Journal of Health Economics 82 (2022) 102593

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/jhealeco

Do disability benefits hinder work resumption after recovery?☆

Pierre Koning[a,b,c], Paul Muller[a,b,c,*], Roger Prudon[a,b]

[a] Vrije Universiteit Amsterdam
[b] Tinbergen Institute
[c] IZA, Bonn

ARTICLE INFO

JEL classification:
J08
I1
J22

Keywords:
Disability insurance
Mental health
Labor supply
Health shocks

ABSTRACT

While a large share of Disability Insurance recipients in OECD countries are expected to recover, outflow rates from temporary disability schemes are typically negligible. We estimate the disincentive effects of disability benefits on the response to a (mental) health improvement using administrative data on all Dutch disability benefit applicants. We compare those below the DI eligibility threshold with those above and find that disincentives significantly reduce work resumption after health improves. Approximately half of the response to recovery is offset by benefits. Estimates from a structural labor supply model suggest that disincentives are substantially larger when the worker's earnings capacity is fully restored.

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 and 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 (Autor and Duggan, 2006; Claussen et al., 1993).

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 inflow or the use of schemes.[1] At the same time, studies that do consider the outflow of DI largely ignore the potential importance and consequences of health

☆ We thank Owen O'Donnell, Maarten Lindeboom and seminar participants at the Vrije Universiteit Amsterdam, University of Amsterdam, Gothenburg University, EEA Congress and EALE conference for useful comments. The authors gratefully acknowledge the financial support from Instituut Gak for this research project, as well as UWV for providing access to their data on DI applications. Results in this paper are based on non-public microdata from Statistics Netherlands (CBS). In principle, part of these microdata are accessible for statistical and scientific research. For further information, contact microdatacbs.nl.

* Corresponding author.
E-mail addresses: p.muller@vu.nl (P. Muller), r.e.k.prudon@vu.nl (R. Prudon).

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

https://doi.org/10.1016/j.jhealeco.2022.102593
Received 16 July 2021; Received in revised form 10 December 2021; Accepted 19 January 2022
Available online 30 January 2022

improvements (Campolieti and Riddell, 2012; Koning and van Sonsbeek, 2017; Kostol and Mogstad, 2014; Weathers and Hemmeter, 2011).

As one of the few studies that considers DI outflow due to health improvements, this paper 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 suffers from mood, anxiety or personality disorders – we obtain a plausible measurement of health improvement by considering the end date of a 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 (Bandelow et al., 2017; Curry et al., 2011; Leichsenring and Leibing, 2003; Richards, 2011; Curry et al., 2011; Leichsenring and Leibing, 2003; Richards, 2011). 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.

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.[2] Using detailed data on healthcare trajectories, we show that the probability of recovery is indeed almost identical for both groups.

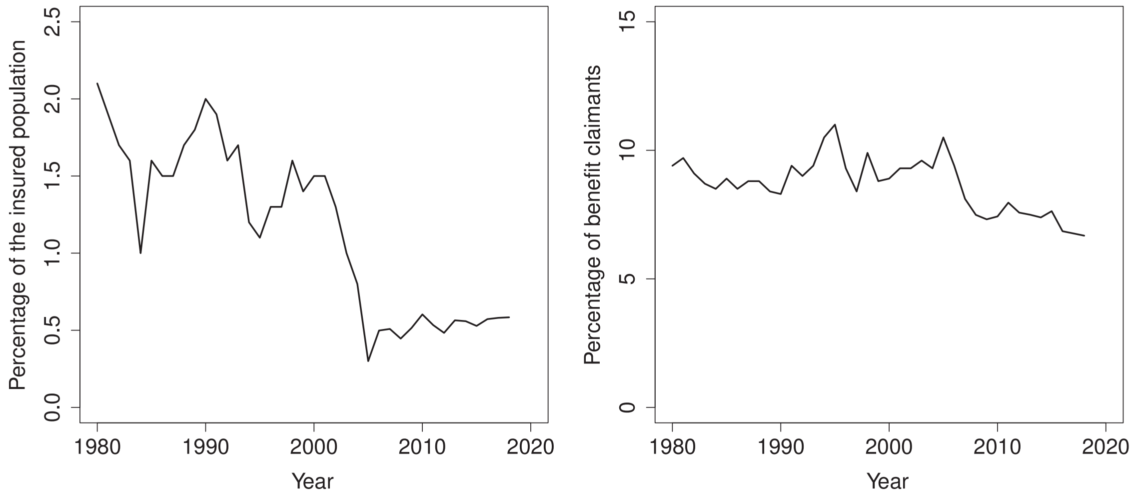
We observe that around the time of recovery the employment rates start to diverge, as those without disability benefits start working at a higher rate than those with disability benefits. The disincentives for work resumption are substantial, amounting to a negative impact of disability benefits of 5.3 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 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 that 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 are 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 by comparing 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 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 and Van der Klaauw, 2008; French and Song, 2014;

[2] 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; Kuhn et al., 2009; Morris and Cook, 1991; Salm, 2009; Schmitz, 2011; Sullivan and Von Wachter, 2009). Again, any biasing effects are mitigated as long as reverse causality is equally strong in both groups.



(a) DI inflow (b) DI outflow

Fig. 1. Inflow and outflow rates of disability benefit programs (UWV 2012, 2018).

Maestas et al., 2013; Von Wachter et al., 2011) and work resumption rates of DI beneficiaries also depend on financial incentives (Campolieti and Riddell, 2012; Koning and van Sonsbeek, 2017; Kostol and Mogstad, 2014; Weathers and Hemmeter, 2011). To understand the lack of outflow from DI schemes and work resumption, it is essential to consider work resumption rates of DI applicants once 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. (2020) 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 and López Nicolás, 2006; Lindelow and 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 5 to 7 percentage-points after a negative health shock.

Lastly, we add to the existing medical literature on the effects of mental health treatments.[3] 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 (Curry et al., 2011; Richards, 2011), personality disorders (Leichsenring and Leibing, 2003) and anxiety disorders (Bandelow et al., 2017)). Treatment is found to be effective for the 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 paper – 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 paper is organized as follows. Section 2 illustrates the institutional background of the Dutch disability insurance system and Section 3 describes the data. Section 4 provides a description of the DiD estimator, presents results and discusses robustness checks. In Section 5 we compare our results to simulations from an estimated structural labor supply model. Section 6 concludes.

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 and 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 Fig. 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.

[3] Section 4.2 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.

