

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11410

Earnings Responses to Disability Benefit Cuts

Silvia Garcia Mandico
Pilar Garcia-Gomez
Anne C. Gielen
Owen O'Donnell

MARCH 2018

DISCUSSION PAPER SERIES

IZA DP No. 11410

Earnings Responses to Disability Benefit Cuts

Silvia Garcia Mandico

Erasmus University Rotterdam and Tinbergen Institute

Pilar Garcia-Gomez

Erasmus University Rotterdam and Tinbergen Institute

Anne C. Gielen

Erasmus University Rotterdam, Tinbergen Institute and IZA

Owen O'Donnell

Erasmus University Rotterdam, Tinbergen Institute and University of Macedonia

MARCH 2018

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Earnings Responses to Disability Benefit Cuts*

Using Dutch administrative data, we assess the work and earnings capacity of disability insurance (DI) recipients by estimating employment and earnings responses to benefit cuts. Reassessment of DI entitlement under more stringent criteria removed 14.4 percent of recipients from the program and reduced benefits by 20 percent, on average. In response, employment increased by 6.7 points and earnings rose by 18 percent. Recipients were able to increase earnings by €0.64 for each €1 of DI income lost. Female and younger recipients, as well as those with more subjectively defined disabilities, were able to increase earnings most. The earnings response declined as claim duration lengthened, suggesting that earnings capacity deteriorates while on DI. The deterioration was steepest for male, younger and fully disabled recipients. Working while claiming partial disability benefits appears to slow the deterioration of earnings capacity.

JEL Classification: H53, H55, J14, J22

Keywords: disability insurance, health, employment, earnings

Corresponding author:

Anne C. Gielen
Erasmus School of Economics
Erasmus University Rotterdam
PO Box 1738
3000 DR Rotterdam
The Netherlands
E-mail: gielen@ese.eur.nl

* We thank Dinand Webbink, Pierre Koning, Gabriele Ciminelli, Songul Togan, Coen Van de Kraats, Eric French, Gaute Torsvik, Andreas Kostol and participants at various seminars and conferences for valuable feedback.

1 Introduction

Stricter screening of disability insurance (DI) applications can reduce entry to the program (Autor and Duggan 2003; Gruber and Kubik 1997; De Jong et al. 2011), but it does nothing to raise the typically low rate of exit. It can slow the growth of spending on DI, but it will take time to reduce the stock of benefit recipients and so lighten the fiscal burden of a program that has become the largest item of social insurance expenditure in many countries. More immediate program savings can be made by cutting the benefits paid to current recipients of DI. However, the fiscal impact of this policy, as well as its consequences for the wellbeing of program beneficiaries, depends on the earnings capacity of the benefit recipients. Only if a sufficient number of them have the potential to replace a substantial fraction of their benefit income with labor market earnings, will it be possible to make cuts to DI without jeopardizing recipients' financial wellbeing and merely shifting the fiscal burden to other social transfer programs. Yet there is little evidence on the earnings capacity of DI recipients and the extent to which they can be induced to realize it through program retrenchment.

Much of the evidence on the earnings crowd-out from DI comes from studies that follow Bound (1989) in using the earnings of rejected applicants to place an upper bound on the earnings capacity of successful applicants (Chen and Van der Klaauw 2008; Von Wachter et al. 2011). Exploitation of plausibly exogenous variation in the award or appeal probability can eliminate upward bias in estimated earnings capacity at the time of application (Autor et al. 2017; French and Song 2014; Maestas et al. 2013). However, this strategy will underestimate the impact of DI on employment and earnings if, as Autor et al. (2015) demonstrate for the US Social Security Disability Insurance (SSDI) program, work capacity and/or preferences decay during the time it takes for an application to be adjudicated. Further, it will overestimate the earnings potential of the stock of beneficiaries if skills and preferences for work deteriorate with time spent on DI. Evidence obtained from comparison of accepted and rejected applicants is pertinent to the impact of policies that tighten entry to DI. It is less

relevant to assessing the potential of reforms that aim to release the earnings capacity of recipients of DI benefits.

This paper provides evidence on the earnings capacity of the universe of Dutch DI benefit recipients aged 30-44 years. It uses administrative data to estimate the impact on employment, earnings and social transfers of a 2004 reform that increased stringency of the DI program, resulting in the termination of the claims of some benefit recipients and substantial cuts in the amounts paid to others. The Netherlands is known for a DI dependency rate that reached 12% of the insured population at the beginning of the 1990s, while also being commended for a series of reforms, including the one we study, that are claimed to have contributed to a two-fifths reduction in DI dependency (Koning and Lindeboom 2015). Officials and analysts in countries, such as the US, searching for ways to manage the escalating fiscal burden of DI can potentially learn from the Dutch experience (Autor 2015; Burkhauser et al. 2014). By examining a reform that occurred a decade into the rationalization process when the dependency rate had already fallen considerably, we deliver evidence that is more relevant to the situation prevailing in other countries than the evaluation of an earlier Dutch reform conducted by Borghans et al. (2014).

