohorts entering SI. This does not support the idea that marginal applicants with less severe health conditions strategically time their applications. Assuming that the mental health treatments are a valid proxy for health, it suggests that the spike represents initial non-take-up of SI benefits or some acceleration of take-up that would have occurred later with continued entitlement to the SI scheme.

In the following months of SI receipt, however, there is a statistically significantly lower incidence of mental health treatments for the conditional spike cohort, amounting to about three percentage points. This difference cancels out after one year of SI receipt, which may mirror the offsetting effects of the first-year-assessment. The number of treatment minutes after three months of SI receipt is also significantly and substantially (20%) smaller for the spike cohort. Given the smaller incidence differential, we interpret this as an intensive margin effect of treatments. The difference in treatment minutes increases during the first three months of SI, corresponding to a widening of the relative risk of the spike cohort during these months as shown in Figure 4.2. In this period a large fraction of workers leave SI, and this fraction is substantially larger for the pre-spike cohort than for the spike cohort. With differences in exit rates that stabilize after three months of SI receipt, the difference in average treatment minutes also stabilizes over time. This contrasts with the difference in the incidence, which is no longer there after one year of SI receipt. Compared to the pre-spike cohort, it therefore seems that the first-year-assessment screens out SI recipients in the spike without mental health treatments, rather than those with positive but fewer treatment minutes.

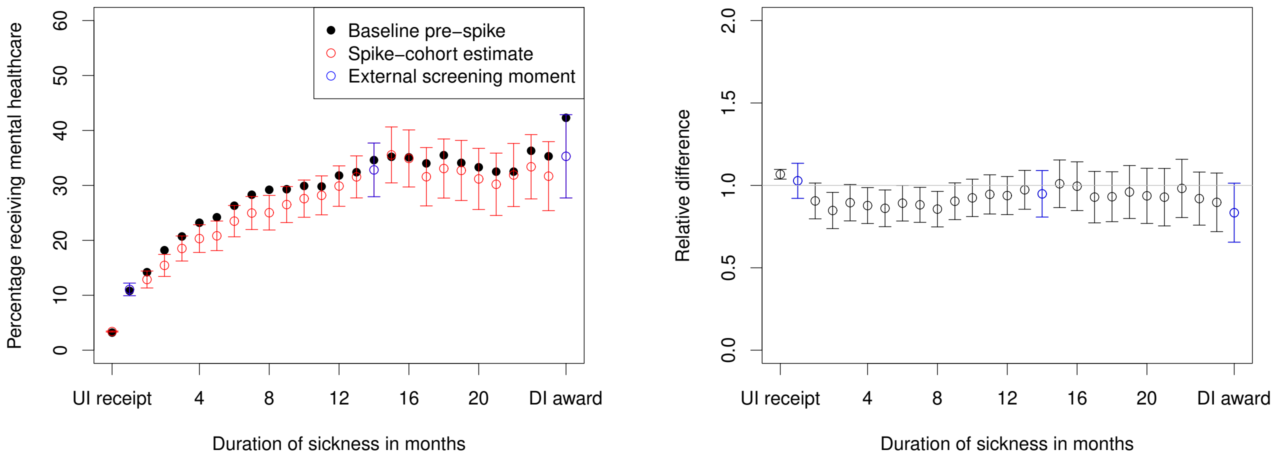
We infer that marginal applicants use considerably less healthcare than non-marginal applicants before the spike by again assuming that non-marginal applicants in the spike cohort are comparable to applicants in the pre-spike cohort. As an extreme case, the pre-spike cohort reaching three months of SI receipt is treated 50 minutes per month on average, as against 30 minutes per month for the spike cohort. With monotonous effects, marginal applicants would receive almost no mental health treatment minutes. While this should be interpreted as an upper bound, it illustrates that marginal applicants in the spike that continue receiving SI are likely to be healthier. With such differences being absent at the start of SI receipt, there appears a stronger impact of self-screening of healthy pre-spike workers that leave SI benefits in the first months.

While our data on non-mental healthcare are less detailed (and only available on an annual basis), they can be used to test whether the differences we observe in mental healthcare utilization are mirrored by similar patterns for non-mental healthcare utilization. Panel (c) of Figure 4.6 shows the annual non-mental healthcare spending measured

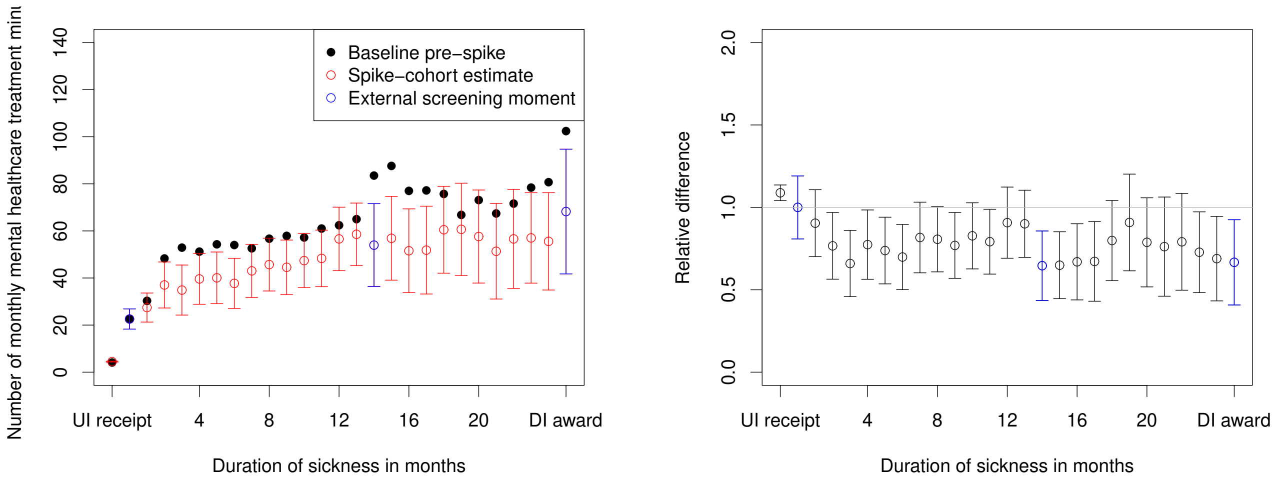
[11] These estimation results are available upon request.

Figure 4.6: Absolute (left) and relative (right) healthcare utilization of pre-spike and spike cohorts (95% confidence intervals)

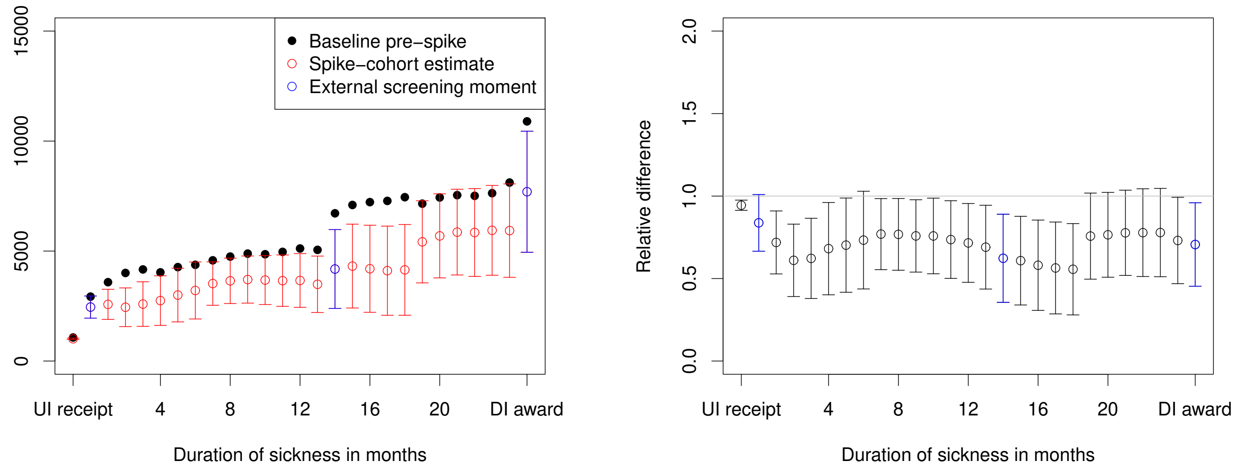
Panel a: Fraction with current mental health treatment



Panel b: Current mental health treatment minutes



Panel c: Annual non-mental healthcare spending



at the moment of inflow into SI for the spike and pre-spike cohorts. The resulting pattern is similar to that observed for mental healthcare utilization: differences in exit rates during the first four months result in significantly lower average amounts of spending on non-mental healthcare for the remaining spike cohort.[12] While the external assessment moments have a significant impact on the spending for both cohorts, the gap in healthcare utilization remains fairly constant. Decreasing sample size does result in insignificant estimates for the sample on SI for more than six months.

With monthly information on mental healthcare utilization, inferences on pre-spike and spike cohort differences represent the joint effect of (initial) selection effects and time-variant changes. This contrasts with the earlier comparisons, where variables were constants that were measured at the end of UI benefit entitlement. The differences in mental healthcare utilization in Figure 4.6 can be driven by selection effects or genuine changes in health conditions of the pre-spike and spike survival cohorts. To investigate this, Figure 4.7 shows differences in mental healthcare utilization as measured at the end of UI. As mental healthcare utilization is fixed at this moment, any differences are driven by selection rather than by differences in the evolution of health. Zooming into the first months of SI receipt, we see that the widening of the difference in the incidence of receiving mental health treatments in panel (a) and in the number of treatment minutes in panel (b) is very similar to the corresponding panels in Figure 4.6. This lends credence to the idea that – at least in the first months of SI receipt – differences are driven by selection effects.[13]

Alternatively, we examine differences in the evolution of health by tracking unconditional cohorts and looking at their healthcare utilization starting from the moment of the start of SI receipt. This approach thus considers all pre-spike and spike workers, irrespective of continued SI receipt. The figures for both the incidence and number of treatment minutes are shown in Appendix Figure 4.A.4. While the incidence levels shown in panel (a) are very comparable, panel (b) of the figure shows that the increase in healthcare utilization is less pronounced for the spike cohort than it is for the pre-spike cohort. This would suggest that part of the differential is driven by genuine health differentials. Still, the effect is small, insignificant, and cancels out over time.

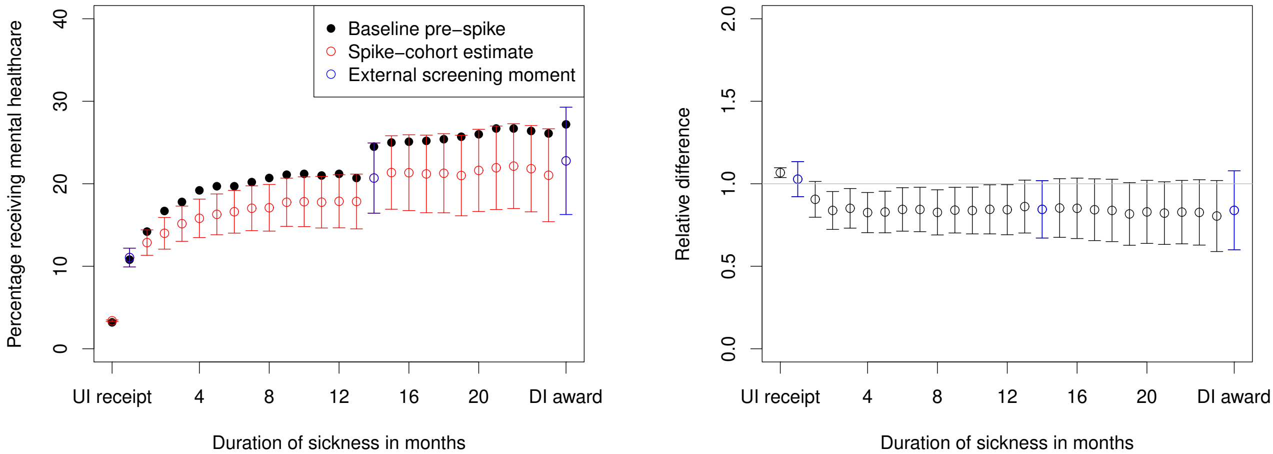
Our results on mental health care utilization entail comparisons of the incidence or minutes for specific months. While being informative on the targeting of SI benefits on workers with more or less severe health conditions, we cannot draw inferences on the

[12] Similar as for mental health treatment, we additionally examine the distribution of healthcare expenditure. For the cohorts flowing into SI, we find no significant differences in the full distribution.

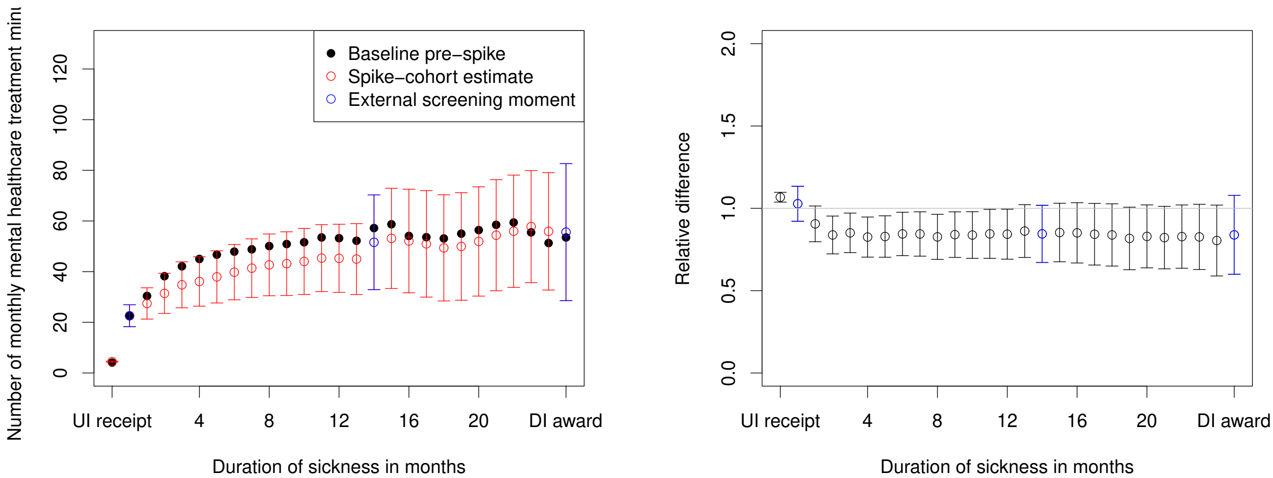
[13] For later months, interpreting differentials in mental health utilization measured at the end of UI benefit receipt becomes more cumbersome, since the initial levels become less relevant for the current ones.

Figure 4.7: Absolute (left) and relative (right) mental healthcare utilization of pre-spike and spike cohorts at the end UI receipt (95% confidence intervals)

Panel a: Fraction with mental health treatment at end of UI



Panel b: Mental health treatment minutes at end of UI

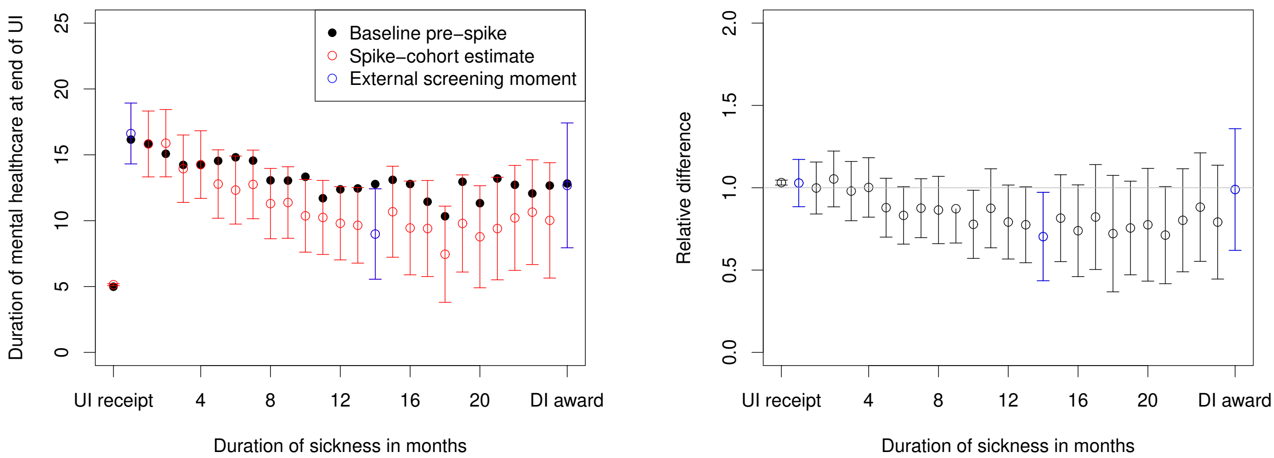


timing of entries in SI compared to the start or end of mental health treatments. We therefore re-estimate our baseline model with the elapsed and the residual duration of mental health treatments at the end of the UI spell for the cross-sections of workers with mental health treatments at that moment. If the spike in SI is caused by workers who could have reported sick earlier without incentives to bunch, one would expect that the average duration of treatment at the moment of sickness reporting would be longer.

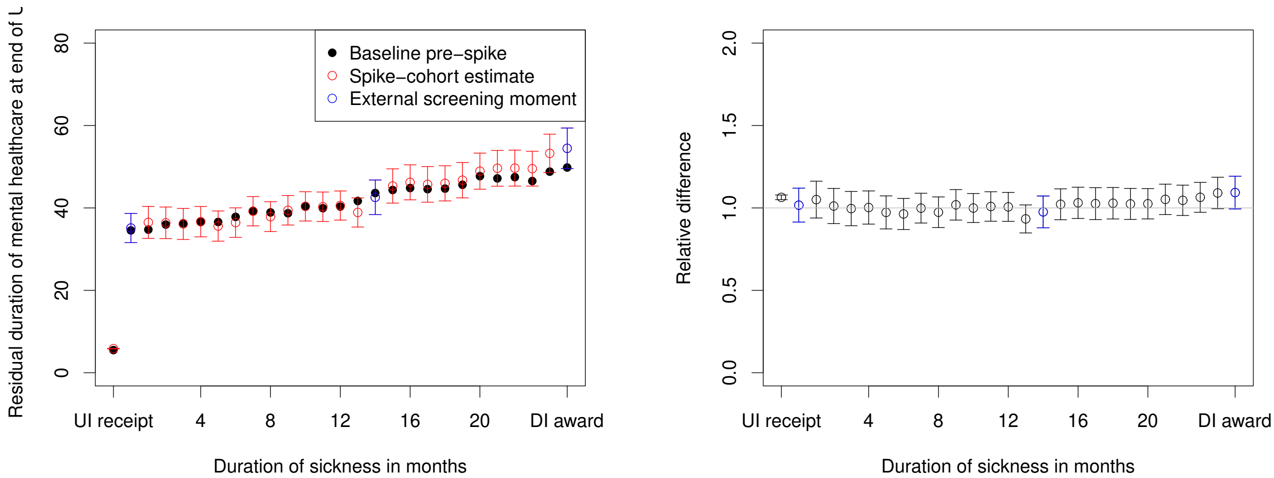
Figure 4.8 shows the results. We find that the elapsed duration is similar during the

Figure 4.8: Absolute (left) and relative (right) differences in elapsed and residual duration of mental health treatments of pre-spike and spike cohorts (95% confidence intervals)

Panel a: Elapsed duration of mental health treatment, measured at end of UI



Panel b: Residual duration of mental health treatment, measured at end of UI



first months of SI receipt and is shorter for the spike cohort in later months. The latter effect most likely reflects the earlier-mentioned selection differentials. To rule out that the equality in mean elapsed duration hides strong differences in the distribution of elapsed duration, we perform an F-test for equality of variances of the pre-spike and spike cohorts at the start of SI receipt. Since we cannot reject the hypothesis of equal variances[14], we conclude there is no combined effect of delayed and faster applications in the spike. This finding does not only confirm that cohorts entering into SI are similar in terms of

[14]The estimated variances amounted to 164 and 173 for pre-spike and spike cohorts, respectively, with a P-value of 0.09.

mental healthcare utilization, but also that the spike does not stem from timing behavior of UI recipients during their mental health treatments. In line with expectations, we also see that selection effects are mirrored by shorter elapsed durations and increased residual durations, as measured at the end of UI receipt. This implies that workers who remain sick started mental health treatment later, but will continue receiving treatments for longer. The impact of these selection effects is similar for the pre-spike and spike cohorts.

4.6.2.3 The impact of external screening

Our findings suggest that the health status of those entering SI before and at the end of their UI entitlement is similar. Lower exit rates out of SI do result in differences in the health status between the surviving cohorts. The figures on the persistence of the spike and the compositional differences seem to indicate that these differences between the cohorts are partly offset during the first year sickness assessment, as more individuals in the spike cohort are screened out. We now formally test whether the observed differences are indeed counteracted by the first-year assessment and the DI assessment that takes place after two years of SI receipt. A natural way to do so is by comparing the award rates of pre-spike and spike workers before and after the first- and second-year assessments. Table 4.1 therefore presents the average award probabilities of pre-spike and spike cohorts for the first-year assessments in columns (1) and (2) and for the final DI assessment in columns (4) and (5), respectively. In addition, it shows the corresponding relative probabilities in columns (3) and (6). The first line shows the observed (unconditional) award rates. Knowing that there are differences in the composition of both cohorts, we subsequently derive award rates that are based on the observed characteristics of the full sample of SI recipients. As such, we increasingly control for differences in observables that may change the implied conditional pre-spike and spike award rates.

For the first-year sickness assessment, workers in the spike cohort have a significantly lower probability of having an assessed degree of disability exceeding 35% and therefore are more likely to be screened out. Their award rate is 58.4%, as compared to 63.5% for the pre-spike cohort. Recall that this difference is mirrored by the drop in the relative risk of SI receipt for at least 12 months in Figure 4.2. Most notably, adding demographic controls results in conditional award rates of both cohorts that are markedly higher. When including both demographic and pre-UI job controls, the conditional relative award rate of the spike cohort increases from 0.92 to 0.95, indicating that part of the lower award rate stems from these variables. As expected, the inclusion of health controls does not significantly affect the difference in the probability to pass the first-year assessment. As an additional check, we interact the spike dummy with health controls to test whether the impact of health differs for both cohorts. The estimated interaction terms are insignifi-

Table 4.1: Decomposition of difference in probability of having an assessed degree of disability above 35% in the first year assessment or the final DI assessment

Specification	First-year assessment award rate			DI assessment award rate		
	Pre-spike cohort	Spike cohort	Ratio	Pre-spike cohort	Spike cohort	Ratio
Unconditional	63.5%	58.4%	0.92*	67.4%	66.6%	0.99
Cumulative controls[a]:						
Elapsed time in UI	59.6%	53.6%	0.90**	62.3%	59.2%	0.95
+ Demographics	70.5%	66.0%	0.94*	70.7%	69.2%	0.98
+ Pre-UI job characteristics	71.5%	67.6%	0.95	72.6%	71.5%	0.98
+ Health utilization	70.0%	67.1%	0.96	71.8%	71.0%	0.99
+ Interaction terms	69.9%	67.1%	0.96	72.8%	70.3%	0.97

*/** Difference between award rates is significant at a 5%/1% significance level.
[a]: With the estimation of models with subsequently added controls, we derive conditional award rates that are based on the full sample of SI recipients. As such, we control for compositional differences of the pre-spike and spike cohort, as compared to the full sample.

cantly different from zero, implying that health has a similar effect on the probability of passing the first-year sickness assessment for both cohorts. We conclude that the impact of the one-year-assessment is more substantial for workers in the spike cohort.

For the final DI assessment, we find that the observed award rates are comparable between the cohorts. The inclusion of control variables does not affect the remaining difference in the award probabilities. These results suggest that the first-year sickness assessment screens out more workers in the spike cohort, while the final DI assessment affects both groups similarly. In terms of the assessed degree of disability, both groups are therefore comparable. This again renders it unlikely that the spike in sickness reports is fully driven by moral hazard of relatively healthy workers.