Even though some reforms also aimed at increasing work incentives, this drop in inflow rates is accompanied by decreasing outflow rates out of DI that are shown in panel (b).[4]

Since 2006, DI applications can be filed after two years of illness in the Netherlands.[5] 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:

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

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 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.[6] 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.[7]

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 Fig. 2 for income as a function of labor earnings for an individual with 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 someone uses at least 50% of their remaining earnings capacity, the amount of benefits are 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%.[8] 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 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 less 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 margin. 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.[9]

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

[5] 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.

[6] 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.

[7] 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

[8] Note that the structural model in Section 5 also provides further insight in relevant labor supply incentives for partially disabled workers with and without DI benefits.

[9] 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.

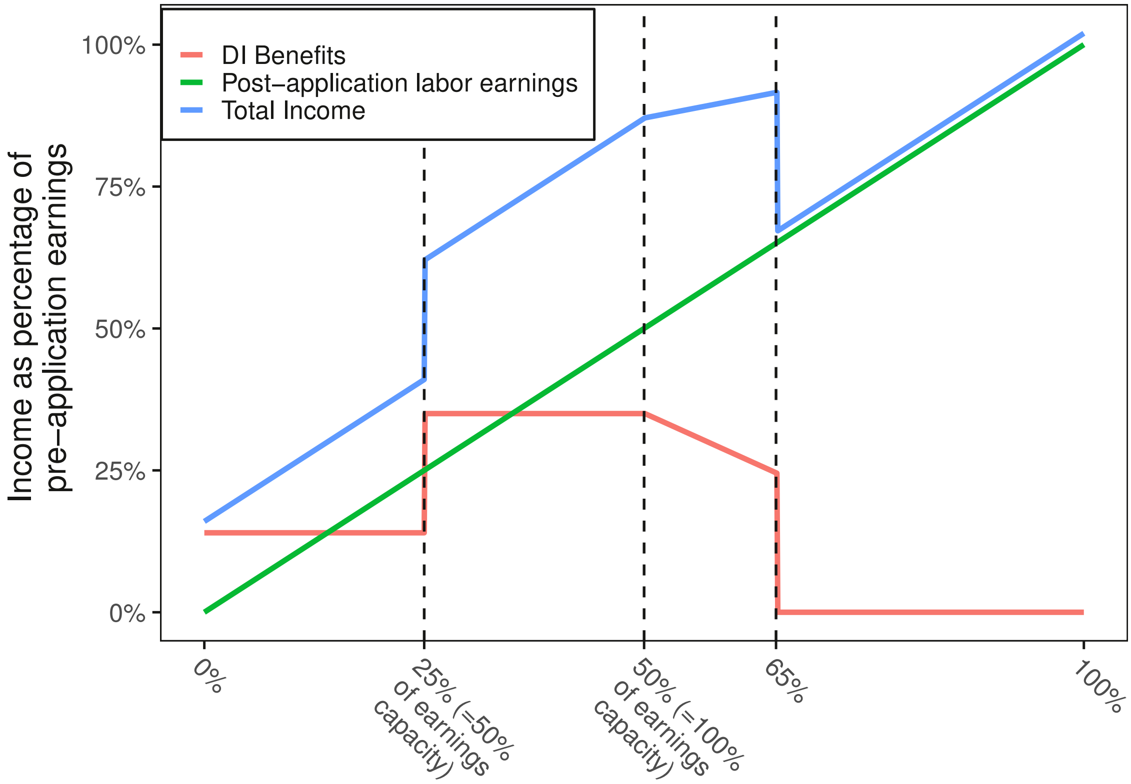


Fig. 2. Income from labor earnings and benefits, as a function of post-application earnings (for an individual with assessed degree of disability of 50%).

3. Data sources and sample selection

We link three data sources to analyze 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.

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 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.[10] 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 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.[11] The CAS-codes are available for 80% of all applications, whereas the UWV codes are available for 98% of all applications. The analysis will, where possible, combine both types of diagnoses 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

[10] 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.

[11] CAS-codes consist of a diagnose 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 thee CAS-codes and three UWV diagnoses group codes.

Table 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

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").

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 individuals' 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.

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" (DBC's) that are used as payment units for complete medical treatments. DBC's comprise all treatments that are deemed necessary to alleviate or solve health problems. We observe all DBC's regarding mental health between 2011 and 2016. Mental health problems in the sample can be considered as severe, with average cost of treatment of approximately € 5,500 and resembling roughly 140 hours 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 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 4.2). 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 the 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.

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 is observed. A selection was made to make the sample suitable for analysis. Table 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.[12] We exclude individuals for whom the end of treatment occurs before the application date. 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). This ensures that our sample 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

[12] 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 4.3, this does not change our results.

Table 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[a]	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.

of 5,003 DI applicants for whom we have a proxy for mental health recovery. The majority of these individuals suffer from mood disorders (40%), personality disorders (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 case of depressions), but it could also be partial recovery that for example restores to some extent the individual's ability to work.[13]

In Table 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.[14] 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 similar assessed probability of health improvement and comparable healthcare expenditures in the year of their DI application.[15]

[13] 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%.
[14] See Online Appendix Table B.4, Table B.3 and Table B.5 for descriptives of groups with specific disorders.
[15] 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.

4. Empirical analysis

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 are 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 Fig. A.9 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).[16] The general event-study specification is as follows:

$$E_{it} = \alpha_{DI} + \alpha_t + \sum_{l=-T+1}^{-1} \beta_l DI_i\, I_{t=l} + \sum_{l=0}^{T} \beta_l DI_i\, I_{t=l} + \theta X_{it} + \varepsilon_{it} \quad (2)$$

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.[17,18] We use a time window of 48 months before and 48 months after recovery in the baseline specification. Individual characteristics X_{it} contain age, gender, nationality, education level and calendar-month fixed-effects.