We identify the effect of the reform by comparing the pre-post change in earnings (and other outcomes) of benefit recipients aged 30-44, whose DI entitlement was reassessed under stricter criteria, with the respective change among older recipients, who were not reassessed. Unlike other studies that rely on difference-in-differences (DID) across age groups to identify effects of more stringent criteria at application for DI (Karlström et al. 2008; Staubli 2011), we adjust for the difference between the age groups in the outcome trend over a period prior to the reform. This trend-adjusted DID (Bell et al. 1999) eliminates age-specific trends, as well as period effects. Identification rests on the assumption that, in the absence of the reform, the age differential in the outcome trends would have been that observed in the earlier period. Consistent with this, we demonstrate that the age differential in the trends is similar over multiple periods prior to the reform. A placebo test also lends credibility to

the identification: implementing the empirical strategy with data on individuals who are not DI benefit recipients, we find no effect of a pseudo reform on earnings. Applying the same strategy with data on DI recipients, we estimate that the increased stringency induced them to raise their earnings by 18 percent.

The paper makes four contributions to evidence on the work and earnings capacity of DI recipients. First, it is one of only a few studies of younger program beneficiaries. DI recipients below middle age account for an increasing share of DI rolls (OECD 2014). They would be expected to have the greatest earnings capacity and are likely to be the principal target of any attempt to cut benefits. Borghans et al. (2014) estimate earnings capacity at the age of 45 using cuts to DI benefits that were made ten years before the reform we use. We find that the effects on the employment and earnings of younger recipients aged 30-39 are roughly twice the magnitude of the effects on recipients aged 40-44. This is not only because the benefits of younger recipients were cut more aggressively. They also replaced a larger fraction of their lost benefits with earnings. The only other paper that has estimated the employment (but not earnings) response to DI benefit cuts (in this case terminations) among recipients younger than 45 is restricted to examination of the 2% of DI beneficiaries in the US who had qualified for the program, at least in part, through alcohol or drug addiction (Moore 2015).¹ This group is not necessarily indicative of the earnings capacity, and preferences, of the majority of benefit recipients with more prevalent conditions.

Second, precisely because the qualification rules were tightened for all claimants irrespective of the nature of their disability, we are able to compare earnings capacity across DI beneficiaries qualifying through different diagnoses. This is relevant to assessing the credibility of one of the main proposed explanations of lengthening DI rolls. That is, loosening of the criteria for entitlement from precisely defined medical diagnoses to the vaguer concept of *work capacity* opened the door to claims based on health conditions that are difficult to

¹Moore also finds that the employment response is negatively associated with age. Two other papers also find a larger (in magnitude) elasticity of labor supply with respect to DI at younger ages (Kostol and Mogstad 2014; Von Wachter et al. 2011), although these studies do not identify from variation arising from benefit cuts.

verify medically (Autor 2015; Autor and Duggan 2006). Lower back pain and stress-related problems are the stereotypical examples. In 2012 across all OECD countries, individuals suffering from mental health disorders constituted one half of DI benefit recipients and 27-47% of new awards (OECD 2012). Musculoskeletal problems are typically the second most common reason given for a DI claim. If this explanation for the growth of DI programs is correct, then claimants with these more subjectively defined conditions should have greater earnings capacity. Evidence on this hypothesis obtained from comparisons between accepted and rejected DI applicants in the US is contradictory (French and Song 2014; Maestas et al. 2013; Von Wachter et al. 2011).² Consistent with the hypothesis, Moore (2015) finds that recipients with a primary diagnosis of a mental health or a musculoskeletal condition were more likely to work after their benefits were terminated. But, as mentioned above, this paper examines only a small proportion of DI benefit recipients who partly qualified though addiction. We provide the first test of the hypothesis using data on a universe of those receiving DI benefits, who are all subjected to the same tightening of entitlement rules that did not involve any change in the medical criteria for qualification. Consistent with the hypothesis, we find that the earnings responses of those with mental health and musculoskeletal conditions are approximately twice as large as the impact on the earnings of those qualifying for DI through all other diagnoses.

Our third contribution is to add to the meager evidence on whether and how the earnings capacity of DI recipients varies with time on the program. The longer someone is claiming DI, the more their skills may be expected to deteriorate and their work preferences dissipate (Bryngelson 2009; Svensson et al. 2010; Vingård et al. 2004). These negative effects of claim duration are potentially offset by time on DI providing the opportunity for health and work capacity to partially recover from any illness that does not cause permanent disability.

²Von Wachter et al. (2011) find in favor of the hypothesis. Maestas et al. (2013) correct for unobservable differences between accepted and rejected DI applicants and find no evidence that those reporting mental health and musculoskeletal conditions have greater work capacity. In fact, they find the work capacity of applicants with musculoskeletal problems to be lower than average. French and Song (2014), who also correct for unobservables, arrive at the opposite conclusion: those with musculoskeletal problems have greater work capacity.