4.6.3 Synthesis

The first step of our analysis has shown that their is a spike in inflow into SI at UI benefit exhaustion. As shown in the second step of the analysis, the spike-cohort reporting sick at UI benefit exhaustion is more likely to have a migration background and a weaker pre-UI labor market status than the pre-spike cohort. We additionally see that outflow rates out of SI are lower for the spike cohort, leading to the lower relative use of mental health care treatments for the remaining cohort of SI recipients. In this final step, we investigate the impact of these compositional differences on the spike in SI inflow. To do so, we perform an Oaxaca-Blinder decomposition on the probability to enter SI (Equation (4.1)).[15]

[15] As an alternative approach, we sequentially control for compositional differences in Equation (4.1), as shown in Appendix Section 4.A. The results are similar.

Figure 4.9: Oaxaca-Blinder decomposition of the spike in sickness insurance inflow

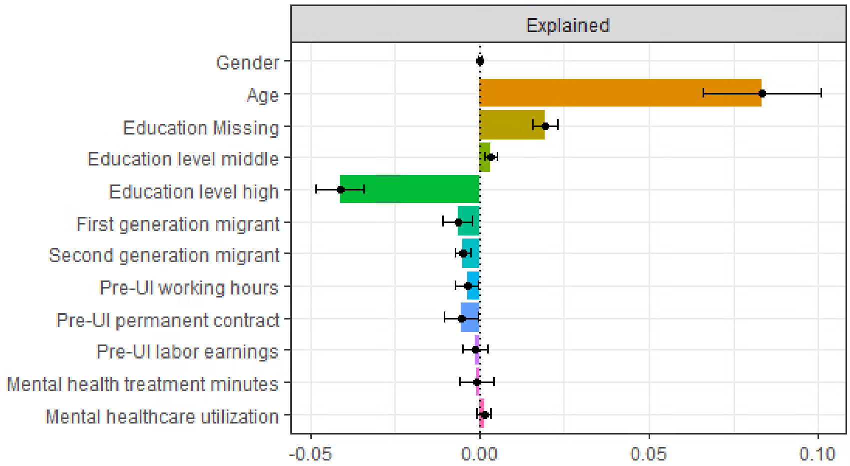


Figure 4.9 shows to what degree compositional differences between both cohorts explain the spike in inflow into SI .[16] The difference in SI inflow between the pre-spike and spike cohorts is 0.55 percentage points. Compositional differences between the two cohorts explain approximately 0.15 percentage points or 27% of the spike in inflow. Most notably, we see that the inclusion of age and educational attainment explains a substantial part of the spike in SI inflow. The other differences in composition we observe, such as the migration background, pre-UI job characteristics, and healthcare utilization do not explain the difference in inflow into and subsequent outflow out of SI.

In light of the limited age differential between the pre-spike and spike cohort, the effect of age on the inflow into SI may be perceived as striking. One potential explanation for this concerns the strong link between the workers' age and pre-UI labor market history. Given the presence of the maximum entitlement period dummies as controls, it is likely that the workers' age provides us with additional information that captures an important part of differences in pre-UI labor market histories. In particular, we can identify older workers with relatively short maximum entitlement periods that have experienced earlier unemployment spells and therefore are likely to have a weaker labor market position. These differences are also informative on the return to work probability of SI recipients.

4.7 Discussion

We have shown that workers receiving UI benefits in their final month of UI entitlement are about 30% more likely to enter into the Sickness Insurance (SI) scheme, as compared to cohorts prior to the end of their entitlement period. From a policy perspective, this raises the question of whether the additional inflow in the spike is caused by moral hazard

[16]Results are similar when looking at the spike into DI, as shown in Appendix Section 4.A.

or whether it represents initial non-takeup. Moral hazard would stem from workers with less severe health conditions that aim to maximize the total period of benefits receipt. Alternatively, the spike may resemble the catch-up of workers with health problems that were initially unaware of the possibility to apply for SI benefits while on UI or are prone to status-quo bias. This explanation is more likely if the health conditions of the spike cohort are similar to those of the pre-spike cohort.

In terms of healthcare utilization, we find similar outcomes for the pre-spike and spike cohort at the start of SI benefit receipt. This is at odds with the idea that relatively healthy workers enter into SI to extend the total period of benefit receipt. To back up this finding, we also hypothesized that the optimal timing of SI benefit applications at the end of UI entitlement would go together with longer elapsed mental health treatment durations. Since these durations are almost equal for spike en pre-spike cohorts, there is no evidence of systematically different application behaviors of the spike cohort.

Without evidence pointing at the strategic timing of SI benefit applications, we consider it likely that workers in the spike initially lacked the information about their entitlement to SI or that they suffered from status-quo bias. We also find that the spike cohort has a substantially higher fraction of workers with a migration background. We suspect that this group was less aware of their eligibility for SI, inducing non-takeup which (partly) reversed due to the information given at the end of their UI entitlement period. There is also evidence that workers in the spike cohort showed weaker labor market status before entering into UI. They worked fewer hours per month, were more likely to have a fixed-term contract, and were more likely to have experienced earlier spells in UI.

Although the health conditions of pre-spike and spike cohorts that enter into SI are similar in all aspects that we have studied, we do find marked differences in terms of exit rates out of SI. This particularly concerns the first four months of SI receipt, wherein exit rates of the spike cohort are considerably lower, and, as a consequence of this, there is a divergence in average (mental) healthcare utilization of the remaining samples. This suggests that this cohort has a weaker labor market position or work preferences that lower the odds of work resumption. While we cannot fully capture these characteristics in observed controls, it is likely that behavioral failures that explain the spike – such as status quo bias – are also relevant for the (ex-post) behavior in the SI scheme. In light of the increased exits out of SI after 12 months, it appears that these workers in the spike with lower initial exit rates out of SI are partly screened out in the one-year assessment of the SI scheme.

We conclude that the spike into the SI scheme represents workers with very similar health conditions as the inflow in earlier months before benefit exhaustion, which in itself suggests that there is a catch-up effect of initial non-takeup. At the same time, the cohorts clearly show dissimilar labor market attachment. Workers in the spike more often have a migration background and weaker labor market positions and are probably prone to behavioral failures. To avoid unnecessary long SI spells, this therefore calls for timely screening on recovery in the first months of SI receipt. Obviously, this argument is also relevant for the larger population of workers with weak labor market positions that enter the SI scheme at any point in time during the UI spell. Additionally, the indication of non-take-up of SI also calls for (information) interventions for those on UI to avoid unnecessary long UI spells of sick unemployed workers.

4.A Appendix

4.A.1 Classification of healthcare expenditure categories

Table 4.A.1: Construction of mental healthcare expenditures and physical healthcare expenditures based on expenditure categories used by Statistics Netherlands

Expenditure category[a]	Mental healthcare	Physical healthcare
General practitioner		X
Pharmacy		
Hospital healthcare		X
Paramedical healthcare		X
Apparatus		
Hospital transportation		
Birth care		
Health care expenditures incurred abroad		
Other cost		
First-line psychological healthcare	X	
Mental healthcare	X	
Basic-mental healthcare	X	
Specialist mental healthcare	X	
Geriatric rehabilitation healthcare	X	
Nursing without stay		X
Sensory disability healthcare		

Note: (a) Expenditure categories as used by Statistics Netherlands

4.A.2 Theoretical model

We model the decision of unemployed workers to apply for SI or remain on UI. An unemployed worker with some level of health h, will receive SI and potentially DI benefits for n periods if they choose to apply for SI. As sicker workers are more likely to remain on SI and subsequently DI for a longer duration, the valuation of SI benefits depends negatively on the level of health h.

If a worker decides to remain on UI, he/she has to adhere to the UI requirements (apply for jobs, meet with a caseworker, etc.), which is more difficult for workers who are in poor health. Hence the valuation of UI benefit receipt, b^{UI} increases as health increases ($\frac{\partial b^{UI}}{\partial h} > 0$). To simplify the model, we assume that exits into employment are independent of the decision whether to stay on UI or apply for SI. Exits into employment are therefore not incorporated into the stylized model. The money utility values of states in period t can thus be written as:

$$U_t = \begin{cases} \sum_{j=1}^{n} \gamma^{j-1}\, b^{SI}(h) + \gamma^n\, U_{t+n} & \text{if SI and } t \leq T \\ b^{UI}(h) + \gamma\, U_{t+1} & \text{if UI and } t \leq T \\ b^{SA} + \gamma\, U_{t+1} & \text{if } t > T, \end{cases}$$

with SI an indicator for choosing to enter SI, b^{SI} the value of sickness benefits, b^{UI} the value of unemployment benefits and b^{SA} the value of social assistance (SA). SI benefits are received for n months. The subsequent period's utility, U_{t+1}, is discounted by a monthly factor γ.

We now first examine a scenario in which delaying DI applications is not possible. The worker either decides to enter SI today and receive the corresponding benefits for n months or does not apply for SI and receive either UI benefits or social assistance for n months. This yields the following comparison:

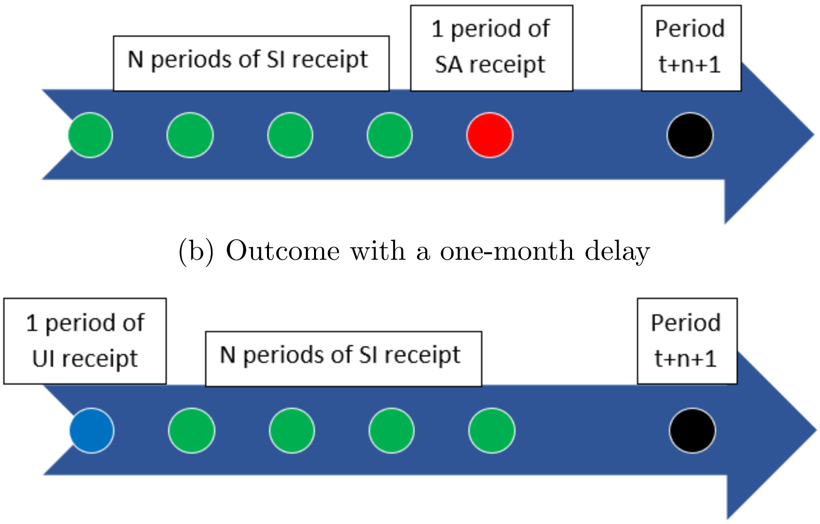
$$\begin{aligned} \sum_{j=1}^{n}\gamma^{j-1}b^{SI}(h) + \gamma^n U_{t+n} &> b^{UI}(h) + \gamma U_{t+1} \\ \sum_{j=1}^{n}\gamma^{j-1}b^{SI}(h) + \gamma^n U_{t+n} &> b^{UI}(h) + \sum_{j=2}^{T-t}\gamma^{j-1}b^{UI}(h) + \sum_{j=T-t+1}^{n}\gamma^{j-1}b^{SA} + \gamma^n U_{t+n} \\ \underbrace{\sum_{j=1}^{n}\gamma^{j-1}b^{SI}(h)}_{\text{Net present money utility value of SI}} &> \underbrace{\sum_{j=1}^{T-t}\gamma^{j-1}b^{UI}(h) + \sum_{j=T-t+1}^{n}\gamma^{j-1}b^{SA}}_{\text{Net present money utility value of UI and SA}} \end{aligned}$$

A worker thus chooses to apply for SI if the net present utility value of choosing to do so, exceeds the net present utility value of receiving either UI benefits or social assistance for the same duration. If it is optimal to apply for SI, workers can choose to do so immediately or to delay the application. Workers only benefit from delaying their SI application if the expected duration of SI benefit receipt n exceeds the remaining UI entitlement period. Additionally, we assume that delaying an SI application does not impact the duration of SI benefit receipt. One could argue that if the moment of recovery is unaffected by delaying SI applications, the duration of SI benefit receipt decreases as applications are delayed. If this would be the case, the incentives to delay would be smaller, as we briefly discuss at the end of this section.

The optimization problem is related to the remaining UI entitlement. Workers who have less than 13 weeks of UI entitlement remaining directly gain in terms of increased UI benefit duration by delaying their SI application. The choice problem faced by these workers is illustrated in Figure 4.A.1. Delaying the SI application by one month implies that workers gain one month of UI benefit receipt, receive their SI benefits one month

Figure 4.A.1: Choice problem for delaying SI application by one month

(a) Outcome without delay



later and lose one month of SA receipt. The theoretical model gives the same outcome, as shown below:

$$\begin{aligned}
b^{UI}(h) + \gamma \sum_{j=1}^{n} \gamma^{j-1} b^{SI}(h) + \gamma^{n+1} U_{t+n+1} &> \sum_{j=1}^{n} \gamma^{j-1} b^{SI}(h) + \gamma^{n} U_{t+n} \\
b^{UI}(h) &> \sum_{j=1}^{n} \gamma^{j-1} b^{SI}(h) - \gamma \sum_{j=1}^{n} \gamma^{j-1} b^{SI}(h) + \gamma^{n}(U_{t+n} - \gamma U_{t+n+1}) \\
b^{UI}(h) &> (1-\gamma) \sum_{j=1}^{n} \gamma^{j-1} b^{SI}(h) + \gamma^{n}(U_{t+n} - \gamma U_{t+n+1})
\end{aligned}$$

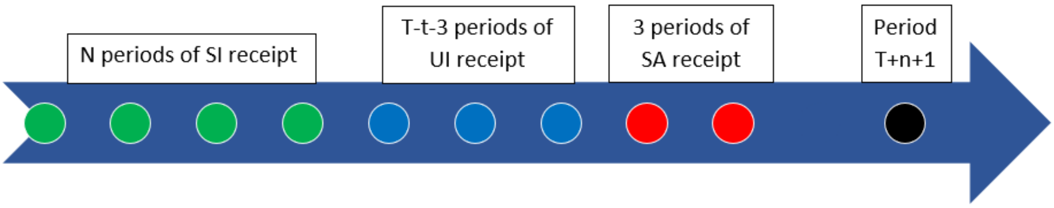
Since delay is only beneficial if the duration of SI benefit receipt exceeds the UI entitlement period, the remaining UI entitlement is zero when workers exit SI. Hence, after the receipt of SI, the worker can no longer return to UI but will receive SA benefits instead. $U_{t+n} - \gamma U_{t+n+1}$ is equal to the utility received in the first month after SI receipt ends and therefore equal to b^{SA}. The comparison then simplifies to

$$\underbrace{b^{UI}(h)}_{\text{Gained UI receipt}} > \underbrace{(1-\gamma) \sum_{j=1}^{n} \gamma^{j-1} b^{SI}(h) +}_{\text{Lost valuation of SI}} \underbrace{\gamma^{n} b^{SA}}_{\text{Gained SA receipt}}$$

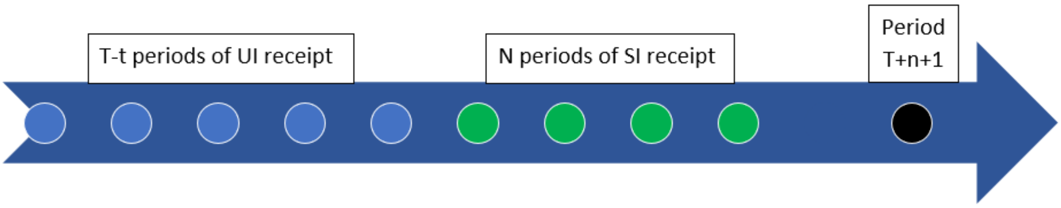
The worker will thus choose to delay the UI application if one additional month of UI receipt, exceeds the loss incurred through discounting by receiving SI benefits one month later, plus one month of SA received after SI benefits are terminated. Note that in the next month, the same condition still holds. Workers who choose to delay one month, are

Figure 4.A.2: Choice problem for delaying SI application by T-t month

(a) Outcome without delay



(b) Outcome with a T-t months delay



therefore inclined to delay as long as possible and apply for SI in their last month of UI entitlement.[17].

If workers are not yet within the last 13 weeks of their UI entitlement, they only gain in terms of UI benefit receipt if their remaining UI entitlement after delaying is less than 13 weeks. Only SI applications within the last 13 weeks of UI entitlement results in a phase-in period which is shorter than 13 weeks and thereby an increased duration of potential UI benefit receipt. For simplicity, we assume that if this worker decides to delay, he/she will delay as long as possible to gain as many additional months of UI receipt as possible. This worker thus has to delay their application for $T - t$ months, to gain three months (13 weeks) of UI entitlement. The choice they face is illustrated in Figure 4.A.2.

Delaying the SI application by $T - t$ month implies that workers gain three months of UI benefit receipt, receive their SI benefits $T - t$ month later and lose three months of SA receipt. The theoretical model gives the same outcome, as shown in the following equations:

[17]Note that alternative specifications of the discount factor can cause a worker to delay in one month, while not choosing to delay in another month

$$\sum_{j=1}^{T-t}\gamma^{j-1}b^{UI}(h)+\sum_{T-t+1}^{n+T-t}\gamma^{j-1}b^{SI}(h)+\gamma^{n+T-t}U_{n+T-t} \quad >$$
$$\sum_{j=1}^{n}\gamma^{j-1}b^{SI}(h)+\sum_{j=n+1}^{n+T-t-3}\gamma^{j-1}b^{UI}(h)+\sum_{j=n+T-t-2}^{n+T-t}\gamma^{j-1}b^{SA}+\gamma^{n+T-t}U_{n+T-t}$$

$$\sum_{j=1}^{T-t}\gamma^{j-1}b^{UI}(h)-\sum_{j=n+1}^{n+T-t-3}\gamma^{j-1}b^{UI}(h) \quad >$$
$$\sum_{T-t+1}^{n+T-t}\gamma^{j-1}b^{SI}(h)-\sum_{j=1}^{n}\gamma^{j-1}b^{SI}(h)+\sum_{j=n+T-t-2}^{n+T-t}\gamma^{j-1}b^{SA}$$

$$\underbrace{\sum_{j=1}^{T-t-3}\gamma^{j-1}b^{UI}(h)-\sum_{j=n}^{n+T-t-3}\gamma^{j-1}b^{UI}(h)}_{\text{Earlier UI receipt}}+\underbrace{\sum_{j=T-t-3}^{T-t-1}\gamma^{j}b^{UI}(h)}_{\text{3 extra months of UI receipt}} \quad >$$
$$\underbrace{(1-\gamma^{T-t})\sum_{j=1}^{n}\gamma^{j-1}b^{SI}(h)}_{\text{Lost valuation of SI}}+\underbrace{\sum_{j=n+T-t-3}^{n+T-t-1}\gamma^{j}b^{SA}}_{\text{Lost SA receipt}}$$

In this case, the worker will thus delay their SI application if the increased valuation of receiving their UI benefit immediately, plus the valuation of the three additional months of UI benefit receipt exceeds the loss in the valuation of postponing SI benefit receipt by $T-t$ months plus the three months of SA received after SI benefits are terminated.

As briefly mentioned, the theoretical model assumes that the duration of SI benefit receipt is unaffected by delaying the application. One could argue that this is unlikely, as workers are deemed to exit SI once they recover. In the most extreme case, this would imply that the duration of SI benefit receipt decreases one-to-one with the duration of the delay. Delaying an SI application would imply a re-labeling of SI receipt to UI receipt, without affecting the total combined duration of the benefit receipt. The amount of benefits received under UI is equal to the amount received under SI while the search requirements are arguably a costly aspect of UI. Under this scenario, delaying a SI application would therefore by highly unlikely. In some hybrid scenario in which the duration of SI benefits reduces less than one-to-one with the duration of the delay, the same reasoning as illustrated above still holds with the expectation of an increased cost of delay. As a result of this, the requirements for the delay to be beneficial become more stringent.[18]

[18]Results using a 0.5 reduction in SI benefit receipt are available upon request.

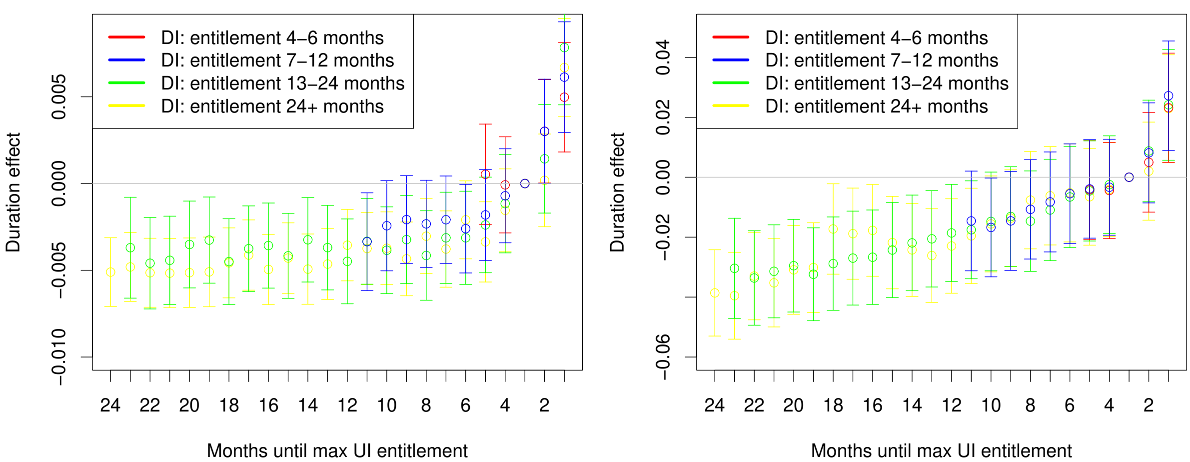
4.A.3 Robustness

In the baseline model, we group workers with maximum entitlement periods in order to separately identify and estimate elapsed and residual time effects in UI and maximum entitlement effects. Such parameter restrictions do not hold for elapsed and residual time effects. This section examines whether the results are also robust to alternative specifications. We examine the robustness of the spike in SI inflow and of the compositional differences. For compositional differences, we focus on the estimates for the cohorts entering into SI.[19]

As a first robustness test, we estimate the spike in exits seperately for groups with different DI entitlement length. As shown in Figure 4.A.3, the spike is similar for all groups. Table 4.A.2 shows the estimates for the various other robustness specifications. As a second test, we alter the definition of the pre-spike cohort. While the observed data shows a clear spike in the last month of UI benefit entitlement, several definitions could be used for the pre-spike cohort. In the baseline specification, the pre-spike cohort consists of workers receiving UI benefits three months before their maximum UI entitlement. Extending the pre-spike cohort to cohorts receiving UI up to six months of residual entitlement yields similar results, as shown in Table 4.A.2. This indicates that cohorts entering into SI in earlier months are in fact very similar to those entering into SI in the three months before their maximum UI entitlement. Alternatively, we could expand the spike cohort by adding one additional month. However, as also seen in the observed data,

Figure 4.A.3: Duration effects for exits into SI (left) and employment (right) across maximum UI entitlement groups, without control variables

(a) Exits into SI (b) Exits into employment



[19] The results for compositional differences of cohorts being sick for at least m months with $1 < m < 24$ are very similar.

the spike is most pronounced in the very last month. Including the cohort receiving UI two months before their maximum UI entitlement attenuates any differences we observe (not shown in the table).

As a third test, we impose a (smooth) polynomial specification for the impact of duration on UI. In our baseline specification, we model the impact of duration on UI as flexibly as possible by using fixed effects for each month of duration. With two polynomials, the resulting estimates are similar to those of the baseline specification.

Fourth, we allow for a more flexible specification for the impact of the maximum UI duration of workers. In the baseline specification, we include dummies for durations between 3 and 6 months, 6 and 12 months, 12 and 24 months, and more than 24 months.[20] A fully flexible specification for maximum UI duration is not possible, given that the elapsed duration on UI plus the residual UI entitlement equals the maximum UI entitlement. The most flexible specification of maximum UI entitlement groups is therefore conducted with two-month intervals. This does not affect any of the estimated coefficients, indicating that heterogeneity due to variation in maximum UI durations is limited.

Finally, we turn to the main parameter of interest, which is the residual duration on UI. In our baseline specification, we again allow for the most flexible specification by including dummies for every residual duration. As shown in Figure 4.1, the estimated duration effects are fairly constant up until five months of residual duration. At this point, the estimated residual duration effects increase, with a clear spike in the last month. If we impose a semi-parametric specification on the residual duration, a polynomial of degree two in addition to a dummy for the spike-cohort, the estimated spike in inflow is largely unaffected. The reason for this is that the polynomial is fitted to the flat left tail of the residual duration. Using a higher-order polynomial does however impact some of the estimated compositional differences. In particular, the estimates for age, migration background, earnings, and fixed-term contract change significantly. For these outcomes, the left tail is not constant and highly susceptive to (small) non-linearities. The fitted value of the polynomial at the spike is therefore very different from the observed value, resulting in spike estimates that significantly differ from the baseline specification estimate. Given that the baseline specification is more flexible, we however believe that the baseline estimates are also more credible.