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 mental health treatments causally increase the probability to recover and to resume work. The medical literature provides considerable evidence in this direction, see e.g. Curry et al. (2011), Leichsenring and Leibing (2003), Bandelow et al. (2017) and de Vries et al. (2018). Taking a broader perspective, these evaluation 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.[19] 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 a 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.[20] For consistent estimation of differences in recovery

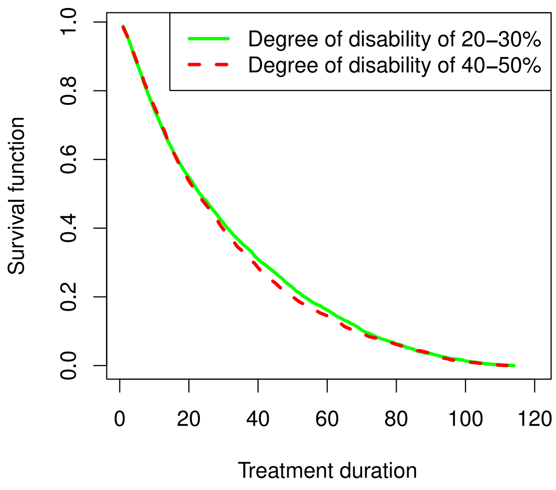
[16] As a robustness check, we also consider larger bandwidths and perform the analysis without the donut around the threshold, using the 25–35% group as control group and the 35–45% as treatment group. Results are similar, as we will show in Table 4.

[17] Given the potential persistence in the outcome variables, we cluster standard errors by individuals (Hausman and 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).

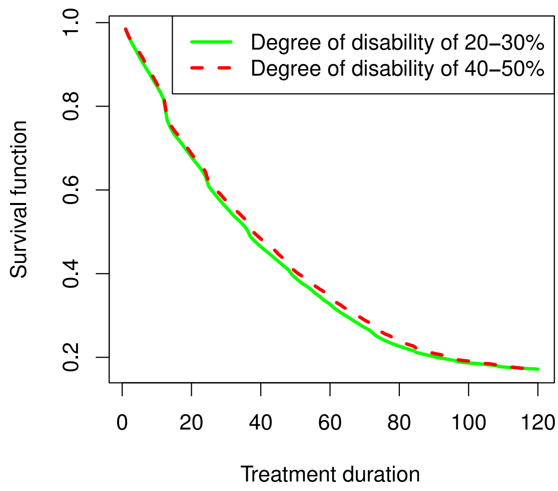
[18] A recent literature has shown that two-way fixed effects estimators are biased in case of staggered treatment implementation and dynamic treatment effects (Borusyak et al., 2021; Callaway and Sant-Anna, 2021; Goodman-Bacon, 2021). In our setting, we compare a single treatment group (those with benefits) 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))

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

[20] 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 $\frac{\beta_2}{\rho}$.



(a) Duration relative to time of application



(b) Duration relative to start of health treatment

Fig. 3. Survival functions of mental health treatment for DI applicants with degree of disability of 20–30% vs. 40–50%.

response effects (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 Fig. 3 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 for this. One such measure is the assessed probability of improvement of UWV, which is very similar for the treatment and control groups (see Table 2). As a second measure, we consider potential differences in the probability that either the treated individuals or the medical doctors 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 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 Fig. 3 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 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.

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.[21] To gauge the validity of this assumption, Fig. 4 shows the trends in employment relative to recovery for the treatment and control group.[22] For both groups, labor supply decreases steadily up to the point of recovery, after which the trend reverses. While

[21] 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 identification of the disincentive effect is not violated.

[22] Trends and event-study estimates for number of working hours are very similar, as shown in Appendix Fig. A.10.

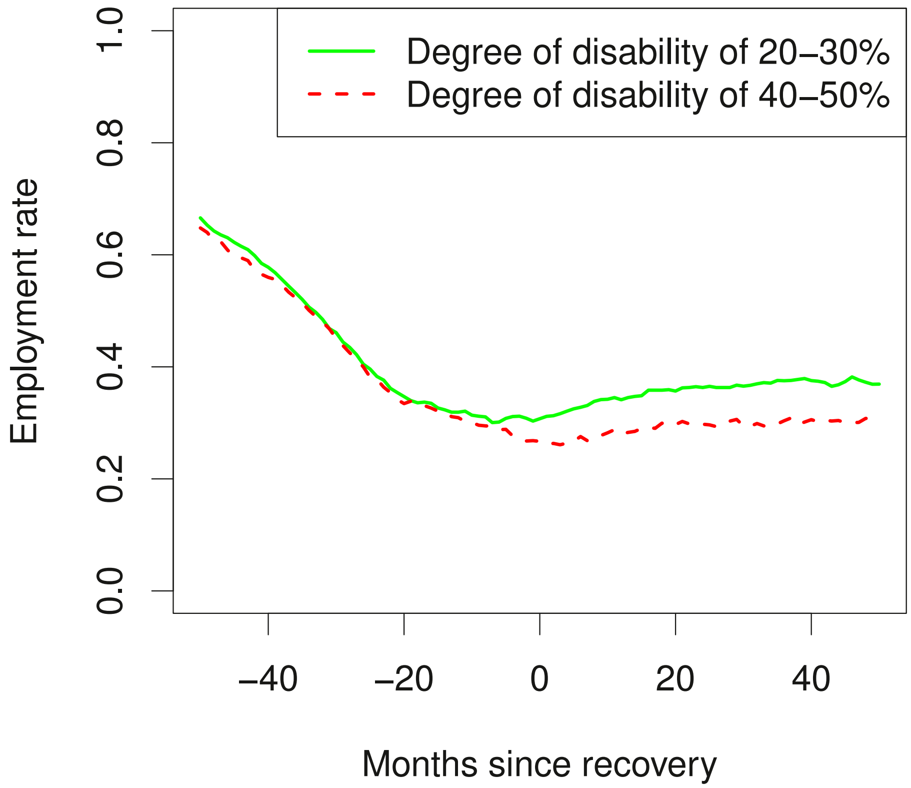


Fig. 4. Employment trends relative to mental health recovery for DI applicants with degree of disability of 20–30% vs. 40–50%.

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 (2) and present the corresponding employment disincentive effects (β_l) in Fig. 5. The results confirm parallel trends before recovery, with non-significant estimates that are close to zero.[23] Employment rates start diverging approximately 6 months prior to recovery, suggesting some anticipation. One year after recovery, the effect of DI benefit receipt on the response to employment has accrued to approximately 5 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 8 working hours. These results are shown in Fig. A.10 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 (2), we find that for individuals without benefits recovery leads to approximately a 10 percentage-point increase in the likelihood of employment; this corresponds to the increase in the green line in Fig. 4 after recovery. The recovery response in working hours for workers without benefits is 11 hours. 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 3. The columns 'Benefit receipt' indicates whether both groups receive DI benefits, and 'Donut' 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)

[23] 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 and Roth (2019), indicates that our estimates are robust to deviations from parallel trends.