Moore (2015) finds that the employment response to DI termination first increases with the duration of a claim before the relationship turns negative after about three years, which is consistent with the health recovery effect first dominating before being overcome by the labor market detachment effect. But there is little or no supporting evidence of this inverted U-shaped relationship between work capacity and DI duration. Autor et al. (2015) find that employment and earnings fall even as the time waiting for a US SSDI application to be decided lengthens. Gelber et al. (2017) estimate income effects of SSDI benefits on earnings and find that they vary little with the duration of a claim. Using claim durations of up to 15 years, which is substantially longer than in these other studies, we also do not find evidence of an inverted U relationship. Over the full sample, the earning response to benefit cuts declines continuously with time on DI. The decline is steeper for male and younger benefit recipients.

The final contribution of the paper is to assess whether the option of partial disability, which permits beneficiaries of the Dutch DI program to permanently earn in excess of the equivalent of the substantial gainful activity (SGA) limit beyond which SSDI benefits are terminated in the US, helps recipients to maintain and reach their earnings capacity.³ The provision for partial disability makes the Dutch program a forerunner of return-to-work incentives later introduced or contemplated elsewhere (Kostol and Mogstad 2014). Autor (2015) argues that SSDI should introduce the option of partial disability to reduce the tendency for benefit recipients to become detached from the labor market. If the measure achieves this, then the earnings capacity of the partially disabled who continue to work should decline less with time on DI. Our estimates are consistent with this hypothesis. The positive impact of reassessment under more stringent rules on earnings declines steeply with claim duration for both the fully disabled and the partially disabled who were not working at the time of reassessment. In contrast, among the partially disabled who were working, the effect on the earnings of those who had been on DI for 15 years is as strong as that after

³Among the recipients targeted by the reform we study, approximately half were permitted to earn in excess of an amount equivalent to the SGA without having their benefits reduced.

a claim spell of only a year.

Our results suggest that some Dutch DI benefit recipients had considerable earnings capacity that cuts to their benefit could induce them to exercise. Application of the more stringent rules is estimated to have reduced the amount of DI income received by 20 percent, raised employment by 20 percent and increased earnings by 18 percent, on average. These estimates imply high elasticities of employment and earnings with respect to DI income.⁴ Each €1.00 reduction in disability benefits is estimated to have been replaced by €0.64 of labor earnings. This is similar to the estimate obtained by Borghans et al. (2014) based on cuts to Dutch DI benefits implemented ten years before those we examine. Even after a decade of retrenchment, some recipients of DI still had considerable earnings capacity they could realize to replace around two thirds of substantial losses in benefits. However, it is important to keep in mind that these were the minority of the stock of DI beneficiaries. Most of those who were subjected to reassessment under more stringent criteria did not have their benefits reduced.

The paper proceeds as follows. Section 2 outlines key features of the Dutch DI program and the reform we evaluate. Section 3 sets out our identification strategy. Section 4 describes the data and examines trends in the outcomes. Section 5 presents the results starting with full sample estimates, then a placebo test, then heterogeneous effects and, finally, examination of relationships between earnings capacity and duration of a DI claim. The final section concludes.

⁴The estimated reduction in benefits includes the withdrawal of payments to claimants who were induced, but not forced, by the cuts to leave DI. It overestimates the cut in DI income that arises mechanically from application of the more stringent rules. Consequently, our estimates imply an elasticity of employment with respect to (offered) DI income that is less than -1, and an elasticity of earnings less than -0.9 (=18/-20).

2 Dutch disability insurance

2.1 Eligibility and benefits

The 2004 reform changed the details but not the general procedures for assessing DI eligibility and benefit entitlement. Before describing the reform, we summarize those procedures.

An application to DI for full disability benefits can be submitted after a period of sick pay, which was one year in 2004.⁵ Application for partial disability benefits can be made while in work. The social security administration makes a medical assessment to establish whether the applicant is completely incapable of work. If the agency's physician judges that the applicant has some residual work capacity, then a vocational expert assesses the applicant to identify specific occupations, from a very detailed list, that she is considered capable of performing, taking into account her educational attainment. Earnings capacity is then approximated by the average salary across the three highest paying of those occupations in which there are jobs available.⁶ *Degree of disability* is defined as the proportionate shortfall of this earnings capacity from pre-disability earnings. If the estimated degree of disability is below a threshold, which in 2004 was 15%, then the claim is rejected.⁷ A degree of disability of at least 80% corresponds to fully disability and maximum benefits. The claimant is compensated for approximately 70% of her lost earnings capacity.⁸

The benefit recipient is permitted to do paid work without the loss of benefits but only up to the maximum earnings consistent with her assessed degree of disability. Earning above that maximum results in downward revision of the degree of disability and a reduced benefit payment. After leaving DI, benefits continue to be received during a three-month trial period before entitlement is terminated.