[20] These groups correspond to the categorization used by the Dutch Employee Insurance Agency.

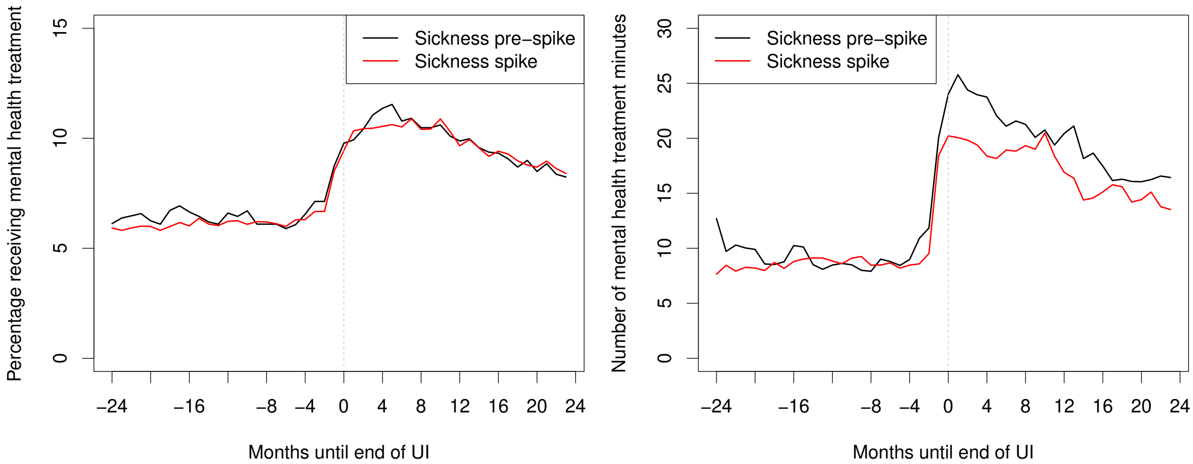
Table 4.A.2: Estimation results of baseline model and robustness analyses

	Risk rates		Demographics			Pre-UI outcomes			Health care utilization		
Specification	**SI inflow**	**DI awards**	**Age**	**Male**	**Native**	**Hours worked**	**Earnings[a]**	**Fixed term**	**MH[b] treatment**	**MH minutes**	**Total spending**
Baseline model	0.66%**	0.07%**	-1.34**	-1.67%	-6.51%**	-2.65*	-5.58%*	4.23%**	0.29%	0.00	474.87
	(0.04)	(0.01)	(0.12)	(1.12)	(0.99)	(1.25)	(2.63)	(1.08)	(0.58)	(2.20)	(256.32)
Pre-spike of 6-3 months	0.66%**	0.07%**	-1.78**	-1.64%	-6.21%**	-2.92**	-6.78%**	3.76%**	0.94%*	0.85	-578.12*
	(0.04)	(0.01)	(0.10)	(0.90)	(0.79)	(1.00)	(2.10)	(0.86)	(0.47)	(1.76)	(205.20)
Polynomial elapsed	0.76%**	0.07%**	-1.30**	-1.85% UI spell	-6.76%**	-2.81*	-5.70%*	4.01%**	0.28%	-0.29	-542.04
	(0.04)	(0.01)	(0.12)	(1.13)	(0.99)	(1.25)	(2.63)	(1.08)	(0.58)	(2.20)	(256.02)
Flexible maximum UI entitlement groups	0.66%**	0.07%**	-1.39**	-2.30%*	-6.82%**	-2.73*	-5.15%	4.21%**	0.52%	0.22	-508.98*
	(0.04)	(0.90)	(0.80)	(1.13)	(1.00)	(1.26)	(2.63)	(1.08)	(0.58)	(2.21)	(257.07
Polynomial residual UI spell	0.63%**	0.05%**	0.41**	-2.20%*	-3.63%**	-2.38*	-2.52%	-0.07%	0.61%	0.32	-219.22
	(0.03)	(0.01)	(0.10)	(0.90)	(0.80)	(1.00)	(2.10)	(0.86)	(0.47)	(1.77)	(205.22)

*/** Difference is significant at a 5%/1% significance level.
[a] The outcome measure for earnings is log(Earnings+1). The estimated coefficient is approximately equal to the percentage difference in earnings between the pre-spike and spike cohorts.
[b]: MH = Mental Health

Figure 4.A.4: Differences in the evolution of mental health treatments of unconditional pre-spike and spike cohorts entering SI: incidence and minutes

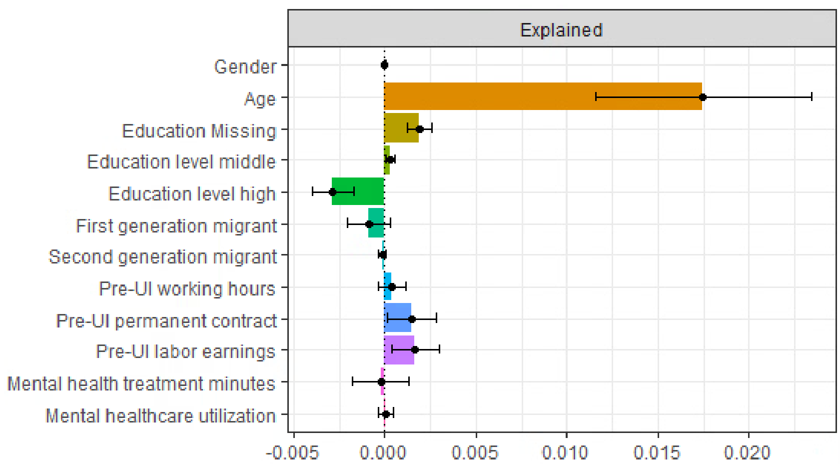
(a) Incidence of mental health treatments (b) Mental health treatment minutes



4.A.4 Synthesis

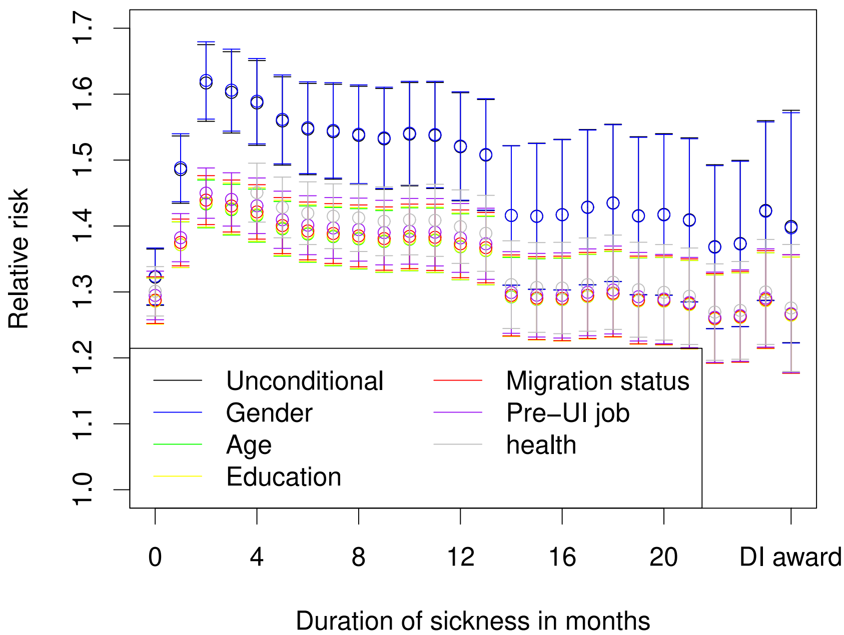
Figure 4.A.5 shows the Oaxaca-Blinder decomposition for the spike into DI. Given substantial outflow during the sickness period, the spike in inflow in DI is substantially smaller than the spike in SI, at 0.04 percentage points. The results for the spike in DI are comparable to the spike in SI. Compositional differences explain approximately 0.01 percentage points or 28% of the spike. Age is again the largest contributor.

Figure 4.A.5: Oaxaca-Blinder decomposition of the spike in DI inflow



As an alternative approach the quantify the impact of composition differences on the spike in inflow into SI, we re-estimate the relative risk of flowing into SI (Equation (4.1)) while sequentially adding control variables. The resulting estimates are shown in Figure 4.A.6. The inclusion of age explains a substantial part of the relative risk while the other compositional differences we observe do not explain the difference in inflow into and subsequent outflow out of SI. After controlling for age, both the risk of entering SI and the risk of being awarded DI benefits are approximately 30% larger for the spike cohort. So conditional on observables, workers in the spike cohort entering SI are equally likely to be awarded DI benefits, relative to their counterparts in the pre-spike cohort. These results are in line with those of the Oaxaca-Blinder decomposition.

Figure 4.A.6: Relative risk of reaching specified SI benefit spells for pre-spike and spike worker cohorts; controls sequentially included (95% confidence intervals)



Chapter **5**

Do disability benefits hinder work resumption after recovery?

5.1 Introduction

One of the largest social insurance schemes in developed economies is Disability Insurance (DI) OECD (2019). In the Netherlands, for example, approximately 9% of the working population received disability benefits in 2017 whereas respectively 4% and 4.5% of the working population received unemployment benefits or social assistance benefits CBS (2019). Total spending on disability benefits amounted to 1.5% of GDP making it the most sizable social insurance scheme in terms of expenditure. In most OECD countries, a substantial share of these expenditures is intended for temporary disabilities Pettersson-Lidbom & Thoursie (2013), of which mental health problems are one of the main causes. Albeit that a large share of those with mental health problems are expected to recover Korpi (2001), their work resumption rate is typically low Claussen et al. (1993); Autor & Duggan (2006).

Given that many disabled workers with mental health conditions are deemed to have residual earnings capacity and/or health conditions that are temporary, a pertinent question is how DI benefits and work incentives should be optimally designed. In the US context, Maestas (2019) points towards the introduction of partial and temporary DI benefits as promising reforms for the public DI system. Still, there is limited understanding of how workers that recover from impairments respond to changes in work incentives. Research on the incentive effects of DI benefits has predominantly focused on policies targeting

[1]This chapter is based on Koning et al. (2022a). The online appendix of this chapter can be found at www.rogerprudon.com/research

inflow or the use of schemes.[2] At the same time, studies that do consider the outflow of DI largely ignore the potential importance and consequences of health improvements (Weathers & Hemmeter, 2011; Campolieti & Riddell, 2012; Kostol & Mogstad, 2014; Koning & van Sonsbeek, 2017).

As one of the few studies that consider DI outflow due to health improvements, this chapter assesses whether disability benefits create disincentives for returning to work once health improves. Comparing DI applicants with degrees of disability below and above the threshold for (partial) DI benefits, we estimate the effect of receiving benefits on the labor supply response to a positive health shock that is measured by the end of mental health treatment. Considering all DI applicants in the Netherlands since 2006, we show that health improvements indeed coincide with an increase in labor supply and that disincentives from receiving benefits do matter: awarded applicants with partial disability benefits show weaker labor supply responses than those without partial benefits.

The main reason for the lack of research on the interplay of disability benefits, health recovery and labor supply is the absence of reliable data on the dynamics of individuals' health. We address this problem by linking three sources of Dutch administrative data covering the entire population of DI applicants since 2006 (over 600,000 individuals): (i) DI application records (including detailed assessment outcomes), (ii) monthly administrative records on employment and social insurance receipt and (iii) administrative records describing mental health treatments. With the DI application data, we are able to identify applicants above and below the disability benefits cutoff in terms of their degree of disability. Inherent with our research approach, we thus focus on disability applicants at the lower end of the disability severity distribution (with a loss of earnings capacity of at most 50%) and for whom the likelihood of recovering is assessed as being high at the moment of application. For those with mental health problems – of whom the majority suffer from mood, anxiety or personality disorders – we obtain a plausible measurement of health improvement by considering the end date of mental health treatment. While certainly not being indicative of full recovery for all workers, we interpret the end of a mental health treatment trajectory as a substantial improvement in health. This is in line with general medical literature, reporting substantial recovery rates of mental illnesses due to treatments Richards (2011); Curry et al. (2011); Leichsenring & Leibing (2003); Bandelow et al. (2017). We compare labor supply responses around the end date of treatment for those that receive disability benefits with those that do not, yielding estimates of the employment disincentive effects of benefit receipt.

[2]By now there is a vast literature on moral hazard and targeting of DI benefits. Specifically, higher benefits lead to higher inflow rates Gruber (2000); Borghans et al. (2014) and more stringent selection criteria reduce inflow Staubli (2011); Godard et al. (2019).

Our estimation approach constitutes a difference-in-differences (DiD) estimator that compares applicants with and without disability benefits, before and after recovery. Using an event-study specification, we show that employment rates for the two groups follow parallel trends until closely before mental health treatment ends. Knowing that the end of mental health treatment is not a perfect proxy for recovery, our DiD estimator only requires the assumption that it proxies recovery equally well in the two groups.[3] Using detailed data on healthcare trajectories, we show that the probability of recovery is indeed almost identical for both groups.

We observe that around the time of recovery, the employment rates start to diverge, as those without disability benefits start working at a higher rate than those with disability benefits. The disincentives for work resumption are substantial, amounting to a negative impact of disability benefits of 5.3 percentage points on employment, relative to baseline employment of around 30%. We interpret this as a large impact since the pre-recovery difference between the groups is small and our proxy measures only partial recovery. Furthermore, the employment response to recovery for individuals without DI benefits is approximately 10 percentage points, indicating that benefits absorb approximately half of the response. For hours worked, the disincentives appear even stronger: they eliminate almost 75% of the recovery response. Using an alternative proxy for health improvement that is based on significant drops in annual healthcare expenditures, we find very similar results that extend to *physical* health improvements. Our findings are robust against a series of alternative specifications, including imposing different 'donuts' around recovery to deal with the imperfect measurement of the exact timing of recovery.

Individuals with DI benefits have a larger assessed loss of earnings capacity than those without DI benefits. Albeit the treatment and control samples are fairly similar in many aspects, the response to recovery might depend on the initial level of health. We therefore perform placebo tests comparing two groups that differ in assessed loss of earnings but both do not receive DI benefits (or both do receive DI benefits). We find that these groups respond similarly to recovery: none of the event-study estimates is statistically different from zero for these groups. This supports the interpretation that indeed the DI benefits cause the difference in recovery response.

For a broader perspective on benefit disincentive effects, we next benchmark our estimates against predictions from a structural labor supply model. We estimate structural model parameters using information from pre-disability labor supply and the assessed remaining earnings capacity of workers. With this information, we validate the model

[3] A similar argument holds for the issue of reverse causality: in some cases, it may be employment that *causes* health improvement (Browning et al. (2006); Sullivan & Von Wachter (2009); Morris & Cook (1991); Kuhn et al. (2009); Schmitz (2011); Salm (2009)). Again, any biasing effects are mitigated as long as reverse causality is equally strong in both groups.

by comparing the predicted labor supply after a disability shock with the observed labor supply after a DI application. We then use the calibrated model to simulate labor supply responses to health recovery. Defining full recovery as a situation where earnings capacity *and* disutility from working return to their pre-application level, we find benefit disincentive effects of disability benefits equal to around 15 percentage points, which is approximately one-quarter of the full recovery response for individuals without benefits. This estimate should be considered as an upper bound of the disincentive effect. Since the end of the mental health treatment does not necessarily coincide with full recovery in all cases, it is not surprising to see considerably smaller reduced-form effects.

Our findings contribute to the broad literature on the effects of financial incentives of DI schemes on employment. So far, this literature has mainly focused on the disincentive effects of disability benefits at the application stage. DI beneficiaries are less likely to be employed than those whose DI application has been rejected (Bound, 1989; Chen & Van der Klaauw, 2008; Von Wachter et al., 2011; Maestas et al., 2013; French & Song, 2014) and work resumption rates of DI beneficiaries also depend on financial incentives (Weathers & Hemmeter, 2011; Campolieti & Riddell, 2012; Kostol & Mogstad, 2014; Koning & van Sonsbeek, 2017). To understand the lack of outflow from DI schemes and work resumption, it is essential to consider work resumption rates of DI applicants once their health improves. Closest to this, some studies consider the effect of changes in qualifying conditions for ongoing DI benefits. In this respect, Moore (2015) exploits changes in the qualifying conditions for DI benefits and Garcia Mandico et al. (2018) evaluate the effects of reassessments of DI benefit claimants.

The second relevant strand of related literature examines the relationship between health and labor supply. This literature either studies the effect of negative health shocks on labor supply or the effect of employment on health outcomes. Using self-assessed health (García Gómez & López Nicolás, 2006; Lindelow & Wagstaff, 2005), road injuries Dano (2005) and acute unscheduled hospitalizations (García Gómez et al., 2013; Lindeboom et al., 2016) a causal relationship has been established between negative health shocks and labor supply. These studies find that employment rates drop by five to seven percentage points after a negative health shock.

Lastly, we add to the existing medical literature on the effects of mental health treatments.[4] The effects of treatment on various health measures and on recovery rates are well established for the main mental health diagnoses. This includes mood disorders (Richards, 2011; Curry et al., 2011), personality disorders (Leichsenring & Leibing, 2003) and anxiety disorders (Bandelow et al., 2017)). Treatment is found to be effective for the

[4]Section 5.4 also discusses this literature, so as to provide a deeper discussion on why the end of treatment can be interpreted as a valid proxy for health improvements.

majority of mental health problems and return-to-work rates are relatively high (> 90%) for individuals with mental health problems (de Vries et al., 2018). With limited access to reliable data on positive health shocks, the causal impact of health *improvements* for individuals receiving DI benefits – as considered in this chapter – has not been examined so far. We contribute to this literature by incorporating disability insurance in the interplay between health and labor and by examining positive health shocks.

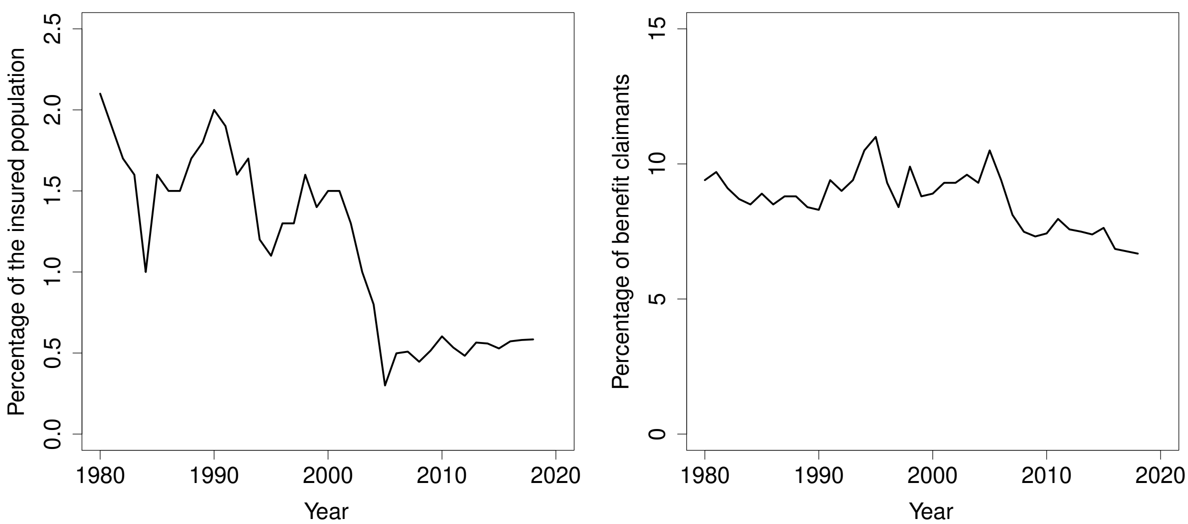
The remainder of the chapter is organized as follows. Section 5.2 illustrates the institutional background of the Dutch disability insurance system and Section 5.3 describes the data. Section 5.4 provides a description of the DiD estimator, presents results and discusses robustness checks. In Section 5.5 we compare our results to simulations from an estimated structural labor supply model. Section 5.6 concludes.

5.2 Disability insurance in the Netherlands

In the Netherlands, the administration of DI benefits is managed by the Employee Insurance Agency (UWV). The DI system has long been characterized as “out of control”, with approximately 12% of the working population receiving benefits by the turn of the century (Koning & Lindeboom, 2015). Since then, a series of reforms that tightened eligibility conditions and increased screening activities before DI application have led to drastic reductions in annual award rates. To illustrate this, panel (a) of Figure 5.1 shows that inflow rates decreased from about 1.5% of the working population in the 90s to about 0.5% since the most recent reform in 2006. Even though some reforms also aimed at in-

Figure 5.1: Inflow and outflow rates of disability benefit programs (UWV (2012, 2018))

(a) DI inflow (b) DI outflow



creasing work incentives, this drop in inflow rates is accompanied by decreasing outflow rates out of DI that are shown in panel (b).[5]

Since 2006, DI applications can be filed after two years of illness in the Netherlands.[6] Using a selection of feasible jobs with wage earnings estimates, medical and vocational experts of UWV assess the remaining earnings capacity of applicants. Based on the pre-application earnings and the assessed earnings capacity, the degree of disability is determined as follows:

$$\text{Degree of disability} = (1 - \frac{\text{Remaining earnings capacity}}{\text{Pre-application earnings}}) * 100\% \tag{5.1}$$

Individuals are assigned to one of the following brackets for the degree of disability: 0–35%, 35–45%, 45–55%, 55–65%, 65–80% and 80–100%. Disability benefits are awarded if the assessed degree of disability exceeds 35%. Benefits are based on the midpoint of the assigned disability interval. As an illustration, consider an individual with pre-application earnings of €3,000 per month. If the remaining work capacity is 16 hours (i.e., 40% of full-time work) at a wage of €2,000 per month on a full-time basis, the remaining earnings capacity is set at €800. The resulting degree of disability is then 73.3%, implying that the relevant degree of disability bracket is between 65% and 80%.

If awarded benefits, benefit conditions differ between the so-called "wage-related period" and the "wage continuation period" UWV (2019b). The wage-related period applies to anyone who worked at least 26 weeks within the 36 weeks prior to falling ill. Benefits amount to 70% of the difference between pre-application earnings and current earnings.[7] The duration of the wage-related period increases with one month for every year worked since the age of 18 and is capped at 24 months.[8]

Once the wage-related period ends, the wage continuation period commences. The benefit level then depends on the utilization of the remaining earnings capacity (see Figure 5.2 for income as a function of labor earnings for an individual with an assessed degree of disability of 50%). If workers earn less than 50% of their remaining earning capacity, DI benefits equal approximately 30% of the minimum wage (500 euros per month). If

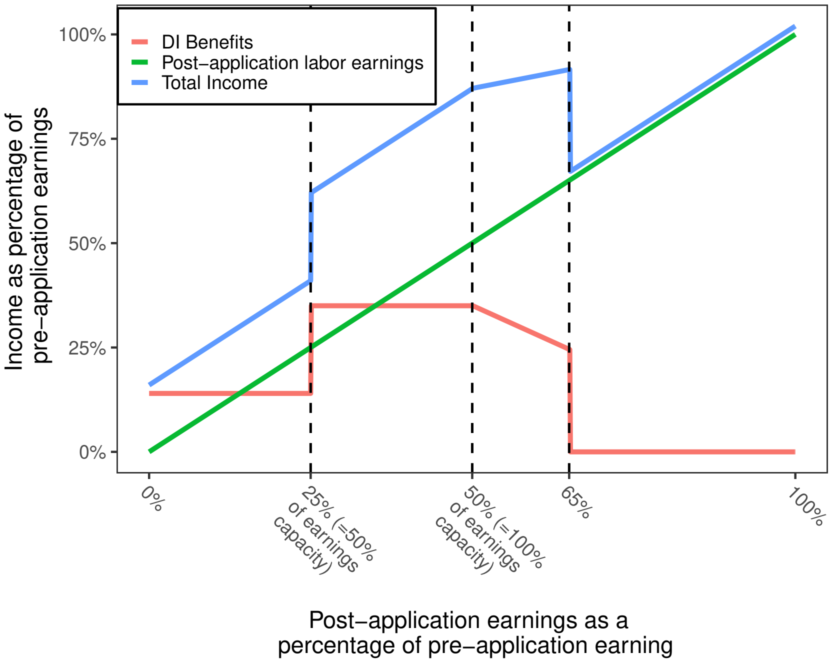
[5]See Koning & Lindeboom (2015) for a more detailed discussion on the reforms prior to 2006.