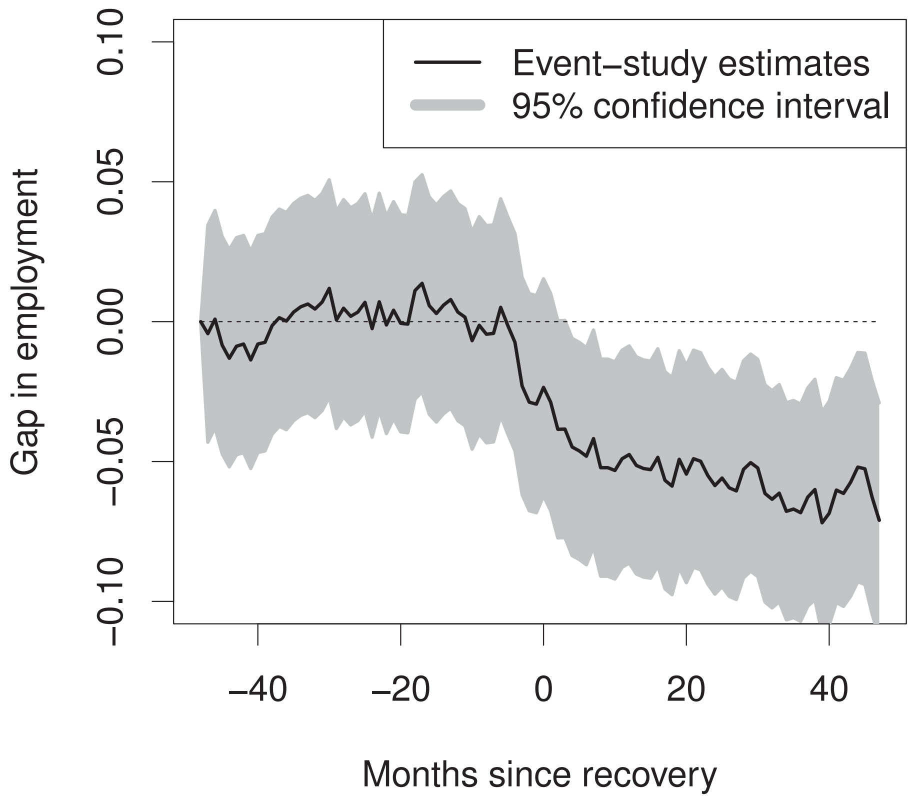


Fig. 5. Event-study estimates (β_l from Equation (2)) mental health recovery for DI applicants with degree of disability of 20–30% vs. 40–50%.

Table 3
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

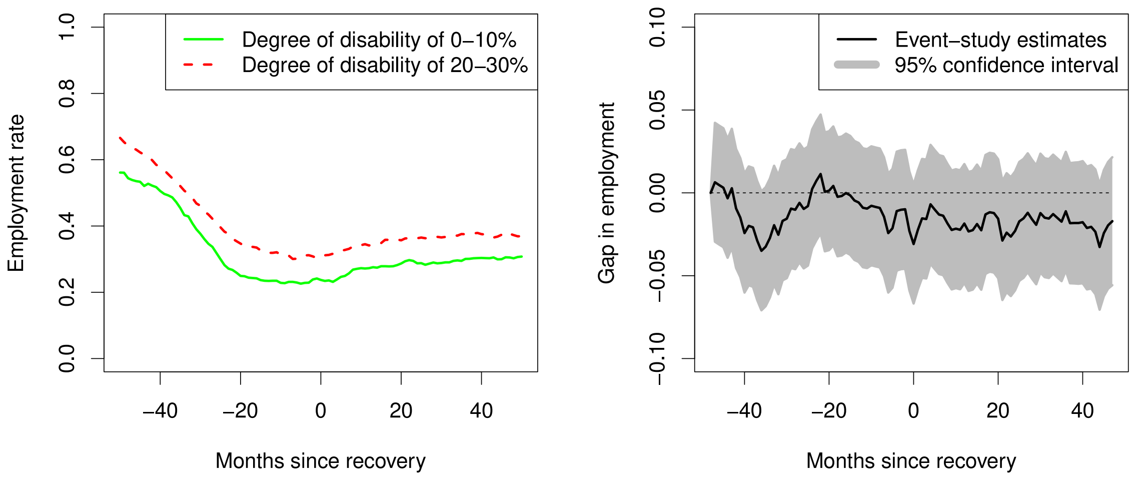
Standard errors in parentheses; * significant at a 5% significance level; **significant after applying the $\frac{1}{45}$ Bonferroni correction factor

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.[24]

The placebo test most similar to the main analysis compares the 0–10% group with the 20–30% group. Fig. 6 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.[25] 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.

[24] 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 Fig. A.13.

[25] See Appendix Figs. A.11 and A.12 for the trends in employment and the corresponding event-study estimates of placebo tests (ii) and (iii).



(a) Employment trends (b) Event-study

Fig. 6. Placebo test. Comparison of employment trends and event-study estimates for groups with equal benefit status but different degrees of disability.

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.[26] 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:

$$E_{it} = \alpha_{DI} + \alpha_t + \beta DI_i \, I_{t\geq 0} + \theta X_{it} + \varepsilon_{it}, \quad (3)$$

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 4 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 4 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, which leaves 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 invidivuals 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 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

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

Table 4
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	–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)

Standard errors shown in parentheses; *significant at a 5% significance level; **significant after applying a $\frac{1}{45}$ Bonferroni correction factor.
[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.