Prior to the reform, outflow from DI was low. Work capacity was reassessed one year

⁵After 2004, the sick pay period was extended to two years.

⁶The salary attached to each occupation is the average paid to those engaged in it.

⁷The minimum degree of disability was increased to 35% in 2006 for new applicants but not existing recipients.

⁸Specifically, the replacement rate is set at 70% of the mid-point of each of the intervals of the degree of disability shown in Appendix A, Table A1.

after a claim was awarded and every five years thereafter. These reassessments were often based on no more than the recipient's response to a postal questionnaire.

2.2 The 2004 Reform

From October 2004, the stock of about 275,000 DI benefit recipients younger than 50 on July 1, 2004 became eligible for reassessment under more stringent criteria.⁹ The outcome could be complete or partial withdrawal of benefits.

Benefit recipients were required to undergo medical assessment by a physician. All had their degree of disability re-evaluated under stricter rules that could result in upward revision of current earnings capacity and downward revision of pre-disability earnings. For any given health condition and associated functional limitations, degree of disability under the new rules could not be greater than it was under the old rules. In many cases, it would be lower and result in a reduction or termination of benefits.

The main reason benefit entitlement could be revised downward was that earnings capacity was now estimated by averaging over the three highest paying occupations the claimant was considered able to perform that each had at least three job vacancies. Previously, only occupations with at least ten vacant jobs had qualified. In addition, jobs requiring Dutch language proficiency and knowledge of information technology were now considered feasible even if the claimant did not possess those skills. Full-time employment and night work were now also considered feasible even if the claimant had not previously engaged in those types of employment. As a result of these expansions of the pool of potential work, the average wage over the three highest paying occupations considered feasible was likely to rise. It could not fall. A rise in assessed earnings capacity meant a fall in benefit entitlement. In addition, pre-disability earnings could be reduced by a new rule that truncated weekly hours at a maximum of 38. If earnings had been inflated by previously working more than this,

⁹The reform was legislated in April 2004 and the initial plan was to start the reassessments from July 2004. Because of strong political opposition and lack of consensus about the reassessment criteria, implementation was pushed back to October. Analysis in section 4.3 of trends in employment and earnings prior to the start of the reassessments does not reveal patterns consistent with anticipation effects on these outcomes.

then there would be a downward revision of lost earnings capacity, and so benefits.

These changes resulted in around one third of all reassessed benefit recipients having their benefits reduced or terminated (Social Insurance Benefits Agency (UWV) 2009). About a fifth had their degree of disability reassessed to be below the 15% minimum threshold. Consequently, their DI entitlement was withdrawn completely. Another 12% were allowed to remain on DI but with lower benefits. Over three fifths (62%) experienced no change in their entitlement. Deterioration in health since the previous assessment resulted in 6% of recipients having their benefits raised despite application of the more stringent rules (see Appendix A Table A2).

If the outcome of reassessment was a downward revision in the degree of disability, then benefits were reduced or terminated two months later. If employment was not secured, a disqualified DI recipient could transfer to unemployment insurance (UI) if she was still eligible for that program. If not, or if the disqualified claimant had less than six months of UI entitlement remaining, then application could be made to a temporary program put in place specifically to cushion the short term impact of the reform. This maintained DI income at the same level for a period of six months (increased to twelve months in 2007). Around 18% of claimants whose entitlements were reduced or terminated were granted benefits from this program (Social Insurance Benefits Agency (UWV) 2009).

The reassessments were undertaken between October 2004 and the end of 2008. Initially, the plan was to reassess all younger benefit recipients before moving to older groups, but this was not adhered to. In 2007, strong criticism of the policy and a change of government resulted in the age threshold for reassessment being revised from less than 50 to less than 45 on July 1, 2004. As a result, around 25,000 recipients aged 45-49 who had already been reassessed were assessed once more under the old, more lenient rules (Ministry of Justice 2007). Consequently, we restrict attention to benefit recipients aged 30-44 on July 1, 2004.

3 Identification & Estimation

3.1 Identification

We estimate labor supply responses to reassessment under more stringent entitlement rules, and use these to infer the employment and earnings potential of benefit recipients. To estimate the average effect of reassessment on recipients aged 30-44, we need a comparison group, or groups, that allows credible identification of the average outcome that would have materialized in the target group if the reform had not been implemented.

Let Y_{it} be the observed outcome of individual i at time t , and let Y_{it}^1 and Y_{it}^0 represent potential outcomes with and without reassessment respectively. Let $t=0$ indicate some time before the commencement of reassessments, such that $Y_{i0} = Y_{i0}^0 \quad \forall i$. In our main analysis, we use annual data and $t=0$ corresponds to 2004. Let $t=4$ be four years later in 2008 when the reassessments were completed. Then, $Y_{i4} = D_i Y_{i4}^1 + (1 - D_i) Y_{i4}^0$, where $D_i = 1$ if i has been reassessed and is 0 otherwise. We wish to estimate the average effect of reassessment on those reassessed: $ATE = \mathbb{E}[Y_{i4}^1 - Y_{i4}^0 \mid D_i = 1]$.