[6]In Online Appendix Section B.1 we provide a detailed description of the disability insurance process from the start of the illness until the actual application.

[7]Benefits are taxed away during the wage-related period in a similar fashion as during the continuation period (at 70%). However, eligibility for benefits cannot be terminated during the wage-related period. If someone earns approximately 100% of his/her pre-application earnings, benefits are effectively reduced to zero. If someone earns more than 65% of the pre-application earnings, eligibility would be terminated once the wage-related period ends.

[8]The setup of the disability system is slightly different for individuals with an assessed degree of disability above 80% whose disability is deemed to be permanent. They receive disability benefits amounting to 75% of their pre-application earnings and no re-assessments are performed. Individuals who don't meet the criteria for the wage-related period, immediately start in the wage continuation period

Figure 5.2: Income from labor earnings and benefits, as a function of post-application earnings (for an individual with an assessed degree of disability of 50%)



someone uses at least 50% of their remaining earnings capacity, the amount of benefits is tied to the pre-application earnings and equal 70% of the loss in earnings capacity (on average 800 euros per month for our sample). If current earnings exceed the remaining earnings capacity, benefits are adjusted accordingly. DI benefits can be terminated if the remaining earnings capacity has increased over 65% of previous earnings. This can be due to increases in actual earnings or medical reassessments. These reassessments can be requested by UWV (because of administered changes in wage earnings), the former employer, or by the recipient. Reassessments due to suspected changes in health are scarce compared to those due to changes in wage earnings UWV (2019a). Workers with DI benefits are not eligible for additional unemployment benefits but might qualify for social assistance (approximately 70% of the minimum wage).

The DI system creates various (dis)incentives to work that may differ in their impact before and after recovery. In Online Appendix Section B.2 we illustrate these incentives for an individual with an assessed degree of disability of 50%.[9] When defining recovery as a situation where the earnings capacity of workers is fully restored to its pre-application level, we expect workers without benefits to resume their initial working hours. For workers with partial benefits, however, the potential impact of substitution and income

[9]Note that the structural model in Section 5.5 also provides further insight into relevant labor supply incentives for partially disabled workers with and without DI benefits.

effects may well increase after recovery and therefore discourage (full) work resumption. Without employment, partial benefits amount to about €500 per month at minimum and €1,000 at maximum. At the same time, the assessed earnings capacity also induces a "cash cliff", since wage earnings exceeding this amount imply the (full) loss of partial DI benefits. This incentive to earn at most 65% of pre-application earnings may also become binding after recovery when hourly wages increase, causing a decrease in working hours. In a similar fashion, the incentive to earn at least 50% of the remaining earnings capacity may relax after recovery, so fewer working hours are needed to exceed this threshold. Overall, we therefore expect benefits to reduce the employment response to recovery. This holds both for intensive and extensive margins. In the empirical analysis, we will estimate the combined income and substitution effect of DI benefits. We refer to these combined effects as the disincentive effects, both on earnings and employment.[10]

5.3 Data sources and sample selection

We link three data sources to analyze the recovery of DI recipients: disability insurance application data (provided by UWV), monthly income data (Statistics Netherlands) and healthcare utilization data (Statistics Netherlands). All data are administrative and cover all Dutch citizens. The following subsections discuss the various data sources and the sample selection.

5.3.1 Disability insurance application data

The disability insurance application data comprises all applications between January 2006 and June 2017. The data contain all information for both awarded and rejected applicants that is needed to determine their earnings capacities and their degrees of disability. It includes the pre-application hourly wage and number of hours worked, as well as the post-application potential hourly wage and number of working hours.[11] Additionally, it includes the timing and the outcome of the award decision by UWV.

Several health-related variables are included in the DI application data. The first group of variables concerns the medical diagnoses of the applicants. The diagnoses are

[10]The disincentive effects we estimate are a weighted average between the disincentive effects in the wage-related and continuation period. Unfortunately, we do not observe whether an individual is in the wage-related period. The maximum duration of the wage-related period is approximately three years, so we do know that individuals who recover more than three years after their application are in the continuation period. The latter is true for 30% of the sample, whereas the other 70% could be either in the wage-related period or in the continuation period.

[11]Some individuals are deemed to be (fully) incapable for work by a medical examiner, based solely on medical grounds. The assessment of the remaining earnings capacity is not conducted for these individuals and their degree of disability is not stated in the application information. Unfortunately, we lack information that would enable us to distinguish these applicants from those that terminate their application before the actual assessment.

either classified by so-called "CAS-codes" or by a categorization created by UWV. The CAS-codes are used by health and safety doctors responsible for the reintegration process of long-term absent workers and are more detailed than the UWV codes.[12] The CAS-codes are available for 80% of all applications, whereas the UWV codes are available for 98% of all applications. The analysis will combine both types of diagnosis information.

The application data also contains information on the type and number of functional limitations of applicants, as assessed by a medical assessor of UWV. These limitations range from physical limitations, such as neck movement and use of hands, to limitations such as cognitive functions and work stress. There are a total of 17 limitation groups, and the severity of every limitation can range from zero, implying no limitation, to 7, implying a severe limitation. The functional limitations are used when determining the potential hourly wage an applicant could earn. The last health-related variable concerns the probability of improvement, as assessed by UWV ("reasonable to good", "small" or "non-existent").

5.3.2 Income data

Income data is provided by Statistics Netherlands. We have access to information on all employment contracts in the Netherlands between 2006 and 2018, including an individual's monthly earnings, hours worked and monthly employment status. We combine multiple employment contracts registered at one and the same month, so as to obtain the total earnings and total number of hours worked. The employment indicator indicates whether an individual worked for at least one hour in a specific month. We enrich the income data with administrative records from Statistics Netherlands on the year of birth, gender, nationality and level of education. Using household identifiers we are also able to link individuals to their partners, which is important to determine welfare eligibility.

5.3.3 Healthcare data

Throughout our analysis, two separate data sources on medical treatments are used. The first source concerns data on mental healthcare treatments, which are derived from so-called "Diagnosis Treatment Trajectories" (DBCs) that are used as payment units for complete medical treatments. DBCs comprise all treatments that are deemed necessary to alleviate or solve health problems. We observe all DBCs regarding mental health between 2011 and 2016. Mental health problems in the sample can be considered as severe, with an average cost of treatment of approximately €5,500 and resembling roughly 140 hours

[12] CAS-codes consist of a diagnosis group letter, e.g. "P" for psychological diagnoses, and a three-digit number indicating the specific diagnosis (32 in total). Applicants can have at most three CAS codes and three UWV diagnoses group codes.

of treatment. DBC entries include the start and end date of treatment. For individuals with multiple treatment trajectories, we use the earliest start date and the latest end date. Unfortunately, information on whether an individual actually recovered because of the treatment is unavailable. The end of treatment should therefore be considered as a proxy for recovery (we discuss the consequences of using a proxy in Section 5.4). As we will argue later on in more detail, it is important to note that the medical literature has found treatment for these disorders to be effective, with recovery probabilities above 50%. Finally, it should be noted that the end of mental health treatment is not determined by fixed treatment durations and has a high level of variability in our data, as shown in Online Appendix Figure B.2.

The second data source concerns the yearly healthcare expenditures covered by basic health insurance for the years 2009 until 2017. Basic health insurance is compulsory in the Netherlands and covers the vast majority of all healthcare. The data shows the spending on various subcategories. We construct measures of mental healthcare expenditures and non-mental healthcare expenditures (see Online Appendix Section B.3). Using the healthcare expenditures in the pre-application waiting period as a baseline, we create a second proxy for recovery based on a substantial drop in healthcare expenditures. Given that the expenditure data is only available on an annual basis, there is more measurement error in the proxy compared to the proxy based on the DBC data.

5.3.4 Sample selection

Merging the three data sources yields a sample of disability insurance applicants for whom the application information, mental health information, and employment history are observed. A selection was made to make the sample suitable for analysis. Table 5.1 illustrates the various sample selection steps that are taken.

As a starting point, we consider all DI applications between January 2006 and July 2017, for which the disability is assessed as being temporary and partial. To ensure that the mental health problems are sufficiently severe to affect the employment status, only those individuals that applied for DI benefits on the basis of mental health problems are included. This yields a sample of 71,854 individuals.

Since the mental healthcare data is only available until 2016, individuals are selected for whom the end of treatment occurred before January 2016; this ensures that a new mental healthcare trajectory does not start shortly after the observed end of treatment.[13] We exclude individuals for whom the end of treatment occurs before the application date.

[13] Given that recurrence/relapses are common for many mental health diagnoses, we alternatively select individuals for whom the end of treatment occurred before January 2014. This ensures that relapses do not occur for at least three years. As shown in Section 5.4, this does not change our results.

Table 5.1: Sample selection criteria

Inclusion criteria	Remaining sample
Temporarily and partially disabled with mental health diagnosis	71,854
Recovered before 01-01-2016	28,046
Recovered after application	15,578
Comparable degree of disability (20%–30% and 40%–50%)	5,003

Last, we focus on DI applicants with an assessed degree of disability of 20–30% (not receiving benefits) and those with an assessed degree of disability of 40–50% (receiving benefits). Our sample therefore consists of "treated" (with benefits) and "untreated" workers that have similar degrees of disability; we motivate this choice in detail in the next section where we present our empirical model. The final sample consists of 5,003 DI applicants for whom we have a proxy for mental health recovery. The majority of these individuals suffer from mood (40%), personality (17%) or anxiety disorders (12%). Considering these types of disorders, we stress that *recovery* could be actual medical recovery from the disorder (for example in the case of depression), but it could also be partial recovery that for example restores to some extent the individual's ability to work.[14]

In Table 5.2 we present descriptive statistics and we show tests for equality of means between the groups with and without benefits. The groups have similar gender and nationality statistics. However, the group without benefits is on average slightly younger and lower educated. The treatment and control groups are very similar in terms of (mental) health. They suffer from similar disorders, and healthcare expenditures are comparable.[15] As expected, there are significant differences in the DI application variables. The pre-application hourly wage and number of working hours, the number of functional limitations and the degree of disability are all lower for those without benefits. Lastly, the groups have a similar assessed probability of health improvement and comparable healthcare expenditures in the year of their DI application.[16]

[14]A comparison between the selected sample and the sample of non-selected DI applicants can be found in Online Appendix Table B.2. Mean characteristics are comparable to the non-selected sample, but – by construction – the non-selected sample contains a more diverse set of individuals with degrees of disability ranging from 0% to 80%.

[15]See Online Appendix Tables B.3, B.4 and B.5 for descriptives of groups with specific disorders.

[16]This is all confirmed in one of the robustness tests to our model, where we show that the inclusion of individual control variables has a negligible impact on our findings.

Table 5.2: Descriptive statistics of DiD treatment and control groups

	Degree of disability:		
	20–30%	40–50%	P-val[a]
Demographics:			
Age	47.6	49.0	0.000
Female	53.8%	54.9%	0.474
Dutch native	65.5%	69.2%	0.010
Education:			
Unknown	8.7%	13.8%	0.000
Low	26.9%	23.9%	0.022
Middle	43.7%	36.7%	0.000
High	20.7%	25.7%	0.000
(Mental) health:			
Treatment duration[b]	32.9	35.3	0.000
Mood disorder[c]	39.0%	41.7%	0.068
Anxiety disorder[c]	12.9%	10.5%	0.011
Personality disorder[c]	16.6%	16.8%	0.837
Treatment termination instigated by:			
Patient	9.0%	8.2%	0.332
Therapist	4.5%	3.7%	0.157
Joint decision	34.9%	33.5%	0.317
Other	51.6%	54.6%	0.041
Mental healthcare expenditures[d]	€3,037	€3,542	0.136
Physical healthcare expenditures[d]	€1,428	€1,612	0.095
DI application:			
Pre-application hourly wage	€15.70	€17.71	0.000
Pre-application hours	33.1	34.0	0.003
Earnings capacity: hourly wage	€12.21	€12.10	0.023
Earnings capacity: hours	32.4	27.1	0.000
Number of functional limitations	9.7	11.6	0.000
Degree of disability	25.2%	44.6%	0.000
Assessed probability of health improvement:			
NA	27.8%	25.8%	0.130
Reasonable to good	64.1%	67.9%	0.007
Small	7.8%	5.9%	0.011
Non-existent	0.3%	0.4%	0.617
Observations	3,346	1,656	

[a]P-value of two-sample t-test for equality of means; [b]Duration of the mental health treatment in months; [c] Percentage with specific mental health disorder; [d]Expenditures in euros in the year of DI application.

5.4 Empirical analysis

5.4.1 Model specification

We use a difference-in-differences (DiD) model with an event-study specification to estimate the impact of disability benefits on the employment response to mental health improvement. We compare the response to recovery while receiving DI benefits (the treatment group) to the response to recovery in the absence of DI benefits (the control group). Under the assumption that the control group and the treatment group follow parallel trends in the outcome in the absence of treatment, any divergence between the groups can be attributed to the causal impact of DI benefits on the response to recovery.

To select a control and treatment group that is most likely to exhibit parallel trends, we consider individuals with a degree of disability relatively close to the benefit threshold of a degree of disability that is equal to 35%. To balance similarity in characteristics with sufficient sample size, we include all applicants with a degree of disability between 20% and 50%. Assessors may have some leeway to affect the degree of disability, and one may worry about potential manipulation around the threshold. If manipulation is partly based on the probability to recover, individuals just below and just above the threshold may have different recovery rates. Since there is a discontinuity in the density around the threshold – see Figure 5.A.1 in the appendix for this – some of these concerns appear to be justified and we therefore use a “donut”-design around the threshold and exclude all applicants with degree of disability between 30% and 40%. In effect, this means we compare applicants with a degree of disability of 20–30% (the control group) to applicants with a degree of disability of 40–50% (the treatment group).[17] The general event-study specification is as follows:

$$E_{it} = \alpha_{DI} + \alpha_t + \sum_{l=-T+1}^{-1} \beta_l DI_i\, I_{t=l} + \sum_{l=0}^{T} \beta_l DI_i\, I_{t=l} + \theta X_{it} + \varepsilon_{it} \tag{5.1}$$

in which i subscripts the individual and t denotes the time relative to the month of recovery (with $t = 0$ being the month of recovery). E_{it} is an employment outcome (employment or hours worked), DI_i is an indicator for receiving disability benefits and $I_{t=l}$ indicates whether an observation is in month l relative to recovery. α_t captures the evolution over time for individuals without DI benefits while β_l, the parameters of interest, capture deviations over time for individuals with DI benefits prior to and after recovery. We use a time window of 48 months before and 48 months after recovery in the baseline

[17] As a robustness check, we also consider larger bandwidths and perform the analysis without the donut around the threshold, using the 25-35% group as control group and the 35-45% as treatment group. Results are similar, as we will show in Table 5.2.

specification. Individual characteristics X_{it} contain age, gender, nationality, education level and calendar-month fixed effects.[18,19]

5.4.2 End of treatment as a proxy for recovery

A specific feature of our analysis concerns the use of the end of mental health treatments as a proxy for health improvements. The end of treatment is an effective proxy if these treatments causally increase the probability of recovery. The medical literature provides considerable evidence in this direction, see e.g. Curry et al. (2011), Leichsenring & Leibing (2003), Bandelow et al. (2017) and de Vries et al. (2018). Taking a broader perspective, these studies are complemented with evidence that an important fraction of individuals recover from mental diseases (Norder et al., 2015; Roelen et al., 2012).

In line with this, medical assessors from the Employee Insurance Agency (UWV) consider the expected length of mental treatments as one of the key determinants for the severity and permanence of impairments. Specifically, protocols for the claims assessment of applicants with mental impairments prescribe that information obtained from medical doctors should be considered by the medical assessors from UWV.[20] At this point, it is important to stress that the guidelines do not necessarily presume that mental health treatments are effective for all relevant DI benefit applicants. For those applicants with more intensive treatments, the end of treatment may also demarcate the end of a period wherein the available time and effort to resume work was limited anyway. Such "lock-in" effects may thus go together with increases in work resumption at the end of treatments, also for workers for whom the overall treatment was not effective.

Albeit most likely that the end of treatment is an accurate proxy for recovery for many workers, our analysis acknowledges that it is not a measure of recovery for *all* workers, nor is it necessarily a measure of *full* recovery. Specifically, estimates of recovery effects need to be interpreted as lower bounds or "Intention-To-Treat" (ITT) effects.[21] For consistent

[18]Given the potential persistence in the outcome variables, we cluster standard errors by individuals Hausman & Rapson (2018). As a robustness test, two additional methods will be used to account for the serial dependence: (i) the model will also be estimated on mean levels and (ii) the analyses will be conducted non-parametrically by not including any control variables in the regressions and thereby comparing the differences in unconditional means. The latter should circumvent any time-series characteristic issues such as heterogeneity Lechner (2011).

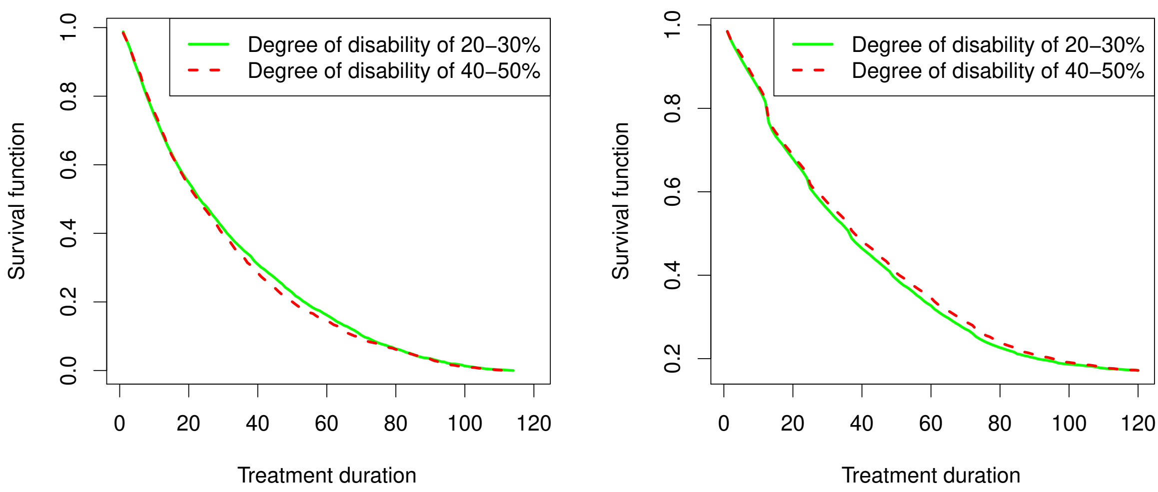
[19] A recent literature has shown that two-way fixed effects estimators are biased in case of staggered treatment implementation and dynamic treatment effects Goodman-Bacon (2021); Callaway & Sant'Anna (2021); Borusyak et al. (2021). In our setting, we compare a single treatment group (those with benefits) to a single control group that is never treated (those without benefits) and thus these concerns do not apply (see for example Baker et al. (2021))

[20]The medical assessor should consider (i) the expected course of the disease in the absence period, (ii) the actual limitations, and (iii) current mental health treatments (Gezondheidsraad, 2016).

[21]If a share of ρ recovers in both the treatment and control group, the actual disincentive effects from disability benefits on the labor response to mental health recovery equals $\frac{\beta_2}{\rho}$.

Figure 5.1: Survival functions of mental health treatment for DI applicants with degree of disability of 20–30% vs. 40–50%.

(a) Relative to time of application (b) Relative to start of health treatment



estimation of differences in recovery response (β_l) we require the assumption that the proportion of individuals that recovers is equal in the treatment and control group. To reassure that this assumption is likely to hold, we now show that (i) the DI benefit award decision does not affect the subsequent duration of the mental health treatment and (ii) the underlying rate of recovery does not differ between the control and treatment group.

First, one might worry that the outcome of the DI application award decision affects the subsequent duration of the mental health treatment. If so, the end of a mental health treatment may have implications that differ between the groups with and without benefits. To assess whether this is the case, panel (a) in Figure 5.1 shows the Kaplan-Meier estimates of the mental health treatment duration since the time of the DI application. It is reassuring to see that the survival functions are almost identical, with a log-rank test indicating insignificant group differentials (P-value = 0.70). This suggests that applicants with and without benefits spend equal amounts of time in mental health treatment after their DI application.

Our second concern is that the actual underlying rate of recovery differs between the control and treatment group. Since actual recovery is unobserved, we compare alternative recovery measures to test this. One such measure is the assessed probability of improvement of UWV, which is similar for the treatment and control groups (see Table 5.2). As a second measure, we consider differences in the probability that either the treated individual or the medical doctor decided to terminate the treatment. Even though recovery is not registered, we do observe who initiates the end of the treatment. If underlying rates of recovery would differ between both groups, we would expect to see differences in the distribution of termination types as well. However, this is not the case (see Table 5.2).

Lastly, we can also compare the duration of completed mental health treatment programs. If individuals without DI benefits are more likely to recover, the duration of mental health treatments should be shorter on average for that group. Furthermore, if reverse causality is stronger in one of the groups – meaning more people recover *because of* working – the duration of treatments should be shorter on average for that group. To shed light on this, panel (b) in Figure 5.1 shows the survival functions for mental health treatment for both groups. A log-rank test shows no significant difference. The strong resemblance between health treatment duration for the two groups renders it unlikely that the end of treatment proxies health recovery to a different degree in the groups with and without DI benefits. This strengthens the idea that divergence in employment outcomes between control and treatment group is due to differential responses to recovery.

Combining these results, we have three pieces of evidence that suggest that recovery rates are comparable between the treatment and control group. Nevertheless, ultimately we cannot observe the actual recovery and remain dependent on the assumption that there are no systematic differences between the two groups.

5.4.3 Main estimation results

Inherent with the DiD approach, our estimation strategy relies on parallel trends of control and treatment groups prior to recovery.[22] To gauge the validity of this assumption, Figure 5.2 shows the trends in employment relative to recovery for the treatment and control group.[23] For both groups, labor supply decreases steadily up to the point of recovery, after which the trend reverses. While the negative trend prior to recovery is virtually identical for treatment and control group, the increase after recovery is substantially larger for the group that does not receive DI benefits.

We next turn to the estimation of Equation (5.1) and present the corresponding employment disincentive effects (β_l) in Figure 5.3. The results confirm parallel trends before recovery, with non-significant estimates that are close to zero.[24] Employment rates start diverging approximately six months prior to recovery, suggesting some anticipation.

[22] It should be stressed that the parallel trend assumption does not impose a non-anticipation assumption. As long as recovery is anticipated in a similar way by those with and without benefits, the parallel-trends assumption needed for the identification of the disincentive effect is not violated.

[23] Trends and estimates for the number of working hours are similar (Appendix Figure 5.A.2).

[24] As suggested by Roth (2019), we also consider to what extent deviations from parallel trends would affect our results. Given the long pre-period considered, we are able to reject deviations from parallel trends above –0.005 p.p. with a power of 80%. Using these hypothetical deviations, our post-recovery event-study estimates remain significant. Directly incorporating potential deviations from parallel trends using the HonestDiD method proposed by Rambachan & Roth (2019), indicates that our estimates are robust to deviations from parallel trends.

Figure 5.2: Employment trends relative to mental health recovery for DI applicants with degree of disability of 20–30% vs. 40–50%.