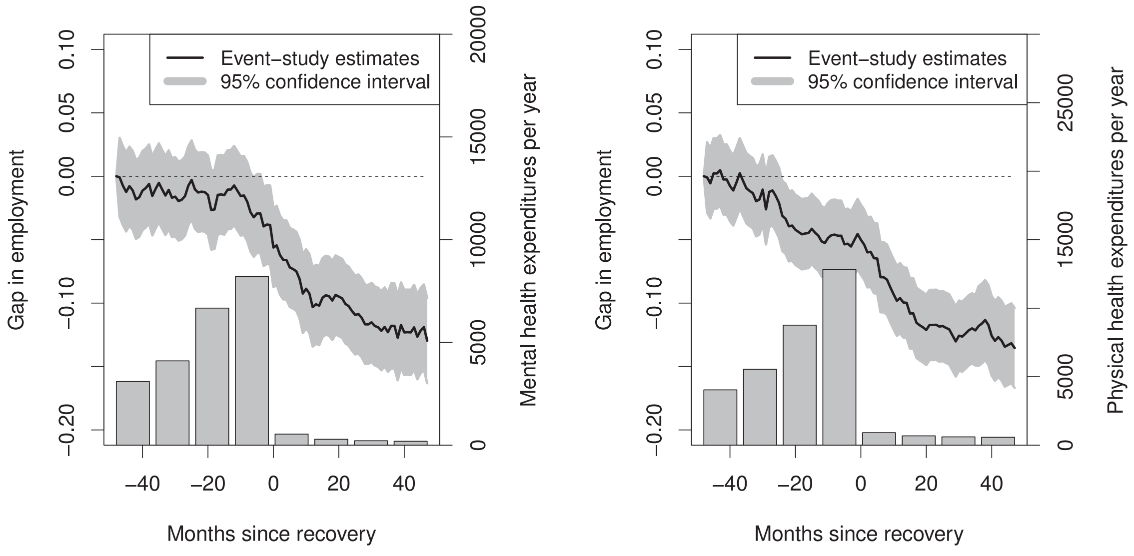
Table 5
Heterogeneity by mental health disorder: DiD estimates for employment and number of working hours.

	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	1,356	669	–0.081**	–10.860**
					(0.002)	(0.213)

Standard errors are shown in parentheses; *significant at a 5% significance level; **significant after applying a $\frac{1}{45}$ Bonferroni correction factor.
[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.

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 Fig. 6, which also find 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.

To shed light on heterogeneous disincentive effects and to make the control and treatment groups more comparable at the same time, Table 5 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 diagnosis, and a separate sample of individuals suffering from disorders other than



(a) Mental health (b) Physical health

Fig. 7. Event-study estimates (β_l) from Equation (2)), using a recovery proxy based on healthcare expenditures.

the three most prevalent ones. Within these sub-samples, the characteristics of the control and treatment group are very similar.[27] 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 disability 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 accumulate 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.

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, health care 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.[28] We distinguish between mental healthcare expenditures and non-mental healthcare expenditures (see Online Appendix Section A.6 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 7418 individuals with mental health-recovery and 9747 individuals with non-mental health-recovery.

For both samples and corresponding types of recovery, Fig. 7 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).[29] 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

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

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

[29] See Appendix Fig. A.14 for the actual trends in employment

Table 6
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 3). Standard errors are shown in parentheses; *significant at a 5% significance level; **significant after applying a $\frac{1}{45}$ Bonferroni correction factor.
[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.

Table 7
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.

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

Table 6 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 the 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.

5. Labor supply effects in a structural model

Our estimation results point at 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 compensate the large employment drop that has occurred in earlier years. We stated earlier that the response effects can be characterized as ITT-effects, since the end of the treatment implies the partial or full recovery only for a part of the sample. This raises the question how large employment effects of full recovery are and what the maximum discouraging impact of DI benefits are.

For insight in 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 and 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.

5.1. 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.[30] Since eligibility to 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.[31] The utility maximization problem for individual i is as follows:

$$\max_{E_i} u(L_i, I_i) = L_i^{\lambda_i+\delta_g} I_i^{1-\lambda_i-\delta_g} \tag{4}$$

$$\text{s.t. } L_i = T - E_i \tag{5}$$

$$I_i(E_i) = E_i w_i + DI(E_i) + \tilde{I}_i \quad \text{if} \quad E_i w_i + DI(E_i) + \tilde{I}_i \geq SA_i \tag{6}$$

$$I_i\big(E_i\big) = (1-F)\, SA_i \qquad \text{if} \quad E_i w_i + DI\big(E_i\big) + \tilde{I}_i < SA_i \tag{7}$$

We set $T = 60$, the highest observed pre-application number of working hours. Income from social assistance is discounted by 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, Fig. 8 shows the choice set for individuals with DI benefits.[32] 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

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

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

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

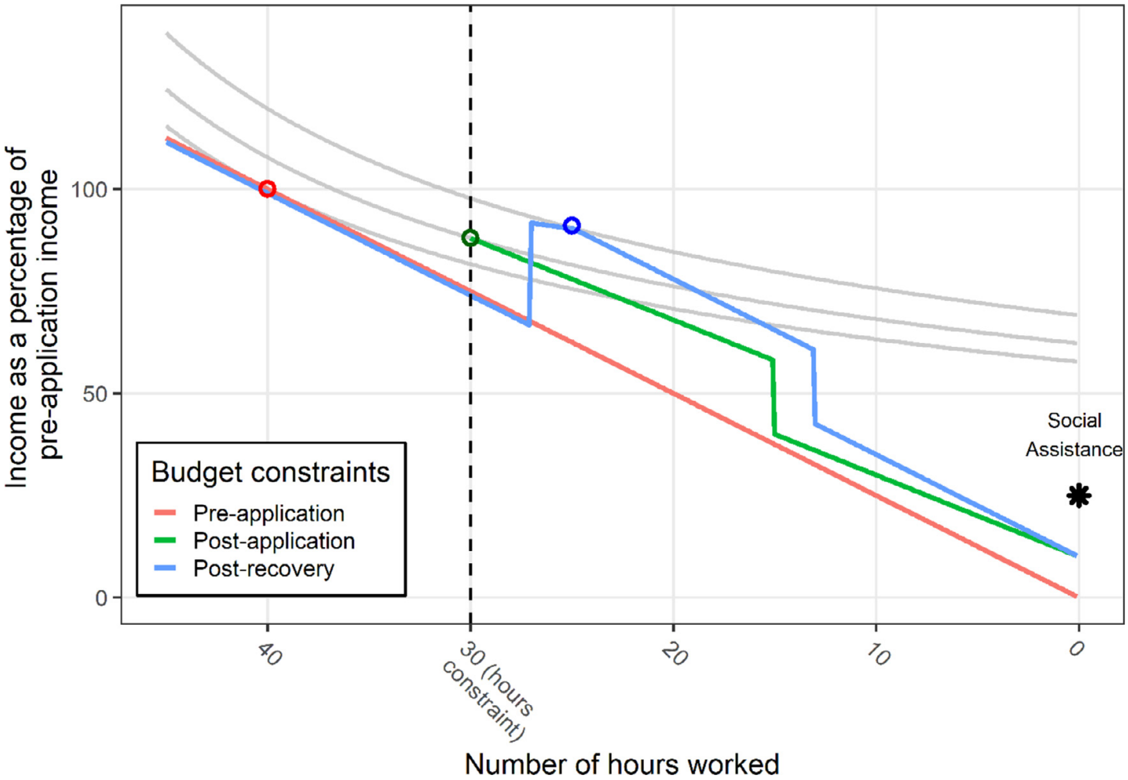


Fig. 8. Budget constraints and utility indifference curves of a fictitious individual awarded partial disability benefits with a degree of disability of 40%.