One identification strategy would rely on a difference-in-differences (DID) comparison between younger benefit recipients (30-44 on July 1, 2004) who were subject to reassessment and older recipients (50+ on July 1, 2004) who were not.¹⁰ This is likely to be problematic, particularly as the age gap widens. Older DI beneficiaries have a lower probability of returning to work and recovering their earnings than younger recipients, even when the latter are not subject to reassessment.

An alternative comparison group would be DI recipients who are the same age as those targeted by the reform but who are observed in a period that ends before the reassessments begin. If in the absence of the reassessments the mean outcome of this age group would have changed during the period in which the reform took place by the amount observed for

¹⁰Those aged 45-49 on July 1, 2004 are not useful either as a treatment group or a comparison group since some of them were first reassessed under the new, stricter rules and then (after 2007) assessed once again under the initial, more lenient rules.

the same age group in the earlier period of equal length, then the DID across periods will identify the average effect of the reassessments. The threat to this identification strategy comes from period-specific labor market conditions and any earlier changes in DI that would invalidate using the earlier period to identify counterfactual employment and earnings of the target age group in the reform period.¹¹

The strategy we adopt makes use of both comparison groups – older benefit recipients in the same period and recipients of the same age in an earlier period – to identify the impact of increased stringency under an assumption that is plausibly (although not necessarily) weaker than each assumption required to construct the counterfactual from one of the two comparison groups alone. We use a four-year interval running from 1999 to 2003 ($PERIOD_i = 0$) that precedes the reform to identify the extent to which the trend in the average outcome of younger DI recipients aged 30-44 ($AGE_i = 1$) differs from the trend of older recipients, whom we define as aged from 50 to 53 ($AGE_i = 0$). Effectively, we subtract this differential trend in the *non-reform period* from the DID of the outcome between the age groups over the four-year *reform period* from 2004 to 2008 ($PERIOD_i = 1$) during which the younger age group was reassessed. This differential trend adjusted difference-in-differences (DADID) (Bell et al. 1999; Blundell and Costa Dias 2002) relaxes the assumption of common trends in earnings (/employment) across age groups in the absence of the reform. It also avoids the assumption of common trends in earnings for a given age group across periods. The assumption that is required for identification of the *ATE* is that the age differential in the

¹¹One DI reform implemented in 2002 is credited with having substantially reduced the rate of inflow into DI (De Jong et al. 2011; Van Sonsbeek and Gradus 2012; Koning and Lindeboom 2015). It made the employer and the employee jointly responsible for taking active measures to enable the latter to continue working during the waiting period for DI. Any impact on the DI exit rate, as well as the employment and earnings of existing DI claimants, which are relevant in the present context, would be indirect. Such effects cannot be entirely ruled out and so comparing 30-44 year old DI recipients in 2004-2008 with their counterparts in an earlier period spanning 2002 could possibly be problematic. However, there is no reason to expect the 2002 reform to have affected younger recipients differentially from older recipients. Provided it did not, then it does not jeopardize the validity of the identification strategy we use.

trends in earnings would have been common across periods in the absence of the reform:

$$\begin{aligned} & \mathbb{E} [Y_{i4}^0 - Y_{i0}^0 \mid AGE_i = 1, PERIOD_i = 1] - \mathbb{E} [Y_{i4}^0 - Y_{i0}^0 \mid AGE_i = 0, PERIOD_i = 1] \quad (1) \\ & = \mathbb{E} [Y_{i4}^0 - Y_{i0}^0 \mid AGE_i = 1, PERIOD_i = 0] - \mathbb{E} [Y_{i4}^0 - Y_{i0}^0 \mid AGE_i = 0, PERIOD_i = 0] \end{aligned}$$

If this assumption holds, then any widening of the age differential in the trends that occurs in the *reform period* relative to the *non-reform period* can be attributed to a positive impact of reassessment on the earnings of younger benefit recipients. The average effect of reassessment on those reassessed is then given by the DADID:

$$\begin{aligned} & \mathbb{E} [Y_{i4} \mid AGE_i = 1, PERIOD_i = 1] - \mathbb{E} [Y_{i0} \mid AGE_i = 1, PERIOD_i = 1] \\ & - \left(\mathbb{E} [Y_{i4} \mid AGE_i = 0, PERIOD_i = 1] - \mathbb{E} [Y_{i0} \mid AGE_i = 0, PERIOD_i = 1] \right) \\ & - \left\{ \left(\mathbb{E} [Y_{i4} \mid AGE_i = 1, PERIOD_i = 0] - \mathbb{E} [Y_{i0} \mid AGE_i = 1, PERIOD_i = 0] \right) \right. \quad (2) \\ & \left. - \left(\mathbb{E} [Y_{i4} \mid AGE_i = 0, PERIOD_i = 0] - \mathbb{E} [Y_{i0} \mid AGE_i = 0, PERIOD_i = 0] \right) \right\} \end{aligned}$$