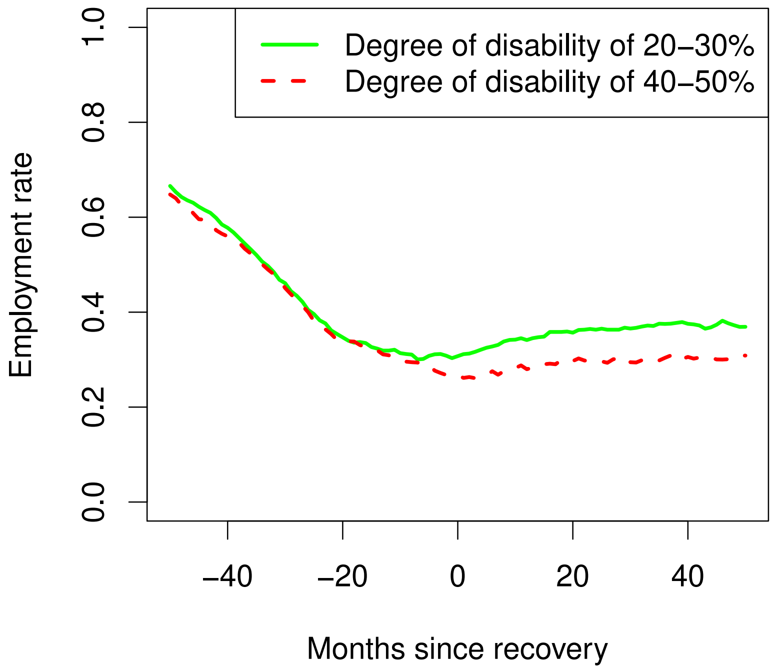
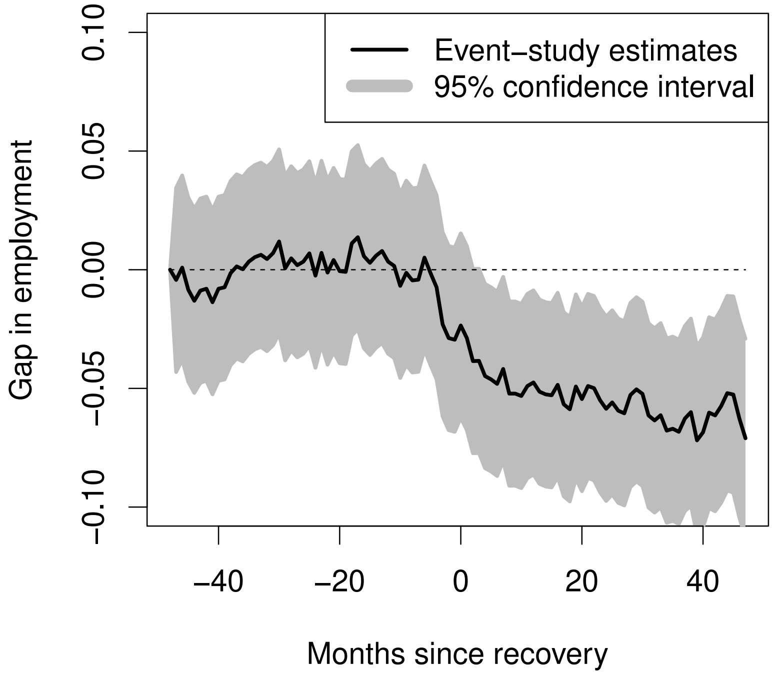


Figure 5.3: Event-study estimates (β_l from Equation (5.1)) mental health recovery for DI applicants with degree of disability of 20–30% vs. 40–50%.



One year after recovery, the effect of DI benefit receipt on the response to employment has accrued to approximately five percentage points in terms of employment. This disincentive effect is statistically significant for the months after recovery. The corresponding analysis on working hours yields a very similar picture, where the disincentive effect after recovery is approximately eight working hours. These results are shown in Figure 5.A.2 in the appendix.

To assess the magnitude of these findings, we perform a back-of-the-envelope calculation that compares the disincentive effect to the recovery response of individuals without benefits. From the estimated α_l parameters in Equation (5.1), we find that for individuals without benefits, recovery leads to approximately a 10 percentage-point increase in the likelihood of employment and an of approximately 11 working hours; this corresponds to the increase in the green line in Figure 5.2 after recovery. The disincentive effect of benefits receipt (5 percentage points) thus offsets approximately 50% of this increase in employment. The negative impact of DI benefits on working hours (8 hours) is even larger in this case: it eliminates almost 75% of the positive effect.

Given that treatment and control groups are in different degree-of-disability classes, one remaining concern may be that (some of) the disincentive effect stems from health differences between the treatment and control group. To some extent, those receiving benefits are in worse health. To test for differentials in recovery impacts, we therefore perform placebo tests and compare individuals with a similar difference in degree of disability, but with the same DI benefit status. In doing so, we also acknowledge the fact that the assessed degrees of disability may be endogenous close to the threshold of 35%, using donut estimates.

The placebo comparisons and the resulting DiD estimates are shown in Table 5.1. The columns ‘Benefit receipt’ indicates whether both groups receive DI benefits, and ‘Donut’

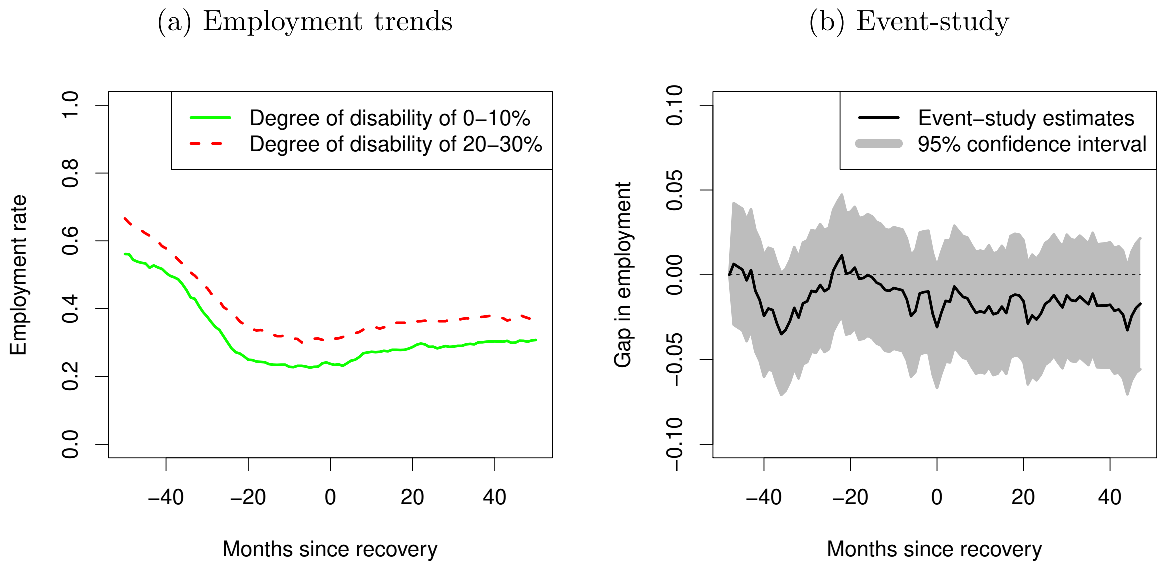
Table 5.1: DiD estimates on employment of possible placebo comparisons; groups defined by the degree of disability (%) of applicants

Test	Placebo group	Control group	Treatment group	Benefit receipt	Donut
		20–30% (N=3,346)	40–50% (N=1,656)		
(i)	0–10% (N=2,057)	– 0.009** (0.001)		NO	YES
(ii)	10–20% (N=2,597)	– 0.006* (0.001)		NO	NO
(iii)	50–60% (N= 1,509)		– 0.011** (0.001)	YES	NO

Standard errors in parentheses; * significant at a 5% significance level; **significant after applying the $\frac{1}{45}$ Bonferroni correction factor indicates whether there is a 10% donut in degree of disability between the two groups. The table shows three placebo comparisons of groups with the same DI benefit status but a different degree of disability (tests (i), (ii) and (iii)). The first two placebo tests compare groups without DI benefits, with or without a donut in the degree of disability. The third placebo test compares groups with DI benefits, without a donut in the degree of disability.[25]

The placebo test most similar to the main analysis compares the 0–10% group with the 20–30% group. Figure 5.4 shows that these groups follow a very similar employment pattern both pre- and post-recovery. While there is a level difference between the two groups, their response to recovery is very similar and the event-study estimates are not significantly different from zero. The same holds for the other two placebo tests.[26] It is therefore unlikely that differences between groups – other than benefit status – cause a substantially different response to an improvement in mental health. In the next section, we confirm this finding by showing that findings are robust against including the continuous degree of disability as a control variable.

Figure 5.4: Placebo test. Comparison of employment trends and event-study estimates for groups with equal benefit status but different degrees of disability



[25] A comparison between groups with DI benefits, and a donut in the degree of disability is not a valid placebo test as the group with a higher degree of disability also receives a higher amount of DI benefits. For completeness, we do show employment trends and resulting event-study estimates for this group in Appendix Figure 5.A.5.

[26] See Appendix Figures 5.A.3 and 5.A.4 for the trends in employment and the corresponding event-study estimates of placebo tests (ii) and (iii).

5.4.4 Robustness

We now turn to an assessment of the robustness of our findings by considering alternative specifications. For ease of comparison, we perform the robustness analysis on a simple DiD specification instead of an event-study specification.[27] In the DiD specification, we use a single dummy for pre- and post-recovery. In effect, the DiD estimate corresponding to our main event-study specification equals the average of the post-recovery event-study estimates, minus the average of the pre-recovery event-study estimates. The DiD specification is:

$$E_{it} = \alpha_{DI} + \alpha_t + \beta DI_i \, I_{t \geq 0} + \theta X_{it} + \varepsilon_{it}, \tag{5.2}$$

in which $t = 0$ is the month of recovery. The DiD specification allows for time-constant pre-recovery level differences between the groups with and without benefits (α_{DI}) and flexible time trends α_t (monthly dummies). The set of individual controls X_{it} are the same as in the event-study specification. The parameter of interest is β, which is the difference between the effects of recovery for the groups with and without disability benefits.

Table 5.2 presents the results of the robustness analyses. Given that the event-study estimates prior to recovery are approximately zero, the baseline DiD estimates are very similar to the average event-study estimates after recovery and amount to 5.3 percentage points in terms of employment and 7.8 working hours (row 1). The robustness analyses consider the effect of changes in the time window around recovery, the exclusion of observations close to the time of recovery and changes in the samples of treatment and control groups. Our DiD estimates for employment and number of working hours in Table 5.2 are generally robust to these changes.

A symmetric donut specification that excludes observations within 12 months around the moment of recovery (row 2) is implemented to account for potential measurement error in the timing of recovery. The resulting DiD estimates are slightly larger than in the baseline specification. Shortening the time window around the moment of recovery to 12 months decreases the DiD estimates somewhat (row 3). Presumably, this reflects the fact that a relatively large part of the pre-recovery divergence in outcomes is attributed to the pre-recovery difference in means, leaving less room for post-recovery treatment effects.

Finally, our estimates are robust to changes in the samples of treatment and control groups. For this, we first extend the groups with degrees of disability that are further away from the 35% disability threshold (row 4), which yields similar results. Next, we include all individuals with degree of disability close to the threshold of 35% (row 5). As mentioned earlier, we omit the 30–40% support in the baseline analysis due to concerns about manipulation around the threshold. If we do include these individuals, the resulting

[27] Robustness analyses on the actual event-study specification yield very similar results.

Table 5.2: Robustness analyses for the DiD effects for employment and working hours as outcome measures

		Specification				Outcome measures	
		Window^a	Donut	N_c^b	N_t^c	Employment	Hours
(1)	Baseline model	48	0	3,346	1,656	–0.053**	–7.786**
						(0.001)	(0.141)
(2)	Time window donut	48	12	3,346	1,656	–0.059**	–8.760**
						(0.002)	(0.193)
(3)	12-month time window	12	0	3,346	1,656	–0.035**	–4.984**
						(0.002)	(0.207)
(4)	15–30% vs. 40–55%	48	0	4,726	2,480	–0.058**	–8.922**
						(0.001)	(0.115)
(5)	25–35% vs. 35–45%	48	0	3,270	1,788	–0.022**	–4.041**
						(0.001)	(0.141)
(6)	No relapses	48	0	2,234	1,551	–0.067**	–9.970**
						(0.002)	(0.222)
(7)	Degree of disability control	48	0	3,346	1,656	-0.047**	-3.993**
						(0.004)	(0.461)

[a]Incorporated number of months before and after recovery; [b]Number of individuals in the control group; [c]Number of individuals in the treatment group; Standard errors shown in parentheses; *significant at a 5% significance level; **significant after applying a $\frac{1}{45}$ Bonferroni correction factor.

estimates become smaller, indicating that manipulation could indeed be a problem. Still, even in this specification, the estimates remain negative and statistically significant. In row 6, we limit the sample to individuals for whom we observe that they do not relapse for at least three years (in the main analysis, this is at least one year). The resulting estimates are a bit larger, indicating that part of the main sample indeed relapses. Note that Table B.6 in the online appendix complements these findings with another set of robustness tests (a specification without covariates, mean-level estimation and non-parametric estimation), which also yield results that are very similar to those for the baseline model.

To avoid contamination due to differences in health, we also use a specification that controls for the (continuous) degree of disability as measured during the DI assessment. Specifically, we allow the degree of disability to have differential impacts before and after recovery. Note that this setup still allows for identification of the disincentive effect of benefits: the degree of disability is continuous, while benefits are awarded once the degree passes the threshold of 35%. When controlling for degree of disability in this way, the DiD estimate for employment hardly changes, as we find an estimate of –4.7 percentage points (row 7). These results confirm the results of the placebo regressions shown in Figure 5.4, which also find a limited effect of differences in health. For hours worked the inclusion of degree of disability reduces the disincentive effect a bit, although we still find a negative significant estimate of four hours per month.

Table 5.3: Heterogeneity by mental health disorder: DiD estimates for employment and number of working hours

	Specification				Outcome measures	
	Window[a]	Donut	N_c^b	N_t^c	Employment	Hours
Baseline model	48	0	3,346	1,656	–0.053**	–7.786**
					(0.001)	(0.141)
Mood disorders[d]	48	0	1,304	690	–0.043**	–7.511**
					(0.002)	(0.224)
Anxiety disorders[e]	48	0	433	174	–0.048**	–9.897**
					(0.004)	(0.435)
Personality disorders[f]	48	0	556	279	–0.064**	–8.245**
					(0.003)	(0.355)
Other disorders	48	0	1356	669	–0.081**	–10.860**
					(0.002)	(0.213)

[a]Incorporated number of months before and after recovery; [b]Number of individuals in the control group; [c]Number of individuals in the treatment group; [d]Mood disorders include (but are not limited to) depression, manic disorder and bipolar affective disorder; [e]Anxiety disorders include (but are not limited to) panic disorders, generalized anxiety, agoraphobia and social phobia; [f]Personality disorders include (but are not limited to) paranoia, schizophrenia, dissocial personality disorder and borderline; Standard errors are shown in parentheses; *significant at a 5% significance level; **significant after applying a $\frac{1}{45}$ Bonferroni correction factor.

To shed light on heterogeneous disincentive effects and to make the control and treatment groups more comparable at the same time, Table 5.3 shows the results of the baseline models estimated on sub-samples that are stratified by types of mental health diagnoses. We consider the three most prevalent mental health diagnoses and a separate sample of individuals suffering from disorders other than the three most prevalent ones. Within these sub-samples, the characteristics of the control and treatment group are very similar.[28] For all sub-samples, DI benefits significantly reduce the response to recovery, but the magnitude ranges from 4.3 to 8.1 percentage points in terms of employment and 7.5 to 10.9 working hours for the various disorders. Individuals diagnosed with mood disorders are affected least by DI benefits, while individuals in the “other disorders” sample are affected most severely by DI benefit receipt.

Summing up, we find robust disincentive effects of having DI benefits on the employment response to mental health recovery. The estimated effects range between –2.2 to –5.3 percentage points in terms of employment and between –3.9 and –8.9 working hours per month. Part of these effects accumulates already before the end of treatment, so individuals either anticipate their recovery or recovery happens before the end of treatment. Since the relative effects on the average number of working hours exceed those on the employment rate, there are both intensive and extensive margin effects on labor supply.

[28]See Online Appendix Tables B.3, B.4 and B.5 for descriptive statistics for sub-samples.

5.4.5 An alternative recovery proxy based on healthcare expenditure data

So far we have employed a proxy for recovery based on the end of mental health treatment trajectories. An alternative approach is to consider health expenditures and define recovery as a substantial decrease in such expenditures. The apparent downside of this is that these expenditures are only available on an annual basis, allowing for a less precise measurement of the time of recovery (the mental health trajectory data includes the exact dates). On the other hand, healthcare expenditures offer the advantage that they contain both mental and non-mental (physical) expenditures, allowing for an analysis of all types of health recovery. We therefore conduct similar analyses of recovery effects, but now with drops in physical and mental healthcare expenditures as proxies for recovery. Specifically, we define the year of recovery as the first year in which expenditures drop below and stay below 20% of the healthcare expenditures in the year before the DI application.[29] We distinguish between mental healthcare expenditures and non-mental healthcare expenditures (see Online Appendix Section B.3 for further details), yielding two proxies for recovery. We select applicants for whom we observe physical or mental recovery and that have a degree of disability in the relevant treatment and control groups that have been explained earlier. The resulting sample contains 7,418 individuals with mental health recovery and 9,747 individuals with non-mental health recovery.

For both samples and corresponding types of recovery, Figure 5.5 shows the event-study estimates relative to the year of recovery (see the left Y-axis) as well as bars indicating the average mental or physical healthcare expenditures for each year (right Y-axis).[30] Prior to recovery, expenditures increase to approximately €8,000 (mental healthcare) and €13,000 (physical healthcare) per year. After the decrease, the expenditures remain low. For mental health, the trends in the employment rates are very similar for the control and treatment group. A divergence in employment rates starts approximately one year before recovering, corresponding with the fact that recovery most likely occurs in the year prior to the first low-cost year. For physical health, the trends are less similar for the control and treatment group, and divergence in employment rates starts approximately two years before the first low-cost year. Such early divergence may (to some extent) result from the imprecise measurement of the timing of healthcare expenditures.

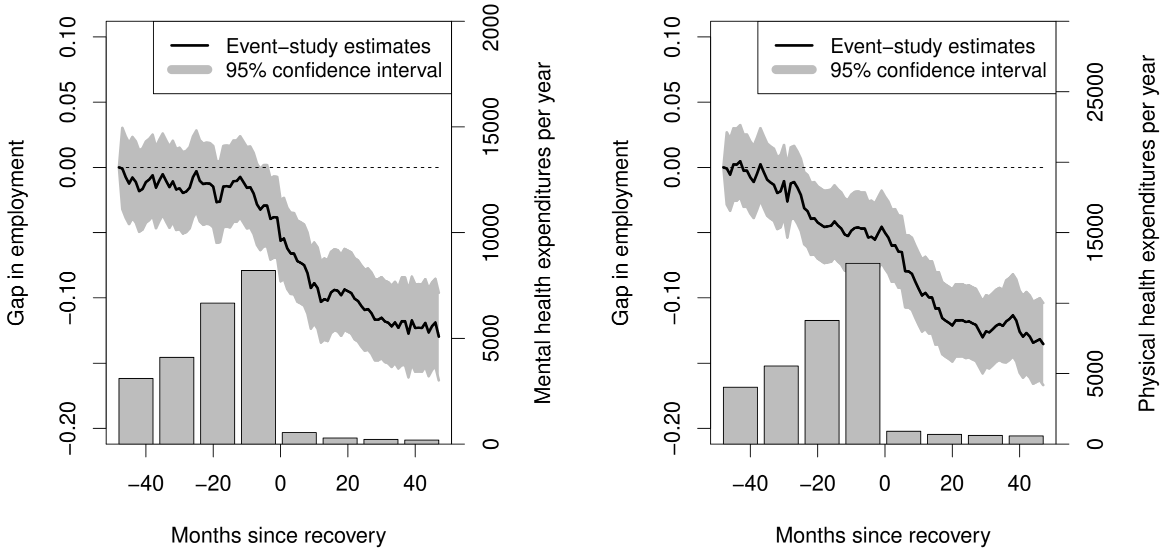
The divergence in employment is statistically significant for both proxies and the difference between the two groups remains relatively constant at approximately 10 percentage points after recovery. The effect size is larger compared to our estimates using the end

[29] Results for other thresholds are similar, see Table B.7 in the online appendix.

[30] See Appendix Figure 5.A.6 for the actual trends in employment

Figure 5.5: Event-study estimates (β_l from Equation (5.1)), using a recovery proxy based on healthcare expenditures

(a) Mental health (b) Physical health



of mental health treatment as proxy for recovery. One potential explanation for this is that healthcare expenditures capture a wider range of treatments (e.g. they also include pharmaceutics) and therefore proxy recovery more accurately.

Table 5.4 displays the DiD estimates for both proxies for recovery. Panel (a) shows the results for recovery based on a decrease in mental health expenditures and panel (b) shows the results for recovery based on a decrease in physical health expenditures. Similar to our earlier baseline model, we incorporate 48 months before and 48 months after recovery. Knowing that recovery could either occur in the first low-cost year or in the year prior to the first low-cost year, we also present outcomes where we incorporate a donut of 12 months before and 12 months after the moment of recovery. All results support our earlier baseline findings: we find significant negative effects of benefits on the response to recovery for both employment and hours worked. Both for mental health recovery (panel (a)) and for physical health recovery (panel (b)) the estimates are somewhat larger in magnitude than in our baseline results when a donut is included. This possibly reflects the fact that the mental health treatments that we use in our baseline, are only a subset of all mental healthcare expenditures.

It should be stressed again that healthcare expenditures are only available on an annual basis, which inevitably leads to larger imprecision in defining the time of recovery. Therefore these results carry more uncertainty than our baseline estimates that are based on the exact end dates of mental health treatment trajectories. Nevertheless, the similarity in results from both types of proxies is reassuring.

Table 5.4: DiD estimates based on annual mental and physical healthcare expenditures for employment and number of working hours

	Specification:				Outcome measures	
	Window[a]	Donut	N_c^b	N_t^c	Employment	Hours
(A): Mental health recovery						
Baseline specification	48	0	4,735	2,683	–0.085**	–13.077**
					(0.001)	(0.121)
Donut specification	48	12	4,735	2,683	–0.099**	–15.588**
					(0.001)	(0.166)
(B): Physical health recovery						
Baseline specification	48	0	6,301	3,446	–0.080**	–13.030**
					(0.001)	(0.110)
Donut specification	48	12	6,301	3,446	–0.100**	–16.309**
					(0.001)	(0.153)

All estimates are based on regressions including degree of disability controls (specification 5.2). [a]Incorporated number of months before and after recovery; [b]Number of individuals in the control group; [c]Number of individuals in the treatment group. Standard errors are shown in parentheses; *significant at a 5% significance level; **significant after applying a $\frac{1}{45}$ Bonferroni correction factor.

5.5 Labor supply effects in a structural model

Our estimation results point to distinct employment effects of recovery for individuals with and without DI benefits, both for health improvements that are proxied by the end of mental health treatments and by substantial drops in medical consumption. When interpreting these results, recall that the employment increase after recovery only partly compensates for the large employment drop that occurred in earlier years. We stated earlier that the response effects can be characterized as ITT-effects, since the end of the treatment implies partial or full recovery only for a part of the sample. This raises the question of how large the employment effects of full recovery are and what the maximum discouraging impact of DI benefits is.

For insight into the full effects of recovery, we develop a structural labor supply model that incorporates both health shocks and the subsequent recovery from health. With the data on the assessed earnings capacity and possible hours restrictions of workers at the moment of DI application, the impact of health changes can be modelled as changes in budget constraints and changes in the maximum number of hours that can be worked. Following Low & Pistaferri (2015), we also allow for changes in utility preference parameters that stem from health shocks. These health shocks may permanently or temporarily change preferences for leisure. In effect, labor supply changes may stem both from productivity losses and a higher disutility from working.

To estimate the structural model parameters, we distinguish three successive stages that are relevant for the individual worker: (1) before the health shock, (2) after the health shock and (3) after recovery. For the first stage, we estimate utility preference parameters for work and leisure for each individual in the sample – based on the observed hours decision (and the hourly wage) before the onset of a disability. In the second stage, the onset of the disability implies a loss in earnings capacity along two observed dimensions: the maximum number of working hours and the hourly wage. If the loss of earnings capacity exceeds 35% of the previous wage earnings, DI benefits are awarded. Together with this decrease in the earnings capacity, we incorporate a shock to the utility parameters in the model. In the third stage, recovery implies that the earnings capacity is restored and utility parameters return to their pre-disability level. Eligibility to DI benefits is maintained for those who were awarded benefits.

Our model abstracts from tax effects and assumes hourly wages to be exogenous. In addition, we assume that individuals receive social assistance if their income from earnings and DI is below the social minimum. In the Netherlands, most workers are entitled to unemployment benefits in the short term, and to social assistance in the long term when UI benefits are exhausted. Our focus is therefore on the long-term effects that occur after the onset of disability. The following subsection describes the specification of the model and presents a graphical illustration. Subsequently, we discuss the estimation of the model parameters and a counterfactual simulation.