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 to 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).

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.[33] 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 are predicted and confronted with 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.[34]

[33] 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.

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

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, note that 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.

5.3. 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.[35] This allows us to assess the discouraging impact that benefit receipt may have on work resumption after application and recovery. Table 7 shows the results that follow from this approach.[36]

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.[37] 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 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 4.3, 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 (2) (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.

6. Conclusion

This paper 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.

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

[36] 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 the Online Appendix Section B.4 we have conducted such analyses.

[37] 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 slight higher employment.

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 and Van der Klaauw, 2008; French and Song, 2014; Maestas et al., 2013). 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 1800 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 and 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 and Van der Klaauw, 2008; French and Song, 2014; Maestas et al., 2013)). 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 a capacity constraints, however, such reassessments are rare. By using healthcare data, the limited reassessment capacity could by 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.[38] 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 paper, further research on these potential policy changes is warranted.

Appendix A

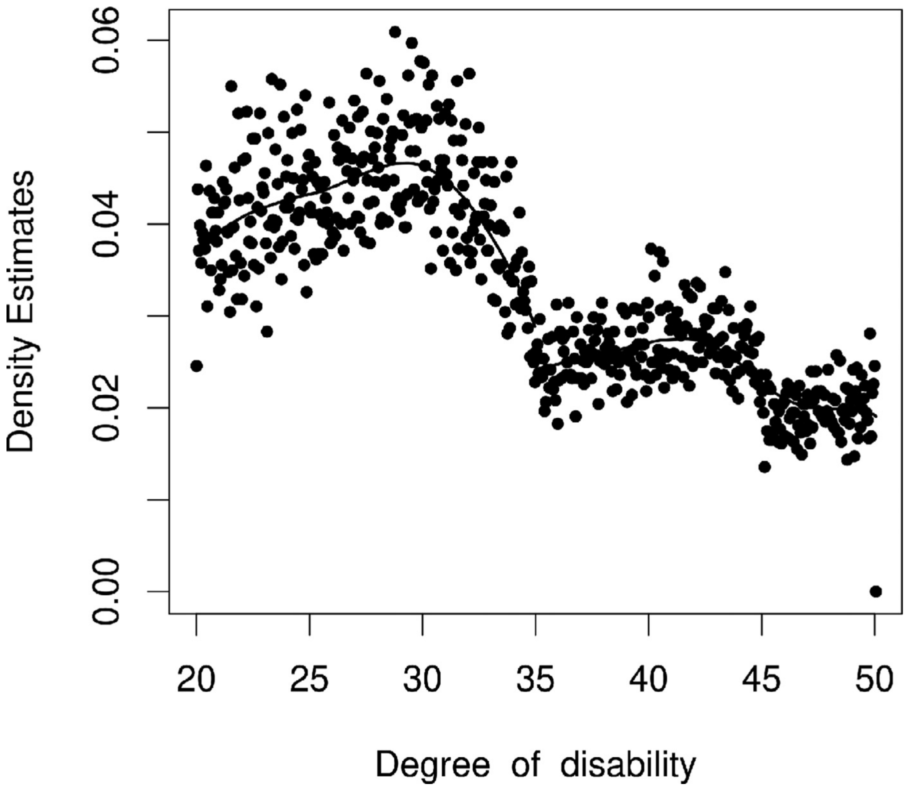
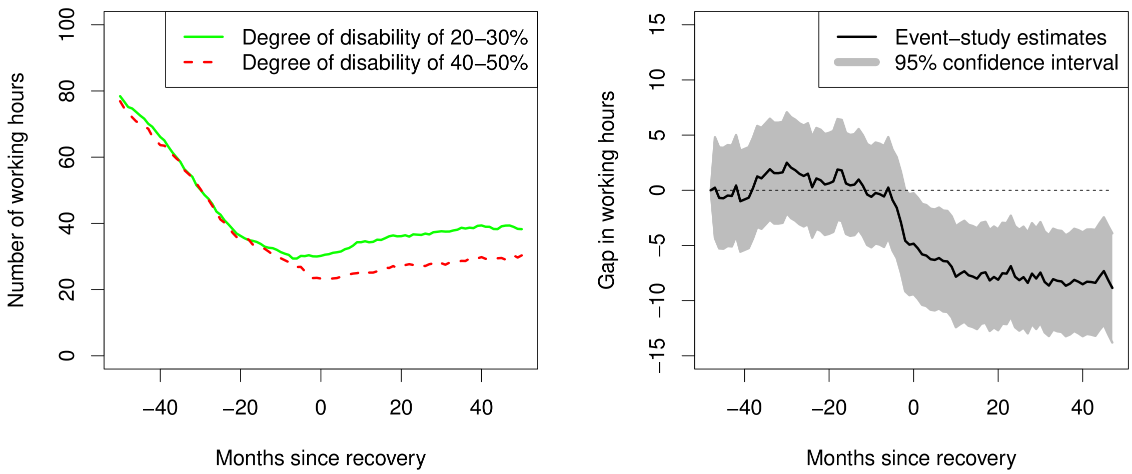