In section 4.3, we assess the plausibility of the assumption that the age differential in the earnings (/employment) trends would have been the same across the periods if no reassessments had been conducted between 2004 and 2008 by comparing age differences in outcome trends across periods in which there was no reform. If the reform was anticipated by benefit recipients who reacted by leaving DI and entering employment already in 2004, then our strategy will deliver lower bound estimates of its impact. But examination of pre-reform trends does not reveal patterns across the outcomes indicative of anticipation. In section 5.2, we further assess the credibility of the strategy by checking that it gives a zero *effect* on the earnings (/employment) of individuals who were not DI benefit recipients and so were not exposed to the reform.

3.2 Estimation

To estimate the effects, we pool balanced panels of DI recipients from the reform and non-reform periods. At entry to the panel, which is January 1, 2004 and January 1, 1999 for the reform and non-reform periods respectively, every observation is receiving DI benefits. In the reform period panel, the treated recipients are aged 30-44 on July 1, 2004. The comparison group obtained from this panel is aged 50-53 on July 1, 2004. In the non-reform period panel, we distinguish between those aged 30-44 and those aged 50-53 on July 1, 1999.

We use least squares to estimate fixed effects models with the following structure,

$$\begin{aligned}
 Y_{it} = & \sum_{t=1}^4 \beta_t AGE_i \times PERIOD_i \times YEAR_t + \sum_{t=1}^4 \theta_t YEAR_t \\
 & + \sum_{t=1}^4 \gamma_t AGE_i \times YEAR_t + \sum_{t=1}^4 \delta_t PERIOD_i \times YEAR_t + \mu_i + \varepsilon_{it},
 \end{aligned} \tag{3}$$

where $YEAR_t$ is an indicator of the within panel year of the observation, such that $YEAR_0 = 1$ & $PERIOD_i = 1$ indicates 2004, $YEAR_0 = 1$ & $PERIOD_i = 0$ indicates 1999 and $YEAR_4 = 1$ indicates 2008 or 2003 depending on the value of $PERIOD_i$, μ_i is an individual fixed effect and ε_{it} is an idiosyncratic error. In addition to period effects and age effects that differ between the periods, both of which are captured by the fixed effects, this model allows within panel time effects (θ_t) that differ across age groups (γ_t) and periods (δ_t). The period-specific level effects and trends allow for the fact that the periods 1999-2003 and 2004-2008 span different phases of the business cycle. Growth was decelerating in the earlier period and accelerating in the later period. The age-specific trends allow for the possibility that, within each period, average earnings (employment) of the younger group of DI recipients does not move in parallel to that of the older group.

Subject to the identification assumption (1), β_t corresponds to the average effect of the reassessments t years after they started. Prior to $t = 4$, corresponding to 2008 in the reform period, these are intention to treat effects since not all benefit recipients in the target group aged 30-44 had been reassessed before then. The evolution of these intention to treat effects is

not so interesting - it simply reflects the cumulative increase in the number of reassessments (See Appendix A Table A3). We focus on the estimate of β_4 , which corresponds to the *ATE* when the reassessments had been completed.

Taking differences from 2004 introduces a slight inaccuracy because 1% of reassessments were carried in the last quarter of 2004. Further, while effectively all recipients aged 30-44 had been reassessed by the end of 2008,¹² around 3% were reassessed during that year. The full effect of reassessment on these recipients may not be reflected in earnings averaged over 2008. To allow for both potential inaccuracies, we test robustness of the main estimates to using monthly data that allow us to take differences between September 2004 and December 2008.¹³

4 Data

4.1 Sources and measures

We obtain data on all recipients of DI benefits from social security files, which record degree of disability, benefit amount, claim duration and main diagnosis. We use these data to estimate the effect of the reform on the probability of receiving DI and the (annual) amount received. Diagnosis recorded on entry to DI is used to distinguish claimants in the two diagnostic groups that include the most subjectively defined disabilities - musculoskeletal conditions and mental disorders. We lump all other disabilities together. The social security files are also used to identify benefits received from other social insurance and social assistance programs, which we aggregate to obtain annual net of tax income from social transfers other than DI.

Information on employment, days worked and annual net of tax earnings is taken from

¹² Only 91 out of 160,194 claimants aged 30-44 in July 2004 were not reassessed until the first five months of 2009.

¹³We do not use the more disaggregated data throughout because they are more noisy and the dataset becomes extremely large, which slows computation considerably on the remote server through which the administrative files are accessed.

tax records. We count a person as employed if she was an employee for at least one day in a calendar year. Annual earnings are divided by days worked to obtain a daily wage.