5.5.1 Model setup

Our model assumes utility maximization over the number of hours worked. We adopt a Cobb-Douglas utility function with utility weights normalized to one.[31] Since eligibility for social assistance depends on partner income, our focus is on income at the household level. Let T be the total amount of time an individual can divide between leisure L_i and employment E_i. I_i is total income, consisting of labor income ($E_i\, w_i$), potential DI benefits ($DI(E_i)$) and partner income ($\tilde{I}_i$). If income falls below the social assistance level, it is supplemented up to this level.[32] The utility maximization problem for individual i is as follows:

[31]Since the utility function is estimated on a single employment decision, we can identify one preference parameter at most.

[32]The social assistance level is approximately €1,000 per month for singles and €1,500 for couples.

$$\max_{E_i} u(L_i, I_i) = L_i^{\lambda_i+\delta_g} I_i^{1-\lambda_i-\delta_g} \tag{5.1}$$

$$\text{s.t. } L_i = T - E_i \tag{5.2}$$

$$I_i(E_i) = E_i w_i + DI(E_i) + \tilde{I}_i \quad \text{if} \quad E_i w_i + DI(E_i) + \tilde{I}_i \geq SA_i \tag{5.3}$$

$$I_i(E_i) = (1-F)\ SA_i \quad \text{if} \quad E_i w_i + DI(E_i) + \tilde{I}_i < SA_i \tag{5.4}$$

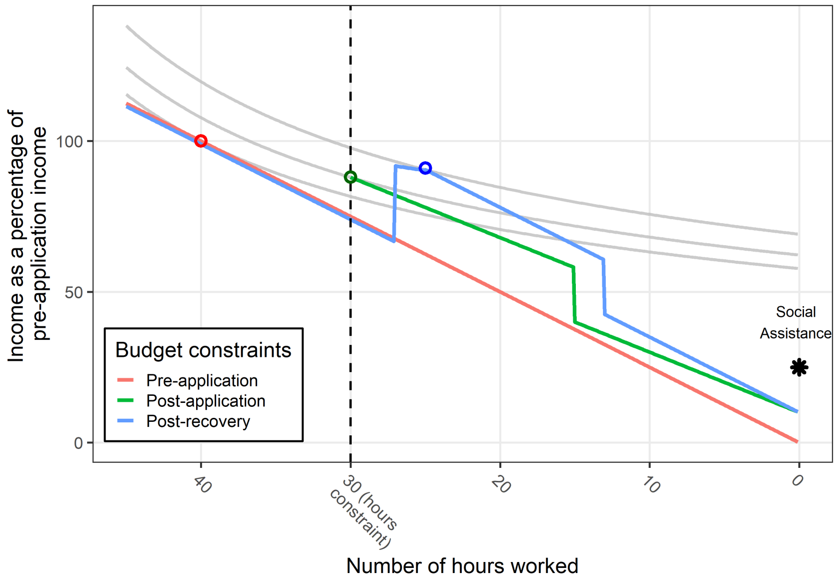
We set $T = 60$, the highest observed pre-application number of working hours. Income from social assistance is discounted by a factor $1 - F$, resembling the stigma attached to receiving social assistance and the (perceived) costs of ongoing eligibility conditions that apply to social assistance recipients (for example job search requirements). The size of F is identified from the size of hidden unemployment in the DI application inflow, i.e. the number of people for whom the pre-application utility level from working is below the pre-application utility level of receiving social assistance. Lastly, λ_i displays the individual-specific utility parameter and δ_g is a group-specific shock to this utility parameter. By definition, δ_g is equal to zero in the pre-application stage.

To illustrate the functioning of the model, Figure 5.1 shows the choice set for individuals with DI benefits.[33] The figure does not include any shocks to the utility function and demonstrates an individual working full-time (40 hours) before the DI application. The individual has an assessed degree of disability of 40% and will therefore be awarded disability benefits. The degree of disability results from an hours restriction of 75% and a 20% reduction in the hourly wage.

The pre-application budget line (red) increases one-to-one with the wage earnings. The post-application budget constraint is indicated by the green line; it shows that the slope of the budget constraint decreases due to the 20% reduction in hourly wage. Furthermore, the disability benefits cause non-linearities at the various thresholds of the DI system. Most notably, there is an incentive to work at least half of one's remaining work capacity (15 hours in this case) and the assessed degree of disability enforces a maximum number of working hours such that the budget constraint ends at 30 hours in this case. Lastly, the post-recovery budget constraint is shown by the blue line. As long as the individual receives disability benefits after recovery, the discontinuities of the post-application period persist. With a higher hourly wage, these discontinuities occur at a lower number of hours worked. In the example shown above, the response to recovery therefore entails a further reduction in the number of working hours. This contrasts with the case of a similar individual without disability benefits, who would increase the number of working hours up to the pre-disability level (see Online Appendix Figure B.3).

[33]See Online Appendix Figure B.4 for the same illustration for an individual without DI benefits.

Figure 5.1: Budget constraints and utility indifference curves of a fictitious individual awarded partial disability benefits with a degree of disability of 40%



5.5.2 Estimation procedure

We briefly describe the estimation procedure, while referring to Online Appendix Section B.4 for further details. Following the timing of the employment decisions an individual faces, the estimation occurs in three stages. In the pre-application step, we estimate the individual utility parameter δ_i using observations of the hourly wages and the number of working hours. Given that every individual in the sample worked prior to applying for DI benefits, there is a unique δ_i for each individual which results in the observed number of working hours. Without any fixed costs of social assistance receipt, this would lead to a substantial fraction of workers for whom utility from receiving social assistance exceeds the utility from working their observed number of hours. For a better fit of the model to the data, we therefore set the fixed costs F such that only 5% of the sample obtains higher utility from social assistance than from their (observed) number of working hours.[34] The resulting fixed cost of receiving social assistance equals 0.46, meaning that one euro received from social assistance is valued the same as €0,54 earned through working.

In the second step, the individual has experienced a health shock and may or may not have been awarded benefits. The budget constraints are altered due to a reduction in potential hourly wage, a restriction on the maximum amount of working hours and the provision of DI benefits for those awarded DI benefits. Using these new budget constraints, the new optimal number of working hours is predicted and confronted with

[34] As a robustness test, we also use a 10% level. This yields similar results for the benefit disincentive effects we obtain in the recovery stage.

the observed post-application number of working hours. To optimize the fit to these data, the group-specific utility parameter shock (δ_j) is set such that the average predicted number of working hours of each group equals the average observed number of working hours. The resulting estimates for the shocks are 0.24 for those without benefits and 0.26 for those with benefits.[35]

In the post-recovery step, the response to recovery is simulated by restoring the earnings capacity to its pre-application level and setting the utility parameters to their original values as well (i.e., we set δ equal to zero). In this respect, the differences in the predicted response between the 20–30% group and the 40–50% group can be interpreted as the analogue of the DiD estimate without degree of disability controls (as they also include the effect of differences in health). These results are presented in Online Appendix B.4.

5.5.3 Counterfactual employment outcomes of awarded individuals

With a larger assessed loss in earnings capacity, individuals with DI benefits are different from individuals without DI benefits to some extent. Similar to the DiD analyses without degree of disability controls, this means that derived differences in response rates to recovery may also reflect compositional differences. Given the structural setup of our model, however, we can offset such effects by constructing counterfactual employment outcomes for the sample of individuals with DI benefits, as if they had not been granted DI benefits instead.[36] This allows us to assess the discouraging impact that benefit receipt may have on work resumption after application and recovery. Table 5.1 shows the results that follow from this approach.[37]

For the post-application stage, we find employment rates that are comparable in the 'true' scenario with DI benefits and in the fictitious scenario without DI benefits. The employment rate is even slightly lower in the absence of DI benefits. The predicted employment rates differ by 1.7 percentage points and the predicted average number of working hours are almost identical.[38] This suggests that the discouraging impact of benefit receipt is small among disabled individuals.

Upon recovery of the earnings capacity, there is a widening in the difference in the

[35]Note that the similarity in shock effects indeed suggests groups in the DiD analysis are similar.

[36]A similar exercise was conducted for applicants without DI benefits, giving similar results.

[37]Given our structural parameters, we can also broaden our analysis to recovery effects with benefit conditions that differ from the Dutch context. In this respect, one may argue that incentive structures are different with different replacement rates or earnings caps and without minimum earnings requirements. In Online Appendix Section B.4 we have conducted such analyses.

[38]Recall that the DI scheme inhibits an incentive to exploit at least half of the remaining earnings capacity. This may explain why benefits lead to slightly higher employment.

Table 5.1: Predicted employment and working hours effect of DI receipt for workers with degree of disability of 40–50%; counterfactual analysis based on structural model

	Pre-illness	Post-application	Recovery of earnings capacity	Recovery of earnings capacity and utility parameters
Employment				
Without DI benefits	96.6%	31.0%	54.1%	96.6%
With DI benefits	96.6%	32.7%	50.8%	81.3%
DI benefit effect		1.7%	-3.3%	-15.3%
Weekly working hours				
Without DI benefits	33.3	5.2	10.0	33.3
With DI benefits	33.3	5.5	6.6	16.0
DI benefit effect		0.3	-3.4	-17.3

All values concern predicted outcomes based on the estimated parameters. By construction, pre-application values are equal to the full-recovery values in absence of DI benefits.

employment rates of the two groups. Assuming that utility preference parameters stay constant but the earnings capacity is restored (as shown in the second column), DI benefits seem to discourage 3.3 percentage-point from work resumption, as compared to a maximum work resumption effect (without benefits) equal to about 23 percentage points. When utility parameters are also restored to their pre-application values and there is “full” recovery (third column), the discouraging impact increases to 15.3 percentage points. Roughly speaking, this is about a quarter of the impact of full recovery for those without benefits of about 65 percentage points. For working hours, we also see a widening of disincentive effects after recovery of the earnings capacity and after full recovery. The relative size of these effects is larger, indicating that the structural model predicts decreases in working hours for those employed and receiving benefits. In this way, these individuals avoid the loss of DI benefits.

Comparing these simulations with our empirical findings from Section 5.4, we obtain two main insights. First, it indeed appears as if our proxy for recovery measures partial recovery, as the estimated employment response based on the regression results of Equation (5.1) (in the absence of benefits) was around 10 percentage points while here it is 23 percentage points. Second, our simulations confirm our earlier result that the disincentive effect (measured as the share of the employment response that is undone by benefits) is larger for hours worked than for employment.

5.6 Conclusion

This chapter studies whether labor supply responses to improvements in mental health are partly eliminated by the disincentives of disability benefits. In doing so, we aim to deepen our understanding of low work resumption rates of DI benefit recipients, particularly in schemes intended for those deemed temporarily disabled. Applying a difference-in-differences (DiD) framework, we compare Dutch DI applicants below the degree of disability eligibility threshold as a control group with those above the degree of disability threshold. The control and treatment groups have parallel trends in employment leading up to recovery, as proxied by the end of mental health treatment.

Our analyses rely on the assumption that this proxy captures recovery equally well for the control and treatment group. While we conduct a variety of tests that support this claim, we cannot conclude with full certainty that no remaining differences exist in this proxy. Under the assumption that recovery rates are indeed similar for both groups, we find that the disincentive effects of disability benefits amount to approximately half of the recovery response for individuals without DI benefits for employment. For hours worked, the disincentives eliminate 75% of the recovery response.

Knowing that our reduced form results should be interpreted as Intention-to-Treat estimates that resemble the effect of partial recovery, we construct and estimate a structural labor supply model that allows us to uncover full recovery responses. We make use of the fact that we observe pre-DI-application labor supply, earnings capacity after falling ill (from the DI application assessment) and post-application labor supply. Using the estimated parameters, we simulate the response to full recovery of earnings capacity for the two groups with distinct budget constraints (due to DI benefits). We then find that disability benefits reduce the response to recovery by approximately 15 percentage points in terms of employment, suggesting that DI benefits absorb at least a quarter of the response in case of full recovery.

Compared to earlier findings, the reduced form estimates of disincentive effects in our setting amount to an employment reduction of about five percentage points. In the literature, estimates of employment reductions typically range between 20 and 30 percentage points Bound (1989); Chen & Van der Klaauw (2008); Maestas et al. (2013); French & Song (2014). We argue that there are two explanations for this difference. First, the end of mental health treatments represents only partial recovery. Based on our structural model, the disincentive effect is substantially larger and can amount to 15 percentage points in case of full recovery. Second, we consider partial DI benefits, while most other papers on the disincentive effects of DI benefits examine the effects of full DI benefits. Benefits in our setting amount to approximately 500 euros a month, whereas for example, SSDI benefits amount to between 800 and 1,800 dollars per month on average. Disincentive effects of partial DI benefits that occur prior to recovery are found to be relatively small, amounting to approximately three percentage points in terms of employment for the Netherlands Koning & Vethaak (2021). For full SSDI benefits, these effects are estimated to range between 20 to 30 percentage points in terms of employment (Bound, 1989; Chen & Van der Klaauw, 2008; Maestas et al., 2013; French & Song, 2014)). Relative to the small disincentive effects of partial benefits, the additional disincentive effects upon recovery are therefore substantial.

Since benefit disincentive effects to resume work after recovery are found to be substantial, a case can be made for several policy options. First of all, DI benefit providers could use healthcare data to more selectively target and reassess potentially recovered DI beneficiaries. Currently, all DI beneficiaries should be reassessed periodically. Due to capacity constraints, however, such reassessments are rare. By using healthcare data, the limited reassessment capacity could be used for those individuals who are most likely to have experienced an improvement in health. Second of all, changes to the setup of the DI benefit scheme could decrease the disincentive effects to resume work, particularly those inherent with the cash cliff of benefit receipt. Policymakers might for example consider increasing the benefits which are awarded upon (full) use of the remaining earnings capacity.[39] Specifically, allowing DI beneficiaries to earn more than their remaining earnings capacity for a longer time period, without cutting DI benefits, might alleviate the fear of losing DI benefits if one attempts to rejoin the workforce. Given the disincentives found in this chapter, further research on these potential policy changes is warranted.

[39]Changes to the DI benefit setup could increase work resumption rates upon recovery. We explore potential changes through the structural model, as shown in Online Appendix Section B.5. These changes could however also influence the inflow rates into DI.

5.A Appendix

Figure 5.A.1: Discontinuity test of application densities at the 35% threshold of degree of disability

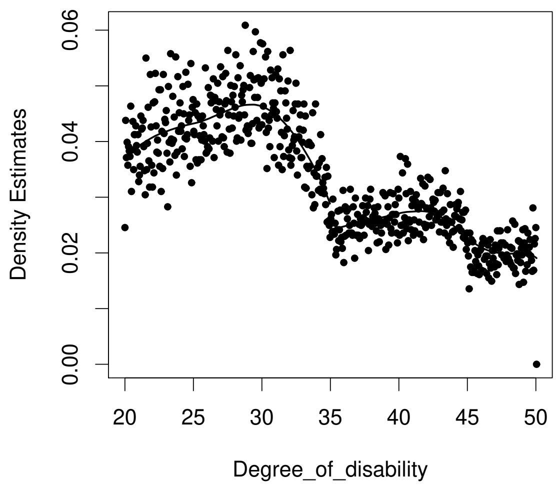
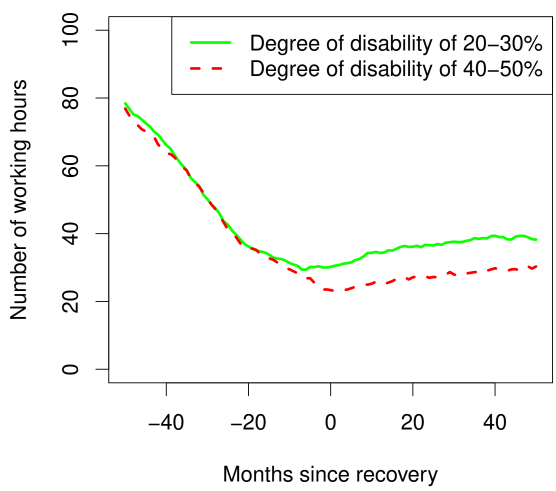


Figure 5.A.2: Trends in number of working hours (panel (a)) and event-study estimates (b) relative to mental health recovery for DI applicants with degree of disability of 20–30% vs. 40–50%.

(a) Trends



(b) Event-study

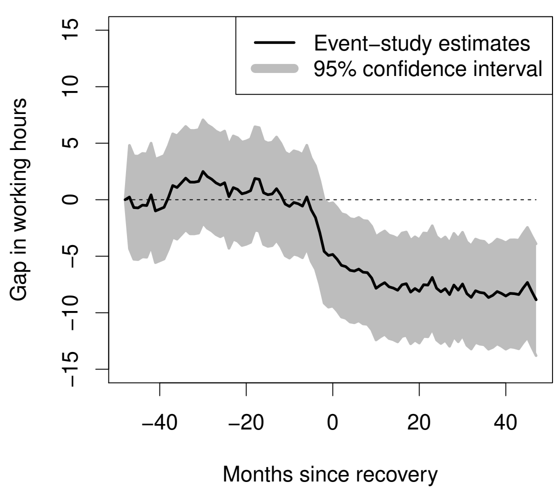


Figure 5.A.3: Trends in employment (a) and event-study estimates (b) relative to mental health recovery for placebo groups with degree of disability of 10–20% vs. 20–30%.

(a) Trends (b) Event-study

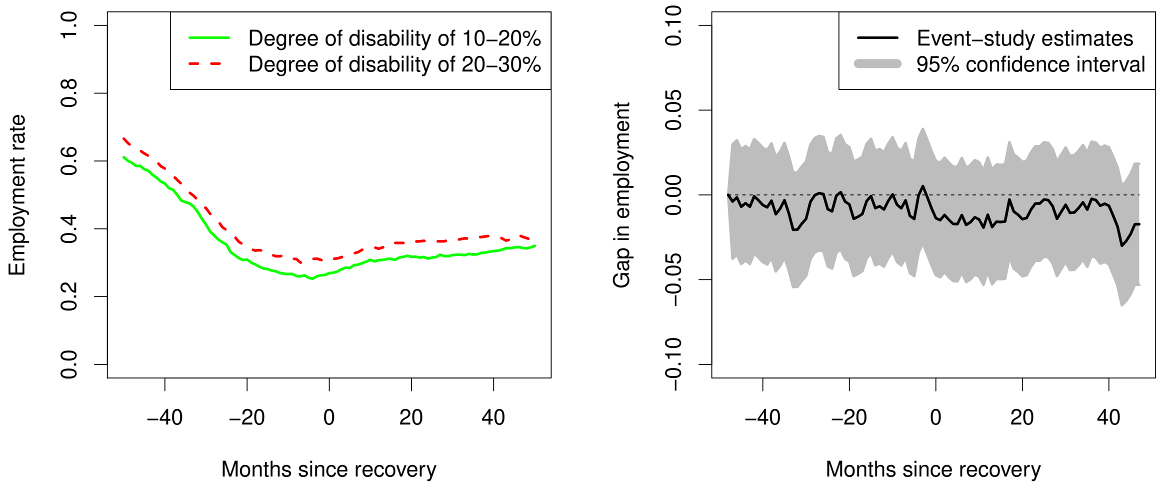


Figure 5.A.4: Trends in employment (a) and event-study estimates (b) relative to mental health recovery for placebo groups with degree of disability of 40–50% vs. 50–60%.

(a) Trends (b) Event-study

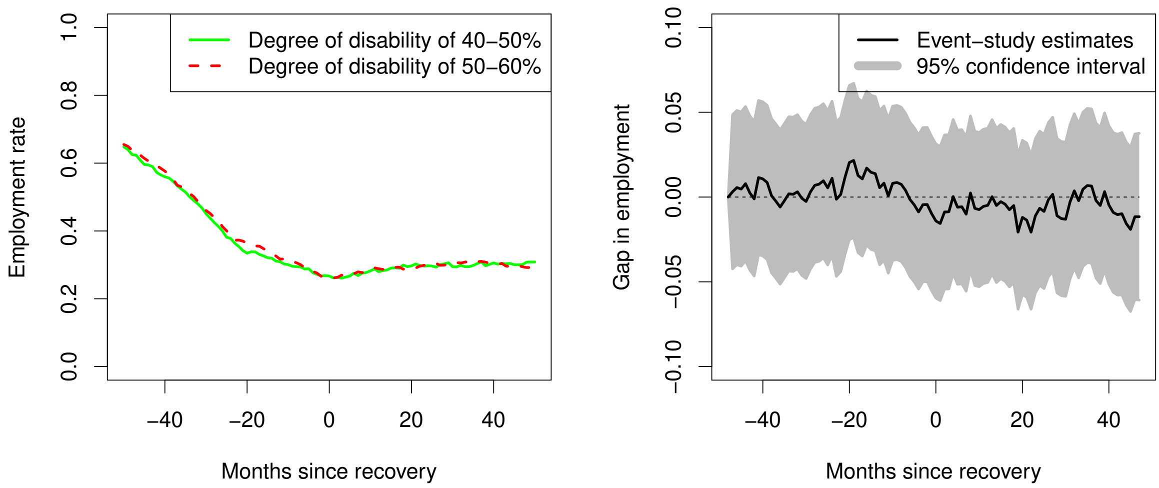


Figure 5.A.5: Trends in employment (a) and event-study estimates (b) relative to mental health recovery for groups with degree of disability of 40–50% vs. 60–70%.

(a) Trends (b) Event-study

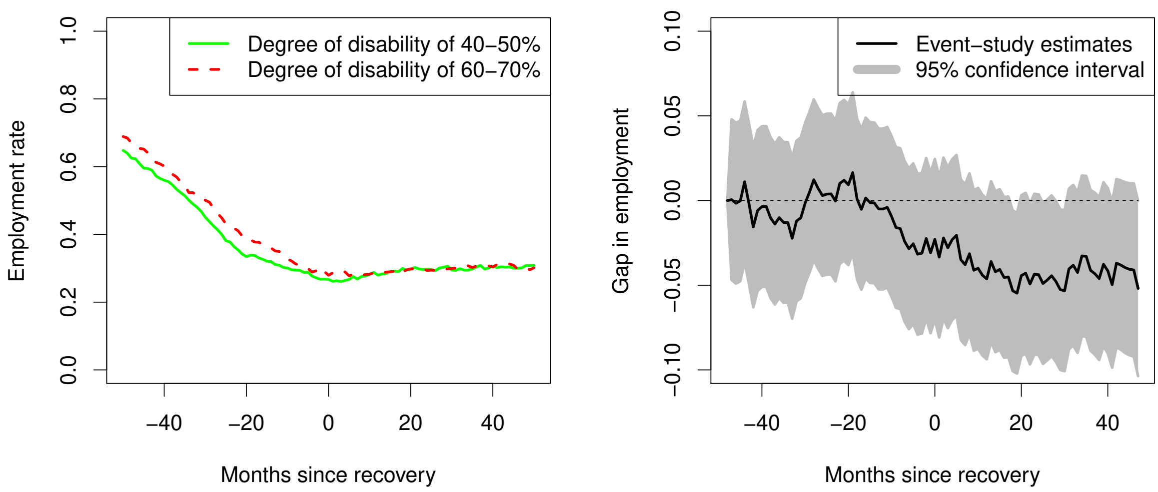
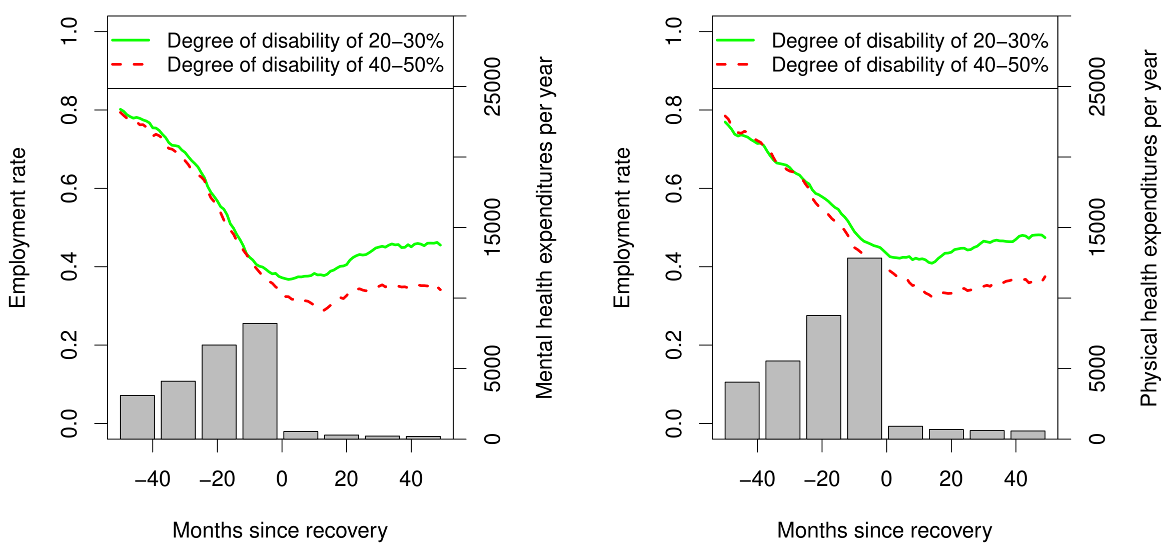


Figure 5.A.6: Employment rates relative to recovery based on healthcare expenditures

(a) Mental health (b) Physical health



Summary

Despite a series of reforms between 1996 and 2006 aimed at reducing the inflow into disability insurance in the Netherlands, a substantial share (7%) of the Dutch labor force still receives disability benefits. To gain insights into which groups are most at risk of entering DI, what measures can be taken to prevent this, and how to support them to return to work, this dissertation consists of four chapters that empirically analyse the process from the onset of health problems, through the DI application until the potential recovery of DI recipients in the Netherlands. In this final chapter, each chapter is briefly summarized and the main conclusions are presented.