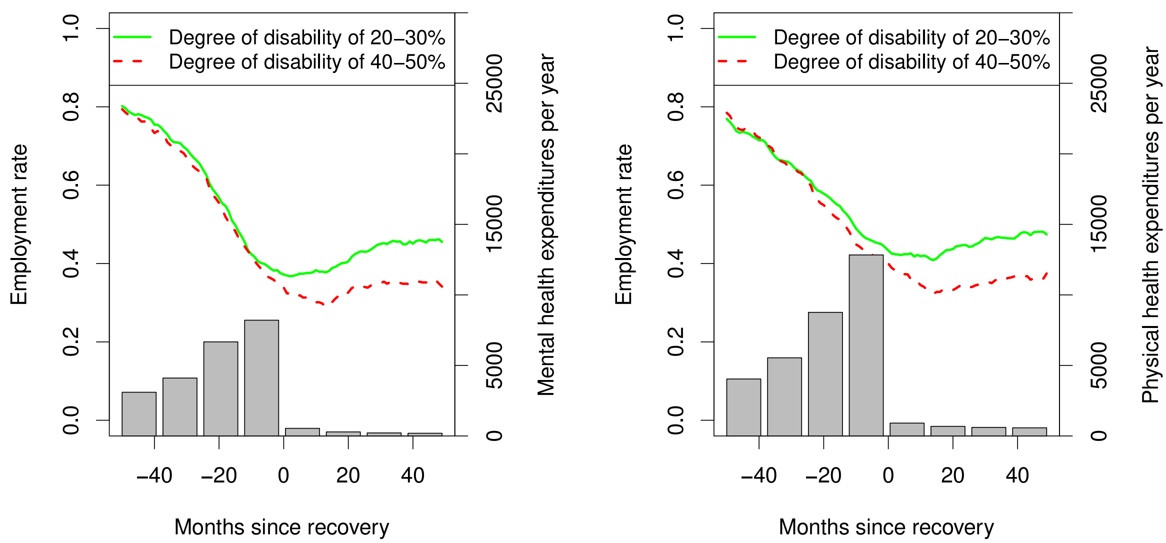
Fig. A1. Discontinuity test of application densities at the 35% threshold of degree of disability.

[38] 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.



(a) Trends (b) Event-study

Fig. A2. 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) Mental health (b) Physical health

Fig. A3. 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%.

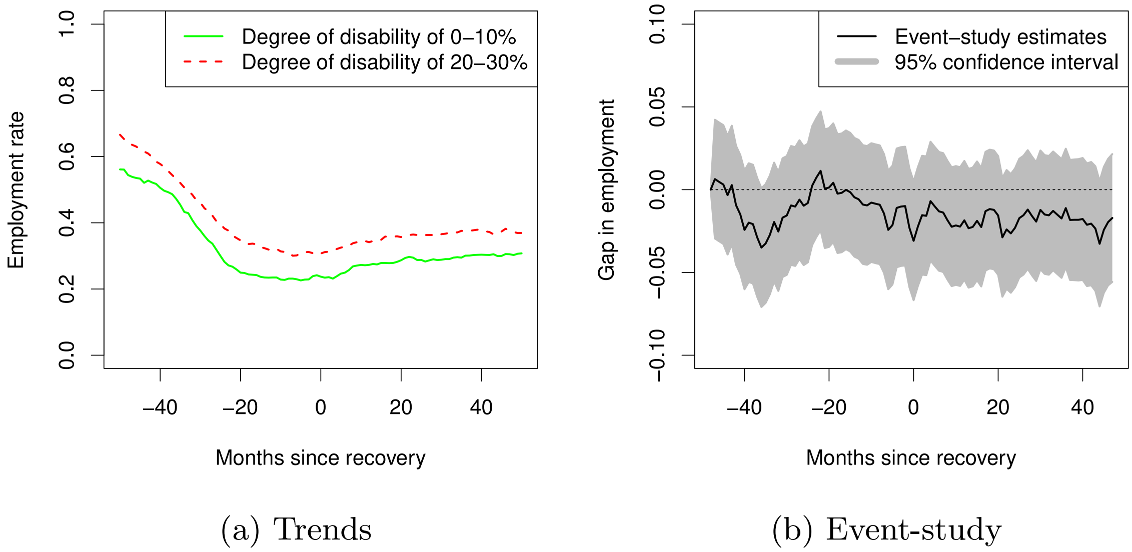


Fig. A4. Trends in employment (a) and event-study estimates (b) relative to mental health recovery for placebo groups with degree of disability of 40–50% vs. 50–60%.

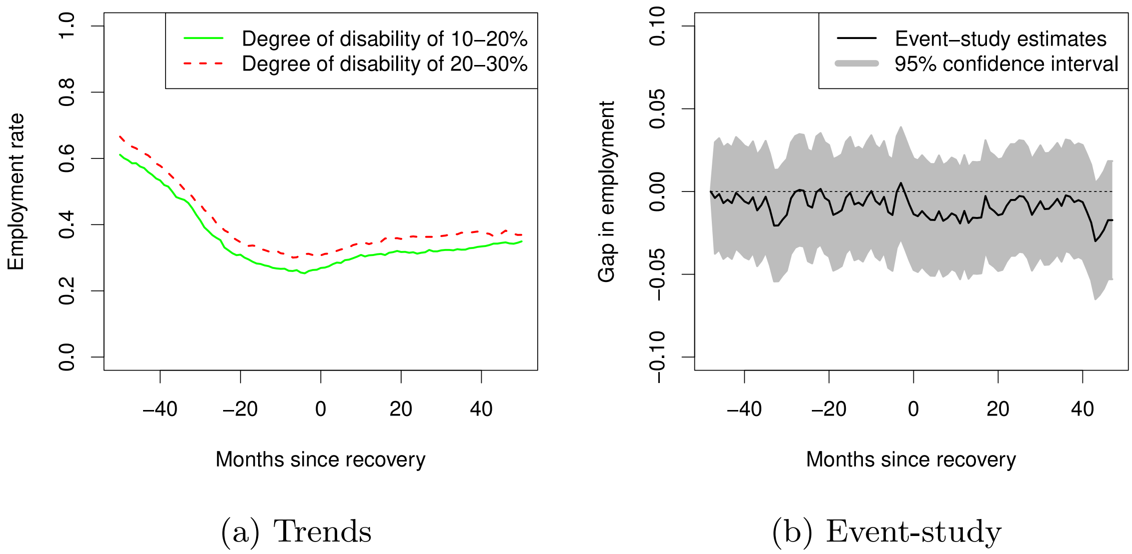


Fig. A5. Trends in employment (a) and event-study estimates (b) relative to mental health recovery for groups with degree of disability of 40–50% vs. 60–70%.