Municipal registers are used to identify date of birth and gender. Deaths are identified from the mortality register. The administrative files are linked using a unique individual identification number (RIN-code) that is issued on compulsory registration with the municipality at birth or after immigration. Additional details of the data sources and measures are provided in Appendix A Table A4.

4.2 Treatment and comparison groups

To construct the reform period sample, we select individuals who were claiming DI in January 2004. Of these, 3.9% die before the end of 2008 and are dropped from the panel. Mortality obviously differs between the age groups. But the age differential in mortality rates does not differ between the reform and non-reform periods. Hence, conditioning on survival does not introduce any compositional change that would bias the DADID estimates. We drop benefit recipients aged 45-49 on July 1, 2004 because the 2007 revision to the reform meant that either they were never reassessed under the stricter rules or they had their reassessment reversed. We also exclude recipients younger than 30 because there are very few of them and they typically have had little employment experience. Their employment patterns are likely to differ markedly from the older claimants we use as one comparison group. This leaves a treatment group of 160,194 individuals who were claiming DI in January 2004, were aged 30-44 on July 1, 2004 and so were eligible for reassessment and could be followed to the end of 2008 when the reassessments were completed.¹⁴

One of our comparison groups comprises 94,404 individuals who were claiming DI in

¹⁴The number of benefit recipients from this group who were called for reassessment is 137,419. Most (94%) of the others had a condition that was considered, without being subject to the full reassessment process, to render them completely incapable of work. These cases were reviewed and so are part of the treatment group. The remainder (6%) left DI before being called for reassessment. Since this exit may have been in anticipation of the outcome of reassessment, these individuals can be considered to have been exposed to the reform and are rightly part of the treatment group. Their inclusion will downwardly bias the DADID estimate only if they exited from DI already in 2004 (or October 2004 in the monthly analysis).

January 2004, were aged 50-53 on July 1, 2004 and so were not subject to reassessment. The non-reform period sample consists of individuals who were claiming DI in January 1999, were aged either 30-44 (as the treatment group, 139,524 individuals) or 50-53 (as reform period comparison group, 102,464 individuals) on July 1, 1999, and survived to the end of 2003. We pool this balanced panel with that constructed for the reform period.

Table 1 shows means of characteristics at selection into the samples, i.e. 1999 and 2004, by age group and period. In both age groups, there is a higher fraction of females in the later period. This partly reflects increasing labor force participation of Dutch women and is consistent with the feminization of DI rolls observed in other countries. More relevant to the plausibility of our identification strategy is that the age group difference in the proportion of female benefit recipients is roughly constant across the two periods. The same is true with respect to the average duration of a DI claim and the amount received. There is a discernible age group difference in the proportion of fully disabled claimants only in the earlier, non-reform period. Related to this, only in this period does the employment rate differ across the age groups, with the older benefit recipients being less likely to work (and more likely to be fully disabled). Consequently, the age difference in mean earnings is in the opposite direction in the two periods. These period differences in the gaps in the *levels* of employment and earnings between the age groups do not invalidate the DADID identification strategy. We examine whether there is any sign of the age-specific *trends* diverging up to the implementation of the reform in the next sub-section.

For both age groups, mean incomes from social transfer programs other than DI are higher at the start of the reform period than at the start of the non-reform period, and the age gap is somewhat wider in the reform period. The increase over time may well be due to the rise in the proportion of benefit recipients with mental health problems, who tend to be more heavily dependent on welfare. Combined with recipients with musculoskeletal conditions, they are the majority in all age groups and periods, and more so in the later period. In the earlier period, there is no age difference in the fraction of recipients with either

Table 1: Characteristics of DI recipients by period and age - Means at sample entry

| | Reform period | | Non-reform period | |
|-------------------------------|---------------|-----------|-------------------|-----------|
| | Age 30-44 | Age 50-53 | Age 30-44 | Age 50-53 |
| <i>Demographics</i> | | | | |
| Female | 60.3% | 45.7% | 53.4% | 37.4% |
| Age | 38.7 | 52.1 | 38.8 | 52.1 |
| <i>Disability insurance</i> | | | | |
| Claim duration (years) | 5.44 | 9.52 | 5.90 | 9.96 |
| Benefit amount (€/year) | 8422 | 9950 | 8559 | 10634 |
| Fully disabled | 63.5% | 64.0% | 65.4% | 69.4% |
| <i>Labor market</i> | | | | |
| Employed | 35.9% | 35.8% | 40.7% | 34.6% |
| Earnings (€/year) | 4207 | 5162 | 4947 | 4879 |
| <i>Other social transfers</i> | | | | |
| Benefit amount (€/year) | 1043 | 726 | 724 | 555 |
| <i>Diagnosis</i> | | | | |
| Mental disorders | 43.1% | 33.8% | 34.4% | 27.9% |
| Musculoskeletal | 28.9% | 32.9% | 25.0% | 31.2% |
| Other disabilities | 28.0% | 33.3% | 40.6% | 40.9% |
| <i>Number of Observations</i> | 160,194 | 94,404 | 139,524 | 102,464 |