Chapter 2 examines the first step in the process towards DI, which is the onset of mental health problems and the subsequent treatment of these problems. While mental health issues are becoming increasingly prevalent and constitute a substantial share of DI applications in the Netherlands, there are long waiting lists for the treatment of these problems. Using an instrumental variable (IV) approach which exploits variation in the congestion of the mental healthcare system across municipalities, large and persistent negative effects of waiting times are found. Specifically, a two-month increase in waiting time (i.e., one standard deviation) decreases the probability of being employed by four percentage points, while increasing the probability to receive disability benefits by two percentage points. The largest burden of these increased waiting times falls upon individuals with a migration background and with lower educational attainment. The impact of waiting times on employment is almost twice as large for them as it is for individuals without a migration background or with higher educational attainment, whereas they also have to wait longer on average before receiving treatment.

Chapter 2 shows that health problems and the treatment of these problems are important in understanding the inflow into DI. However, there are still large differences in DI inflow rates between groups that cannot be explained by health conditions. Chapters 3 and 4 examine two specific groups with high DI inflow rates. Chapter 3 compares the DI

application behavior of workers with fixed-term and permanent employment contracts. Workers with a fixed-term contract are approximately 50% more likely to apply for DI benefits, compared to their counterparts with permanent employment contracts.

To understand what causes this large DI risk premium, a decomposition is made into the following four mechanisms that may contribute to the higher DI application risk; (1) differences in characteristics of workers with permanent and fixed-term contracts, (2) the causal impact of contract type on health shocks, (3) employer incentives during the sickness period, and (4) labor market prospects during the sickness period. While workers with the two contract types differ significantly in terms of age and job type, this does not explain the DI risk premium. Rather than that, if the demographic composition of workers with fixed-term contracts would be the same as for workers with permanent contracts, the DI risk premium would be substantially larger. The second mechanism, which would entail a higher risk of physical or mental health shocks with fixed-term contracts, does not explain the DI risk premium either as the probability of experiencing a health shock is similar for both contract types. Most of the gap in DI application probabilities thus arises after the onset of illness. More specifically, the DI risk premium dropped substantially after a reform in 2013 that increased the incentives for employers to re-integrate their sick temporary workers. The DI risk premium which remains after the reform can largely be explained by labor market prospects of sick employees. Permanent workers can always return to their old job after sickness, while temporary workers will often have to find a new job after they have recovered. If it is relatively easy for sick temporary workers to find a new job (in sectors with a tight labor market), the probability to apply for DI benefits is almost equal for both contract types, whereas the DI risk premium is still sizeable in sectors in which it is relatively difficult to find a new job.

Chapter 4 zooms into a group with an even higher risk of entering DI; unemployment insurance (UI) recipients. This group is four times as likely to apply for DI, compared to workers with a temporary contract. Additionally, there is a spike in inflow at this moment, raising the question of whether this is driven by relatively healthy workers trying to extend the total duration of benefit receipt ("moral hazard") or by sick individuals who were initially unaware that they could call in sick while on UI. A comparison between UI recipients entering SI prior to the end of their UI entitlement period (the "pre-spike cohort") and those entering into SI at the end of their UI entitlement period (the "spike cohort") shows that those in the spike cohort are more likely to remain on SI and eventually apply for DI. Overall, they are thus less likely to be screened out during the two-year waiting period. Furthermore, the spike cohort is more likely to have a migration background and appears to have a weaker pre-UI labor market status, whereas both cohorts are similar in terms of healthcare utilization at the moment they call in sick. It is therefore unlikely that the spike in the inflow into SI at the end of UI entitlement is driven by relatively healthy individuals trying to extend the duration of benefit receipt. Rather than that, the analyses point to catch-up of initial non-take-up of specific groups of the population who might not have been aware of the fact that they would be eligible for SI.

Chapter 5 looks into the last stage of the sickness and DI, where the potential recovery of DI recipients may occur. Despite the fact that a majority of DI recipients with partial benefits are expected to recover at some point in time, outflow rates out of DI and into employment are small. To determine whether the DI system itself hinders recipients to return to work upon recovery, a comparison is made between the response to recovery of individuals with and without disability benefits. Zooming into the threshold of DI benefit denial/approval ensures that both groups are comparable. Since recovery is not directly observed, the end of healthcare treatment is used as a proxy for this. In this context, the end of treatment is an "intention-to-treat" effect that captures partial recovery or the recovery for a subsample of the full population. Both the control and treatment groups show similar recovery patterns and almost identical employment trajectories prior to the end of healthcare treatment. However, after the end of healthcare treatment, a smaller share of the treatment group returns to work. A difference-in-difference approach, in which treatment is defined as reaching the end of treatment while receiving DI benefits and control is defined as reaching the end of treatment while not receiving DI benefits, confirms that the DI benefit system hinders half of all DI recipients to return to work, once they recover.

Concluding, all chapters confirm that while experiencing health issues might be a necessary condition for applying for DI, it is certainly not a sufficient condition to explain differentials in the inflow into DI or work resumption across groups. Even though DI inflow rates in the Netherlands have been reduced by a number of drastic reforms, contextual factors (also non-health related) still matter substantially for both the inflow and the potential outflow out of the DI system. The adequate and timely provision of treatments in the sickness period, the commitment of employers to provide preventative and reintegration activities, labor market prospects, and benefit conditions of the DI benefit scheme itself all make a difference when it comes to the application for and receipt of DI benefits and (partial) work resumption. In terms of targeting, this does not imply that "wrong" (relatively healthy) individuals end up receiving DI in the current system. Groups with higher DI application risks due to these contextual factors exhibit very similar health trajectories leading up to their DI application, compared to groups with lower DI application risks. These contextual factors affect individuals across the entire health spectrum. For example, a DI applicant with a temporary contract might be deemed fully and permanently disabled, while this individual might not have had to apply for DI at all if they had a permanent contract. Improvements of the DI system thus also have the potential to affect a large share of the individuals who might apply for DI.

From a policy perspective, this suggests that there is still room for improvements in the design and implementation of the health care and DI systems in order to retain people in the workforce and reduce the number of DI recipients. The analysis in this dissertation has shown that more stringent screening and increased employer incentives have been effective tools. However, the analyses have also shown that despite these tools being in place, a weak labor market attachment can still result in a substantial DI risk. This can be seen in the elevated risks for workers with temporary contracts and UI recipients, but also in the larger impacts of waiting time for mental health treatments on individuals with a migration background and lower educational attainment. Interventions aimed at improving labor market attachment might thus also be effective in reducing DI inflow. Lastly, there still exist considerable disincentives to return to the workforce once DI recipients recover. It should become more favourable to return to work for DI recipients, to ensure that disability benefit receipt becomes a lay-over station instead of an end station.

Bibliography

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). When should you adjust standard errors for clustering? *NBER working paper*, *24003*.

Acemoglu, D., & Angrist, J. (2001). Consequences of employment protection? The case of the Americans with Disabilities Act. *Journal of Political Economy*, *109*, 915–957.

Alexandre, P. K., & French, M. T. (2001). Labor supply of poor residents in metropolitan miami, florida: The role of depression and the co-morbid effects of substance use. *Journal of Mental Health Policy and Economics*, *4*(4), 161–174.

Almond, D., & Doyle, J. J. (2011). After midnight: A regression discontinuity design in length of postpartum hospital stays. *American Economic Journal: Economic Policy*, *3*(3), 1–34.

Autor, D. (2011). The unsustainable rise of the disability rolls in the United States: Causes, consequences, and policy options. *NBER working paper*, *17697*.

Autor, D., & Duggan, M. (2003). The rise in the disability rolls and the decline in unemployment. *The Quarterly Journal of Economics*, *118*(1), 157–206.

Autor, D., & Duggan, M. (2006). The growth in the social security disability rolls: A fiscal crisis unfolding. *Journal of Economic Perspectives*, *20*(3), 71–96.

Autor, D., & Duggan, M. (2014). Moral hazard and claims deterrence in private disaibility insurance. *American Economic Journal: Applied Economics*, *6*(4), 110–141.

Autor, D., Maestas, N., Mullen, K. J., & Strand, A. (2015). Does delay cause decay? The effect of administrative decision time on the labor force participation and earnings of disability applicants. *NBER working paper, 20840*.

Baker, A., Larcker, D. F., & Wang, C. C. (2021). How much should we trust staggered difference-in-differences estimates? *Available at SSRN 3794018*.

Bandelow, B., Michaelis, S., & Wedekind, D. (2017). Treatment of anxiety disorders. *Dialogues in Clinical Neuroscience*, *19*(2), 93.

Bardasi, E., & Francesconi, M. (2004). The impact of atypical employment on individual wellbeing: Evidence from a panel of British workers. *Social Science & Medicine*, *58*(9), 1671-1688.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, *28*(2), 29–50.

Benach, J., Vives, A., Amable, M., Vanroelen, C., Tarafa, G., & Muntaner, C. (2014). Precarious employment: Understanding an emerging social determinant of health. *Annual Review of Public Health*, *35*, 229–253.

Benítez-Silva, H., Disney, R., & Jiménez-Martín, S. (2010). Disability, capacity for work and the business cycle: An international perspective. *Economic Policy*, *25*(63), 483-536.

Berg, G. J. v. d., Hofmann, B., & Uhlendorff, A. (2019). Evaluating vacancy referrals and the roles of sanctions and sickness absence. *Economic Journal*, *129*(624), 3292—3322.

Biasi, B., Dahl, M. S., & Moser, P. (2021). Career effects of mental health. *NBER working paper, 29031*.

Boone, J., & Van Ours, J. (2012). Why is there a spike in the job finding rate at benefit exhaustion? *De Economist*, *160*, 413—438.

Borghans, L., Gielen, A. C., & Luttmer, E. F. (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.

Borusyak, K., Jaravel, X., & Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *Unpublished manuscript*.

Bound, J. (1989). The health and earnings of rejected disability insurance applicants. *The American Economic Review*, *79*(3), 482–503. Retrieved from `http://www.jstor.org/stable/1806858`

Browning, M., Moller Dano, A., & Heinesen, E. (2006). Job displacement and stress-related health outcomes. *Health Economics*, *15*(10), 1061–1075.

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. (2021). Difference-in-differences with a continuous treatment. *arXiv preprint arXiv:2107.02637*.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200-230.

Campbell, D. (2020, Oct). *One in four waiting three months or more for mental health help.* Guardian News and Media. Retrieved from `https://www.theguardian.com/society/2020/oct/07/one-in-four-waiting-three-months-or-more-for-mental-health-help`

Campolieti, M., & Riddell, C. (2012). Disability policy and the labor market: Evidence from a natural experiment in Canada, 1998–2006. *Journal of Public Economics*, *96*(3-4), 306–316.

Card, D., Chetty, R., & Weber, A. (2007). The spike at benefit exhaustion: Leaving the unemployment system or starting a new job? *AEA Papers and Proceedings*, *97*(2), 113–118.

Caroli, E., & Godard, M. (2016). Does job insecurity deteriorate health? *Health Economics*, *25*(2), 131–147.

Caron, C. (2021, Feb). *'nobody has openings': Mental health providers struggle to meet demand.* The New York Times. Retrieved from `https://www.nytimes.com/2021/02/17/well/mind/therapy-appointments-shortages-pandemic.html`

CBS. (2019). *Sociale zekerheid; kerncijfers, uitkeringen naar uitkeringssoort.* (data retrieved from CBS Statline, `https://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=37789KSZ&D1=a&D2=234-274&HDR=T&STB=G1&VW=T`)

Chatterji, P., Alegria, M., Lu, M., & Takeuchi, D. (2007). Psychiatric disorders and labor market outcomes: Evidence from the national latino and asian american study. *Health Economics*, *16*(10), 1069–1090.

Chatterji, P., Alegria, M., & Takeuchi, D. (2011). Psychiatric disorders and labor market outcomes: Evidence from the national comorbidity survey-replication. *Journal of Health Economics*, *30*(5), 858–868.

Chen, S., & Van der Klaauw, W. (2008). The work disincentive effects of the disability insurance program in the 1990s. *Journal of Econometrics*, *142*(2), 757–784.

Claussen, B., Bjørndal, A., & Hjort, P. F. (1993). Health and re-employment in a two year follow up of long term unemployed. *Journal of Epidemiology & Community Health*, *47*(1), 14–18.

Cook, B. L., Trinh, N.-H., Li, Z., Hou, S. S.-Y., & Progovac, A. M. (2017). Trends in racial-ethnic disparities in access to mental health care, 2004–2012. *Psychiatric Services*, *68*(1), 9–16.

Curry, J., Silva, S., Rohde, P., Ginsburg, G., Kratochvil, C., Simons, A., … Mayes, T. (2011). Recovery and recurrence following treatment for adolescent major depression. *Archives of General Psychiatry*, *68*(3), 263–269.

Dano, A. M. (2005). Road injuries and long-run effects on income and employment. *Health Economics*, *14*(9), 955–970.

De Groot, N., & Koning, P. (2016). Assessing the effects of disability insurance experience rating. The case of the Netherlands. *Labour Economics*, *41*, 304–317.

De Groot, N., & Van der Klaauw, B. (2019). The effects of reducing the entitlement period to unemployment insurance benefits. *Labour Economics*, *57*, 195–208.

Deshpande, M., & Li, Y. (2019). Who is screened out? Application costs and the targeting of disability programs. *American Economic Journal: Economic Policy*, *11*(4), 213–248.

Deshpande, M., & Lockwood, L. (2022). Beyond health: Nonhealth risk and the value of disability. *Econometrica*, *90*(4), 82(102593): 1–12.

de Vries, H., Fishta, A., Weikert, B., Sanchez, A. R., & Wegewitz, U. (2018). Determinants of sickness absence and return to work among employees with common mental disorders: A scoping review. *Journal of Occupational Rehabilitation*, *28*(3), 393–417.

Drenik, A., Jaeger, S., Plotkin, P., & Schoefer, B. (2020). Paying outsourced labor: Direct evidence from linked temp agency-worker-client data. *NBER working paper*, *26891*.

Engellandt, A., & Riphahn, R. T. (2005). Temporary contracts and employee effort. *Labour Economics*, *12*(3), 281–299.

Ettner, S. L., Frank, R. G., & Kessler, R. C. (1997). The impact of psychiatric disorders on labor market outcomes. *ILR Review*, *51*(1), 64–81.

French, E., & Song, J. (2014). The effect of disability insurance receipt on labor supply. *American Economic Journal: Economic Policy*, *6*(2), 291–337.

Frijters, P., Johnston, D. W., & Shields, M. A. (2014). The effect of mental health on employment: Evidence from australian panel data. *Health Economics*, *23*(9), 1058–1071.

García Gómez, P., & López Nicolás, Á. (2006). Health shocks, employment and income in the spanish labour market. *Health economics*, *15*(9), 997–1009.

García Gómez, P., Van Kippersluis, H., O'Donnell, O., & Van Doorslaer, E. (2013). Long-term and spillover effects of health shocks on employment and income. *Journal of Human Resources*, *48*(4), 873–909.

Garcia Mandico, S., García Gómez, P., Gielen, A., & O'Donnell, O. (2018). Earnings responses to disability benefit cuts. *Tinbergen Institute Discussion Paper 2018-023/V*.

Gezondheidsraad. (2016). *Verzekeringsgeneeskundige protocollen (in English: "insurance medicine protocols".*

Godard, M., Koning, P., & Lindeboom, M. (2019). Targeting disability insurance applications with screening. *IZA Discussion Paper*, *12343*.

Godard, M., Koning, P., & Lindeboom, M. (2022). Application and award responses to stricter screening in disability insurance. *Journal of Human Resources*, *57*(3).

Godøy, A., Haaland, V. F., Huitfeldt, I., & Votruba, M. (2022). Impacts of hospital wait time on patient health and labor supply. *Unpublished manuscript*.

Goldsmith, D., & Schmieder, J. F. (2017). The rise of domestic outsourcing and the evolution of the German wage structure. *The Quarterly Journal of Economics*, *132*(3), 1165–1217.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254-277.

Gruber, J. (2000). Disability insurance benefits and labor supply. *Journal of Political Economy*, *108*(6), 1162–1183.

Haller, A., Staubli, S., & Zweimüller, J. (2020). Designing disability insurance reforms: Tightening eligibility rules or reducing benefits? *IZA Discussion Paper, 13359*.

Hamilton, V. H., Merrigan, P., & Dufresne, É. (1997). Down and out: Estimating the relationship between mental health and unemployment. *Health Economics*, *6*(4), 397–406.

Hauge, K. E., & Markussen, S. (2021). Early intervention in temporary DI: A randomized natural field experiment reducing waiting time in vocational rehabilitation. *Ragnar Frisch Centre for Economic Research report*.

Hausman, C., & Rapson, D. S. (2018). Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics*, *10*, 533–552.

Hawkins, A., & Simola, S. (2020). Paying for disability insurance? Firm cost sharing and its employment consequences. *Unpublished manuscript*.

Henningsen, M. (2008). Benefit shifting: The case of sickness insurance for the unemployed. *Labour Economics*, *15*(6), 1238–1269.

Hewlett, E., & Moran, V. (2014). Making mental health count. *OECD Health Policy Studies*, 244.

Hofmann, B. (2014). Sick of being "activated". *Empirical Economics*, *47*, 1103–1127.

Howell, D., & Azizoglu, B. (2011). Unemployment benefits and work incentives: The US labour market in the Great Recession. *Oxford Review of Economic Policy*, *27*(2), 221—240.

Hullegie, P., & Koning, P. (2018). How disability insurance reforms change the consequences of health shocks on income and employment. *Journal of Health Economics*, *62*, 134–146.

Ichino, A., & Riphahn, R. T. (2005). The effect of employment protection on worker effort: Absenteeism during and after probation. *Journal of the European Economic Association*, *3*(1), 120–143.

Katz, L. F., & Krueger, A. B. (2019). The rise of alternative work arrangements in the United States. *Industrial and Labor Relations Review*, *72*(2), 382–416.

Kim, I.-H., Muntaner, C., Shahidi, F. V., Vives, A., Vanroelen, C., & Benach, J. (2012). Welfare states, flexible employment, and health: A critical review. *Health policy*, *104*(2), 99–127.

Kinsilla, E. (2021, Sep). *Hailey's partner says he'll see a psychologist when he can afford it. he's been saying that for over a year now.* ABC News. Retrieved from `https://www.abc.net.au/news/2021-09-18/australia-mental-health-wait-times-covid-pandemic/100457162`

Kleven, H. J., & Kopczuk, W. (2011). Transfer program complexity and the take-up of social benefits. *American Economic Journal: Economic Policy*, *3*(1), 54–90.

Koning, P. (2009). Experience rating and the inflow into disability insurance. *De Economist*, *157*(3), 315–335.

Koning, P. (2016). Privatizing sick pay: Does it work? *IZA World of Labor*.

Koning, P., & Lindeboom, M. (2015). The rise and fall of disability insurance enrollment in the Netherlands. *The Journal of Economic Perspectives*, *29*(2), 151–172.

Koning, P., Muller, P., & Prudon, R. (2022a). Do disability benefits hinder work resumption after recovery? *Journal of Health Economics*, *82*.

Koning, P., Muller, P., & Prudon, R. (2022b). Why do temporary workers have higher disability insurance risks than permanent workers? *Tinbergen Institute Discussion Paper, 2022-024/V*.

Koning, P., & Prudon, R. (2023). Sick or unemployed? Examining transitions into sickness benefits at unemployment benefit exhaustion. *Unpublished manuscript*.

Koning, P., & van Sonsbeek, J.-M. (2017). Making disability work? The effects of financial incentives on partially disabled workers. *Labour Economics*, *47*, 202–215.

Koning, P., & Van Vuuren, D. (2010). Disability insurance and unemployment insurance as substitute pathways. *Applied Economics*, *42*(5), 575–588.

Koning, P., & Vethaak, H. (2021). Decomposing employment trends of disabled workers. *B.E. Journal of Economic Analysis and Policy*, *21*(4), 1217-1255.

Korpi, T. (2001). Accumulating disadvantage. Longitudinal analyses of unemployment and physical health in representative samples of the swedish population. *European Sociological Review*, *17*(3), 255–273.

Kostol, A. R., & Mogstad, M. (2014). How financial incentives induce disability insurance recipients to return to work. *American Economic Review*, *104*(2), 624–55.

Kuhn, A., Lalive, R., & Zweimüller, J. (2009). The public health costs of job loss. *Journal of Health Economics*, *28*(6), 1099–1115.

Kyyra, T., Pesola, H., & Verho, J. (2019). The spike at benefit exhaustion: The role of measurement error in benefit eligibility. *Labour Economics*, *60*, 75–83.

Kyyrä, T., & Tuomala, J. (2013). Does experience rating reduce disability inflow? *IZA Discussion Paper*, *7344*.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, *4*(3), 165–224.

Leichsenring, F., & Leibing, E. (2003). The effectiveness of psychodynamic therapy and cognitive behavior therapy in the treatment of personality disorders: A meta-analysis. *American Journal of Psychiatry*, *160*(7), 1223–1232.

Liebert, H. (2019). Does external medical review reduce disability insurance inflow? *Journal of Health Economics*, *64*, 108-128.

Lindeboom, M., Llena-Nozal, A., & van der Klaauw, B. (2016). Health shocks, disability and work. *Labour Economics*, *43*, 186–200.

Lindelow, M., & Wagstaff, A. (2005). *Health shocks in china: Are the poor and uninsured less protected?* The World Bank.

Low, H., & Pistaferri, L. (2015). Disability insurance and the dynamics of the incentive insurance trade-off. *American Economic Review*, *105*(10), 2986–3029.

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.

Maestas, N., Mullen, K. J., & Strand, A. (2013). Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt. *American Economic Review*, *103*(5), 1797–1829.

Marcotte, D. E., Wilcox-Gök, V., & Redmon, D. P. (2000). The labor market effects of mental illness the case of affective disorders. In *The economics of disability.* Emerald Group Publishing Limited.

Marinescu, I., & Skandalis, D. (2021). Unemployment insurance and job search behavior. *The Quarterly Journal of Economics*, *136*(2), 887-931.

Markussen, S., Roed, K., & Schreiner, R. (2017). Can compulsory dialogues nudge sick-listed workers back to work? *The Economic Journal*, *128*, 1276–1303.

Moore, T. J. (2015). The employment effects of terminating disability benefits. *Journal of Public Economics*, *124*, 30–43.

Morris, J. K., & Cook, D. (1991). A critical review of the effect of factory closures on health. *Occupational and Environmental Medicine*, *48*(1), 1–8.

n.d. (2021, June). *Druk op wachtlijsten ggz loopt verder op: 'ik moet 1,5 jaar wachten'.* RTL nieuws. Retrieved from https://www.rtlnieuws.nl/nieuws/nederland/artikel/5237078/wachtlijsten-ggz-zorg-coronacrisis-verwijzingen

Norder, G., Hoedeman, R., De Bruin, J., Van Rhenen, W., & Roelen, C. (2015). Time to recurrence of mental health-related absence from work. *Occupational Medicine*, 574–577.