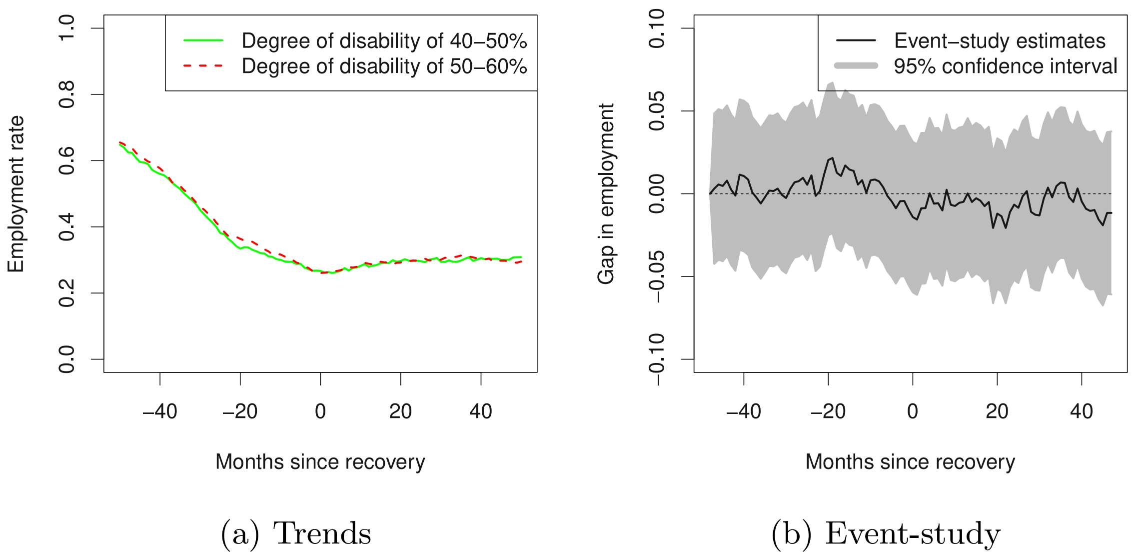


Fig. A6. Employment rates relative to recovery based on healthcare expenditures.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jhealeco.2022.102593

References

Autor, D., Duggan, M.G., 2006. The growth in the social security disability rolls: a fiscal crisis unfolding. Journal of Economic Perspectives 20 (3), 71–96.
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 Clin Neurosci 19 (2), 93.
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.
Bound, J., 1989. The health and earnings of rejected disability insurance applicants. Am Econ Rev 79 (3), 482–503.
Browning, M., Moller Dano, A., Heinesen, E., 2006. Job displacement and stress-related health outcomes. Health Econ 15 (10), 1061–1075.
Callaway, B., Sant-Anna, P.H., 2021. Difference-in-differences with multiple time periods. J Econom 225 (2), 200–230.
Campolieti, M., Riddell, C., 2012. Disability policy and the labor market: evidence from a natural experiment in Canada, 1998–2006. J Public Econ 96 (3–4), 306–316.
CBS, 2019. Sociale zekerheid; kerncijfers, uitkeringen naar uitkeringssoort. CBS Statline.
Chen, S., Van der Klaauw, W., 2008. The work disincentive effects of the disability insurance program in the 1990s. J Econom 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.
Curry, J., Silva, S., Rohde, P., Ginsburg, G., Kratochvil, C., Simons, A., Kirchner, J., May, D., Kennard, B., Mayes, T., 2011. Recovery and recurrence following treatment for adolescent major depression. Arch. Gen. Psychiatry 68 (3), 263–269.
Dano, A.M., 2005. Road injuries and long-run effects on income and employment. Health Econ 14 (9), 955–970.
French, E., Song, J., 2014. The effect of disability insurance receipt on labor supply. American Economic Journal: Economic Policy 6 (2), 291–337.
García Gómez, P., López Nicolás, Á., 2006. Health shocks, employment and income in the Spanish labour market. Health Econ 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., 2020. Earnings responses to disability insurance stringency. Labour Econ 66.
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 No. 12343.
Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. J Econom 225 (2), 254–277.
Gruber, J., 2000. Disability insurance benefits and labor supply. Journal of Political Economy 108 (6), 1162–1183.
Hausman, C., Rapson, D.S., 2018. Regression discontinuity in time: Considerations for empirical applications. Annu Rev Resour Economics 10, 533–552.
Koning, P., Lindeboom, M., 2015. The rise and fall of disability insurance enrollment in the Netherlands. Journal of Economic Perspectives 29 (2), 151–172.
Koning, P., van Sonsbeek, J.-M., 2017. Making disability work? The effects of financial incentives on partially disabled workers. Labour Econ 47, 202–215.
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. Eur Sociol Rev 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), 624–655.
Kuhn, A., Lalive, R., Zweimüller, J., 2009. The public health costs of job loss. J Health Econ 28 (6), 1099–1115.
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: Ameta-analysis. American Journal of Psychiatry 160 (7), 1223–1232.
Lindeboom, M., Llena-Nozal, A., van der Klaauw, B., 2016. Health shocks, disability and work. Labour Econ 43, 186–200.
Lindelow, M., Wagstaff, A., 2005. Health shocks in china: Are the poor and uninsured less protected? World Bank Policy Research Working Paper (3740).
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. Ann Am Acad Pol Soc Sci 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. Am Econ Rev 103 (5), 1797–1829.
Moore, T.J., 2015. The employment effects of terminating disability benefits. J Public Econ 124, 30–43.
Morris, J.K., Cook, D., 1991. A critical review of the effect of factory closures on health. Occup Environ Med 48 (1), 1–8.
Norder, G., Hoedeman, R., De Bruin, J., Van Rhenen, W., Roelen, C., 2015. Time to recurrence of mental health-related absence from work. Occup Med (Chic Ill) 574–577.
OECD, 2019. Social expenditure database.
Pettersson-Lidbom, P., Thoursie, P.S., 2013. Temporary disability insurance and labor supply: Evidence from a natural experiment. Scand J Econ 115 (2), 485–507.
Rambachan, A., Roth, J., 2019. An honest approach to parallel trends. Unpublished manuscript, Harvard University.
Richards, D., 2011. Prevalence and clinical course of depression: A review. Clin Psychol Rev 31 (7), 1117–1125.
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? J Occup Rehabil 409–417.
Roth, J., 2019. Pre-test with caution: Event-study estimates after testing for parallel trends. Unpublished manuscript, Harvard University..
Salm, M., 2009. Does job loss cause ill health? Health Econ 18 (9), 1075–1089.
Schmitz, H., 2011. Why are the unemployed in worse health? the causal effect of unemployment on health. Labour Econ 18 (1), 71–78.
Staubli, S., 2011. The impact of stricter criteria for disability insurance on labor force participation. J Public Econ 95 (9–10), 1223–1235.
Sullivan, D., Von Wachter, T., 2009. Job displacement and mortality: An analysis using administrative data. Q J Econ 124 (3), 1265–1306.
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. Kwantitieve 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.
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. Am Econ Rev 101 (7), 3308–3329.
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. J Occup Rehabil 28 (3), 393–417.
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.