NZA. (2021, Jan). *Informatiekaart wachttijden ggz 2021.* Retrieved from https://puc.overheid.nl/nza/doc/PUC_648825_22/1/

OECD. (2010). *Sickness, disability and work: Breaking the barriers. A synthesis of findings across OECD countries.* OECD Publishing, Paris.

OECD. (2015). *Temporary employment database.* (data retrieved from, https://data.oecd.org/emp/temporary-employment.htm)

OECD. (2019). *Social expenditure database.* (data retrieved from Social Expenditure Database, https://stats.oecd.org/Index.aspx?DataSetCode=SOCX_DET)

Ojeda, V. D., Frank, R. G., McGuire, T. G., & Gilmer, T. P. (2010). Mental illness, nativity, gender and labor supply. *Health Economics*, *19*(4), 396–421.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204.

Parsons, D. O. (1991). Self-screening in targeted public transfer programs. *Journal of Political Economy*, *99*(4), 859–876.

Paserman, M. D. (2008). Job search and hyperbolic discounting: Structural estimation and policy evaluation. *The Economic Journal*, *118*(531), 1418–1452.

Pettersson-Lidbom, P., & Thoursie, P. S. (2013). Temporary disability insurance and labor supply: Evidence from a natural experiment. *The Scandinavian Journal of Economics*, *115*(2), 485–507.

Prinz, D., & Ravesteijn, B. (2020). Employer responsibility in disability insurance: Evidence from the Netherlands. *Unpublished manuscript*.

Prudon, R. (2023). Is delayed mental health treatment detrimental to employment? *Unpublished manuscript*.

Rambachan, A., & Roth, J. (2019). An honest approach to parallel trends. *Unpublished manuscript*.

Reichert, A., & Jacobs, R. (2018). The impact of waiting time on patient outcomes: Evidence from early intervention in psychosis services in england. *Health Economics*, *27*(11), 1772–1787.

Richards, D. (2011). Prevalence and clinical course of depression: A review. *Clinical Psychology Review*, *31*(7), 1117–1125.

Rijksoverheid. (2022, Sep). *Kabinet zet in op preventie en intensievere samenwerking binnen de zorg.* Retrieved from https://www.rijksoverheid.nl/actueel/nieuws/2022/09/20/kabinet-zet-in-op-preventie-en-intensievere-samenwerking-binnen-de-zorg

Riphahn, R. T., & Thalmaier, A. (2001). Behavioral effects of probation periods: An analysis of worker absenteeism. *Jahrbücher für Nationalökonomie und Statistik*, *221*(2), 179–201.

Roed, K., & Westlie, L. (2012). Unemployment insurance in welfare states: The impacts of soft duration constraints. *Journal of the European Economic Association*, *10*(3), 518–554.

Roelen, C., Norder, G., Koopmans, P., Van Rhenen, W., Van der Klink, J., & Bultmann, U. (2012). Employees sick-listed with mental disorders: Who returns to work and when? *Journal of Occupational Rehabilitation*, 409–417.

Roth, J. (2019). Pre-test with caution: Event-study estimates after testing for parallel trends. *Unpublished manuscript*.

Salm, M. (2009). Does job loss cause ill health? *Health Economics*, *18*(9), 1075–1089.

Schmitz, H. (2011). Why are the unemployed in worse health? The causal effect of unemployment on health. *Labour Economics*, *18*(1), 71–78.

Sentell, T., Shumway, M., & Snowden, L. (2007). Access to mental health treatment by english language proficiency and race/ethnicity. *Journal of General Internal Medicine*, *22*(2), 289–293.

Shapiro, B. T. (2022). Promoting wellness or waste? Evidence from antidepressant advertising. *American Economic Journal: Microeconomics*, *14*(2), 439–77.

Staubli, S. (2011). The impact of stricter criteria for disability insurance on labor force participation. *Journal of Public Economics*, *95*(9-10), 1223–1235.

Sullivan, D., & Von Wachter, T. (2009). Job displacement and mortality: An analysis using administrative data. *The Quarterly Journal of Economics*, *124*(3), 1265–1306.

Tompa, E., Cullen, K., & McLeod, C. (2012). Update on a systematic literature review on the effectiveness of experience rating. *Policy and Practice in Health and Safety*, *10*(2), 47–65.

UWV. (2012). *Statistische tijdreeksen 2012.* (data retrieved from UWV, `https://www.uwv.nl/overuwv/kennis-cijfers-en-onderzoek/statistische-informatie/statistische-tijdreeksen-2012.aspx`)

UWV. (2018). *Statistische tijdreeksen 2018.* (data retrieved from UWV, `https://www.uwv.nl/overuwv/Images/uwv-tijdreeksen-2018.pdf`)

UWV. (2019a). *Kwantitatieve informatie 2018.* (data retrieved from UWV, `https://www.uwv.nl/overuwv/Images/uwv-kwantitatieve-informatie-2018.pdf`)

UWV. (2019b). *Wia uitkering.* (data retrieved from UWV, `https://www.uwv.nl/particulieren/ziek/ziek-wia-uitkering/tijdens-wia-uitkering/detail/hoelang-krijg-ik-wga/hoelang-krijg-ik-loongerelateerde-uitkering-lgu`)

Van Sonsbeek, J.-M., & Gradus, R. H. (2012). Estimating the effects of recent disability reforms in the Netherlands. *Oxford Economic Papers*, *65*(4), 832–855.

Virtanen, M., Kivimäki, M., Joensuu, M., Virtanen, P., Elovainio, M., & Vahtera, J. (2005). Temporary employment and health: A review. *International Journal of Epidemiology*, *34*(3), 610–622.

Von Wachter, T., Song, J., & Manchester, J. (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.

Wagenaar, A. F., Kompier, M. A., Houtman, I. L., van den Bossche, S. N., & Taris, T. W. (2012). Employment contracts and health selection: Unhealthy employees out and healthy employees in? *Journal of Occupational and Environmental Medicine*, *54*(10), 1192–1200.

Weathers, R. R., & Hemmeter, J. (2011). The impact of changing financial work incentives on the earnings of social security disability insurance (SSDI) beneficiaries. *Journal of Policy Analysis and Management*, *30*(4), 708–728.

Williams, J., & Bretteville-Jensen, A. L. (2022). What's another day? The effects of wait time for substance abuse treatment on health-care utilization, employment and crime. *IZA Discussion Papers*, *15083*.

Samenvatting (Summary in Dutch)

In de jaren 80 ontving bijna een op de acht mensen in de Nederlandse beroepsbevolking een arbeidsongeschiktheidsuitkering (destijds WAO). De hoge instroom in en lage uitstroom uit het arbeidsongeschiktheidsstelsel werden beide gezien als een van de grootste socio-economische uitdagingen en leidden tot een reeks van hervormingen tussen 1996 en 2006. Deze hervormingen waren duidelijk effectief in het terugbrengen van de instroom, van ongeveer 1,5% van de werkzame bevolking tot 0,5%. Tegelijkertijd bleef het uitstroompercentage vanuit de arbeidsongeschiktheidsregelingen vrijwel onveranderd.[40] Ondanks deze sterke reductie in instroom, ontvangt nog altijd een aanzienlijk deel (7%) van de Nederlandse beroepsbevolking een arbeidsongeschiktheidsuitkering in 2023. De meerderheid van deze arbeidsongeschikten zal een uitkering blijven ontvangen tot aan hun pensioen. Vanuit beleidsperspectief is het daarom nog steeds belangrijk om te weten welke groepen werknemers de grootste kans hebben om arbeidsongeschikt te worden, wat er kan worden gedaan om te voorkomen dat zij het arbeidsongeschiktheidsstelsel instromen, en – als instroom dan toch plaatsvindt – hoe ontvangers van een AO-uitkering geholpen kunnen worden om weer aan het werk te gaan. Om inzicht te krijgen in deze vragen, volgt deze dissertatie de volledige keten van het ontstaan van gezondheidsproblemen, naar de tweejarige wachttijdperiode van ziekte, de arbeidsongeschiktheidsaanvraag, en het mogelijke herstel van de gezondheidsproblemen en werkhervatting van mensen met een arbeidsongeschiktheidsuitkering.

Hoofdstuk 2 kijkt naar de eerste stap in dit proces voor een groot deel van de AO-aanvragers, namelijk het ontstaan van mentale gezondheidsproblemen en de medische behandeling hiervan. Circa 40% van de AO-aanvragers doet een aanvraag met mentale problemen als hoofddiagnose, maar het aanbod van medische behandelingen voor deze problemen is beperkt. Dit heeft geleid tot lange wachtlijsten in de geestelijke gezondheid-

[40] Arbeidsongeschiktheid (AO) refereert naar zowel de WIA, WGA als IVA-regelingen.

szorg (GGZ). Van de gevolgen van deze wachtlijsten op de kans op werk van wachtenden is echter nog niet veel bekend in de literatuur. Door gebruik te maken van gedetailleerde administratieve gegevens over alle gespecialiseerde GGZ-behandelingen in Nederland wordt daarom de causale impact van wachtlijsten in de GGZ in hoofdstuk 2 geschat.

Omdat de duur die een individu moet wachten voordat hij/zij behandeld wordt afhangt van met name de ernst van de problemen, zou het direct schatten van effecten van wachttijd op werk tot verkeerde uitkomsten leiden. Er wordt daarom gebruik gemaakt van een instrumentele variabelen (IV) aanpak, waarbij de (exogene) variatie in de druk op het GGZ-systeem tussen gemeentes leidend is. De schattingen tonen aan dat er significante en persistente negatieve effecten zijn van wachtlijsten. Als mensen een maand langer moeten wachten voordat ze behandeld worden, daalt de kans dat ze nog werken met twee procentpunt en stijgt de kans dat ze een AO-uitkering ontvangen met één procentpunt. De wachttijden hebben de grootste impact op mensen met een migratieachtergrond of met een lager opleidingsniveau. Het effect van een extra maand wachten is voor hen bijna twee keer zo groot, terwijl zij ook nog eens gemiddeld langer moeten wachten totdat hun behandeling begint.

Hoofdstuk 2 laat zien dat gezondheidsproblemen en de daaropvolgende medische behandeling van belang zijn om de instroom in het AO-stelsel beter te begrijpen. Er zijn echter ook grote verschillen in instroomkansen tussen groepen met vergelijkbare gezondheidsproblemen. Hoofdstukken 3 en 4 kijken daarom naar twee specifieke groepen op de arbeidsmarkt met hoge instroomkansen. Hoofdstuk 3 vergelijkt de kans om een AO-aanvraag te doen voor mensen met een tijdelijk contract met diezelfde kans voor mensen met een vast contract. Deze vergelijking laat zien dat mensen met een tijdelijk contract ongeveer een 50% grotere kans hebben om een aanvraag te doen voor een AO-uitkering. Door administratieve gegevens over arbeidsmarktgeschiedenis, demografische karakteristieken en gezondheidszorgverbruik te combineren, wordt vervolgens een decompositie gemaakt van het verschil in AO-aanvraag risico in de volgende vier mechanismen; (1) selectie van mensen met gezondheidsproblemen naar contracttype, (2) de causale impact van contracttype op de gezondheid van werknemers, (3) werkgeversprikkels tijdens de wachttijd om zieke werknemers te re-integreren, en (4) de arbeidsmarktkansen van zieke werknemers gedurende de wachttijd.

De analyse maakt duidelijk dat er selectie naar contracttype is, voornamelijk op basis van leeftijd en het contracttype. Deze selectie verklaart echter niet het verschil in AO-aanvraagrisico, maar maakt het juist groter omdat mensen met een tijdelijk contract jonger zijn en jongeren een lager AO-risico hebben. Het tweede mechanisme, de causale impact van het contracttype op de gezondheid, verklaart het verschil in AO-aanvraag risico ook niet. Beide contracttypes hebben namelijk een vergelijkbare kans op verslechteringen in mentale en fysieke gezondheid. De prikkels van de werkgever tijdens de wachttijd, en de arbeidsmarktkansen van de zieke werknemer verklaren gezamenlijk wel meer dan 80% van het verschil in AO-aanvraag risico. Zo is het verschil in AO-aanvraagrisico gehalveerd na een hervorming in 2013 waarbij de re-integratieverplichtingen voor werkgevers richting hun zieke tijdelijke werknemers is verhoogd. Het verschil in AO-aanvraag risico dat nog overbleef na de hervorming, kan verklaard worden door de arbeidsmarktkansen van zieke werknemers. Vaste werknemers kunnen altijd terugvallen op hun oude baan, terwijl veel tijdelijke werknemers op zoek zullen moeten naar een nieuwe baan als ze na ziekte weer aan het werk willen. Als het relatief makkelijk is om een nieuwe baan te vinden, bijvoorbeeld in een krappe arbeidsmarkt, is het AO-aanvraag risico vergelijkbaar voor beide contracttype. In sectoren waar geen krapte heerst, hebben zieke tijdelijke werknemers nog altijd een hoger AO-aanvraag risico.

Hoofdstuk 4 zoomt in op de instroom in de ziektewet (ZW) en daaropvolgend arbeidsongeschiktheid voor mensen die het einde van hun werkeloosheidsuitkering (WW) bereiken. Ontvangers van een WW-uitkering hebben een vier keer zo grote kans om arbeidsongeschikt te worden in vergelijking met werknemers met een tijdelijk contract. Het voorportaal van een AO-aanvraag voor deze groep is de Ziektewet. Aanvragen voor de Ziektewet stijgen sterk in de laatste maand waarin uitkeringsontvanger nog WW-gerechtigd zijn. In deze laatste maand ontvangen ze een brief waarin vermeld staat dat, als ze ziek zijn, ze zich nog ziek kunnen melden bij het UWV. Dit roept de vraag op of deze stijging veroorzaakt wordt door relatief gezonde mensen die zich ziekmelden om de totale duur van uitkeringsontvangst te verlengen ("moreel gevaar"), of door mensen die ziek zijn maar zich er niet bewust van waren dat ze zich ook tijdens de WW ziek konden melden (inhalen van niet-gebruik van de Ziektewet). Om deze vraag te beantwoorden, word een vergelijking gemaakt tussen de groep die de Ziektewet instroomt voordat hun WW-uitkering verliep (het "pre-piek cohort") en de groep die op het laatste moment de ziektewet instroomt (het "piek cohort").

Deze vergelijking laat zien dat het piek-cohort niet alleen een grotere kans heeft om in de Ziektewet te komen, maar ook daarin te blijven en uiteindelijk een AO-uitkering te ontvangen. Een vergelijking van de kenmerken van beide cohorten laat vervolgens zien dat mensen in het piek-cohort vaker een migratieachtergrond hebben en een zwakkere arbeidsmarktpositie hadden voorafgaand aan hun instroom in de WW. Het gezondheidszorgverbruik van beide cohorten is echter zeer vergelijkbaar op het moment van ziekmelding. Het is daarom onwaarschijnlijk dat de piek in ZW-instroom veroorzaakt wordt door relatief gezonde mensen die de maximale uitkeringsduur proberen te verlengen ("moreel gevaar"). Het lijkt er eerder op dat de piek in ziekmeldingen veroorzaakt wordt door een inhaalslag van ondergebruik van specifieke groepen die mogelijk al ziek waren maar nog niet wisten dat ze in aanmerking kwamen voor de Ziektewet.

Hoofdstuk 5 kijkt tenslotte naar het laatste deel van het AO-proces, gezondheidsherstel en mogelijke werkhervatting. Ondanks het feit dat door keuringsartsen bij een meerderheid van de AO-ontvangers wordt aangenomen dat ze op enig moment zullen herstellen, zijn de uitstroompercentages uit de regeling en naar werk klein. Om vast te stellen of het arbeidsongeschiktheidsstelsel een barrière vormt om terug te keren naar werk bij herstel, wordt er een vergelijking gemaakt tussen de reactie op herstel voor mensen met en zonder een uitkering. Om ervoor te zorgen dat mensen met en zonder uitkering zo vergelijkbaar mogelijk zijn, worden mensen wiens AO-aanvraag is afgewezen omdat hun verlies in verdiencapaciteit net te klein was, vergeleken met mensen wiens AO-aanvraag is toegekend omdat hun verlies in verdiencapaciteit net groot genoeg was. Door groepen te selecteren die beiden dicht bij de arbeidsongeschiktheidsgrens van 35% zitten, hebben ze vergelijkbare gezondheidspatronen en verschillen ze alleen in het wel of niet hebben van een (partiele) AO-uitkering. Een uitdaging binnen het onderzoek is dat herstel van gezondheid vaak niet waargenomen wordt. Daarom wordt het einde van een gezondheidszorgbehandeling gebruikt als een indicatie voor mogelijk herstel. Voorafgaand aan het einde van de behandeling is de kans op werk van beide groepen bijna identiek. Echter rond het einde van de behandeling gaat een kleiner deel van de groep met een AO-uitkering weer aan het werk. Een "verschil-in-verschil" analyse bevestigt dat het arbeidsongeschiktheidssysteem een barrière vormt om weer aan het werk te gaan voor ongeveer de helft van de WGA-ontvangers op het moment dat hun gezondheid verbetert.

De gemeenschappelijke deler van alle hoofdstukken is dat gezondheidsproblemen sterk bepalend zijn voor de instroom in en uitstroom uit arbeidsongeschiktheid, maar dat contextuele factoren ook zeer belangrijk zijn. Ondanks dat hervormingen hebben geleid tot een selectievere instroom in AO, blijken een adequate ent tijdige voorziening van gezondheidszorg, prikkels en verplichtingen tijdens de wachttijd voor werkgevers, de arbeidsmarktkansen van zieke werknemers en de aanwezigheid van perverse prikkels voor uitkeringsontvangers sterk bij te dragen aan de kans op een uitkering en mogelijke werkhervatting. Dit betekend niet dat in het huidige systeem de "verkeerd" (relatief gezonde) mensen in het AO-systeem instromen. Groepen met een hoger AO-risico door deze contextuele factoren vertonen vaak een vergelijkbare evolutie van gezondheid voorafgaand aan hun AO-aanvraag, vergeleken met groepen met een laag AO-risico. De contextuele factoren lijken dus invloed te hebben op mensen in het hele gezondheidsspectrum. Een voorbeeld; iemand met een tijdelijk contract vraag een AO-uitkering aan en wordt beoordeeld als permanent en volledig arbeidsongeschikt terwijl deze persoon misschien helemaal geen aanvraag had gedaan als diegene een vast contract zou hebben gehad. Verbeteringen van het AO-stelsel hebben dus de potentie om invloed te hebben op een groot deel van de mensen die mogelijk een aanvraag gaat doen.

Vanuit beleidsperspectief betekent dit dus dat er ruimte is voor verbeteringen om deze werknemers te behouden voor de arbeidsmarkt en het aantal arbeidsongeschikten terug te dringen. De analyses in deze dissertatie hebben aangetoond dat strengere screening en sterkere prikkels voor werkgevers effectieve instrumenten kunnen zijn. De analyses tonen echter ook aan dat, ondanks deze instrumenten, een zwakkere arbeidsmarktpositie nog altijd kan resulteren in een hoog risico om arbeidsongeschikt te worden. Dit is terug te zien in het verhoogd AO-risico van tijdelijke werknemers en WW'ers, maar ook in de grotere impact van wachttijden voor GGZ behandelingen voor mensen met een migratieachtergrond en een lager opleidingsniveau. Interventies die gericht zijn op het verbeteren van de arbeidsmarktpositie kunnen dus effectief zijn in het terugdringen van het aantal arbeidsongeschikten. Het laatste hoofdstuk laat zien dat het AO-stelsel het lastig maakt voor arbeidsongeschikten om terug aan het werk te gaan nadat hun gezondheid verbeterd. Het zou daarom gunstiger gemaakt moeten worden om weer aan het werk te gaan, om ervoor te zorgen dat arbeidsongeschiktheid geen eindstation, maar een tussenstation wordt voor veel arbeidsongeschikten.

Tinbergen Institute Research Series

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

787 L.M. TREUREN, *Essays in Industrial Economics - Labor market imperfections, cartel stability, and public interest cartels*

788 D.K. BRANDS, *Economic Policies and Mobility Behaviour*

789 H.T.T. NGUYEN, *Words Matter? Gender Disparities in Speeches, Evaluation and Competitive Performance*

790 C.A.P BURIK *The Genetic Lottery. Essays on Genetics, Income, and Inequality*

791 S.W.J. OLIJSLAGERS, *The Economics of Climate Change: on the Role of Risk and Preferences*

792 C.W.A. VAN DER KRAATS, *On Inequalities in Well-Being and Human Capital Formation*

793 Y. YUE, *Essays on Risk and Econometrics*

794 E.F. JANSSENS, *Estimation and Identification in Macroeconomic Models with Incomplete Markets*

795 P.B. KASTELEIN, *Essays in Household Finance: Pension Funding, Housing and Consumption over the Life Cycle*

796 J.O. OORSCHOT, *Extremes in Statistics and Econometrics*

797 S.D.T. HOEY, *Economics on Ice: Research on Peer Effects, Rehiring Decisions and Worker Absenteeism*
798 J. VIDIELLA-MARTIN, *Levelling the Playing Field: Inequalities in early life conditions and policy responses*
799 Y. XIAO, *Fertility, parental investments and intergenerational mobility*
800 X. YU, *Decision Making under Different Circumstances: Uncertainty, Urgency, and Health Threat*
801 G. ANTONECCHIA, *Productivity and Strategies of Multiproduct Firms*
802 H. KWEON, *Biological Perspective of Socioeconomic Inequality*
803 D.K. DIMITROV, *Three Essays on the Optimal Allocation of Risk with Illiquidity, Intergenerational Sharing and Systemic Institutions*
804 J.B. BLOOMFIELD, *Essays on Early Childhood Interventions*
805 S. YU, *Trading and Clearing in Fast-Paced Markets*
806 M.G. GREGORI, *Advanced Measurement and Sampling for Marketing Research*
807 O.C. SOONS, *The Past, Present, and Future of the Euro Area*
808 D. GARCES URZAINQUI, *The Distribution of Development. Essays on Economic Mobility, Inequality and Social Change*
809 A.C. PEKER, *Guess What I Think: Essays on the Wisdom in Meta-predictions*
810 A. AKDENIZ, *On the Origins of Human Sociality*
811 K. BRÜTT, *Strategic Interaction and Social Information: Essays in Behavioural Economics*
812 P.N. KUSUMAWARDHANI, *Learning Trends and Supply-side Education Policies in Indonesia*
813 F. CAPOZZA, *Essays on the Behavioral Economics of Social Inequalities*
814 D.A. MUSLIMOVA, *Complementarities in Human Capital Production: The Role of Gene-Environment Interactions*
815 J.A. DE JONG, *Coordination in Market and Bank Run Experiments*
816 Y. KIM, *Micro studies of macroprudential policies using loan-level data*
817 S.R. TER MEULEN, *Grade retention, ability tracking, and selection in education*
818 A.G.B. ZIEGLER, *The Strategic Role of Information in Markets and Games: Essays in Behavioral Economics*
819 I. VAN DER WERVE, *Panel data model for socioeconomic studies in crime and education*
820 Y. GU, *Roads, Institutions and the Primary Sector in West Africa*

821 Y. LI, *Share Repurchases in the US: An extensive study on the data, drivers, and consequences*

822 R. DIAS PEREIRA, *What Makes us Unique? Genetic and Environmental Drivers of Health and Education Inequalities*

823 H.P. LETTERIE, *Essays on the regulation of long-term care in the Netherlands*

824 D.D. PACE, *Essays on the cognitive foundations of human behavior and on the behavioral economics of climate change*

825 J.N. VAN BRUMMELEN, *On the estimation of parameters in observation-driven time series models*

826 Z. CSÁFORDI, *Essays on Industry Relatedness*

827 B. VAN OS, *On Dynamic Models: Optimization-Based Methods and Practical Applications*

828 D.T. Ó CEALLAIGH, *Self-control Failures and Physical Inactivity: Measuring, Understanding and Intervening*

829 S.B. DONOVAN, *Ties that bind and fray: Agglomeration economies and location choice*

830 A. SOEBHAG, *Essays in Empirical Asset Pricing*

831 H. YUAN, *Essays in Behavioral Economics*

832 A. LENGYEL, *Essays on Government Bond Markets and Macroeconomic Stabilization*

833 S. KÜTÜK, *Essays on Risk Creation in the Banking Sector*

834 E. VLADIMIROV, *Essays on the Econometrics of Option Pricing*