Note: The Reform period panel refers to DI benefit recipients selected in January 2004. The Non-reform period panel refers to those selected in January 1999. Columns within each panel are split by age on July 1, 2004 (Reform period) and July 1, 1999 (Non-reform period). The first column in the Reform period panel corresponds to the treatment group. All others are for comparison groups. Earnings and benefit amounts are annual, net of taxes and inflated to 2015 price levels (Eurostat Netherlands HCPI 2015).

of these two more subjectively defined conditions. But in the later reform period, recipients in the younger group are more likely to have these diagnoses. This gives further reason to perform disaggregated analysis by diagnosis.

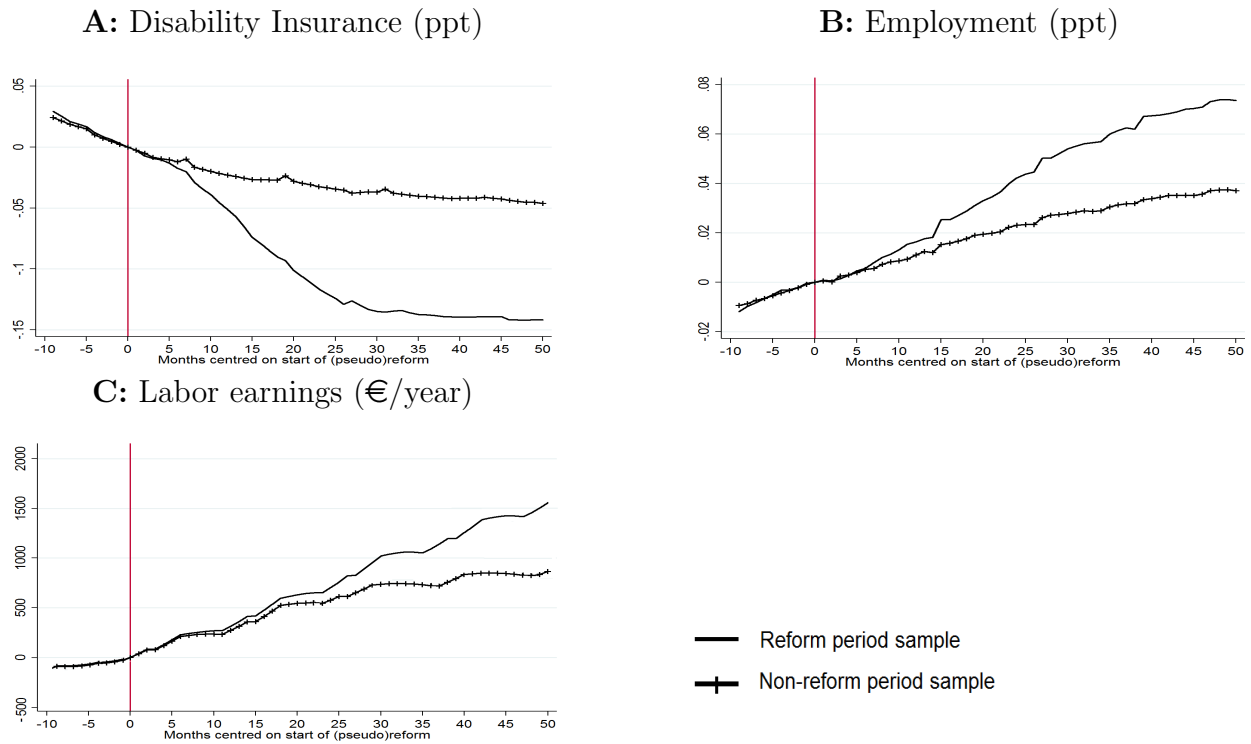
4.3 Trends

Figure 1 shows difference-in-differences in receipt of DI benefits, employment and labor earnings between the two age groups within each period. These figures are drawn using monthly data to allow more detailed assessment of the evolution of the trends before and after the start of the reassessments. Each line traces the age group difference (30-44 years - 50-53 years) in the deviation of the respective outcome from its value in month 0, which is October 2004 in the reform period, when reassessments started, and October 1999 in the non-reform period. After month 0, the difference in the DID between the periods corresponds to the DADID and gives an initial impression of the impact of the reform. By making the same comparison before month 0, the plausibility of the identification assumption can be assessed.

Consistent with the identification assumption, prior to month 0 the age group difference in the trend of each outcome is very similar across the two periods. In fact, up to month 5, i.e. five months after reassessments started in the reform period and by when only 8% of claimants aged 30-44 had been reassessed, there is little sign of the age differential in the trends differing across the periods. After that point, when the pace of reassessments picked up in the reform period (See Appendix A Table A3), the age differentials begin to diverge more markedly across the periods. This is consistent with the application of more stringent eligibility criteria to ever greater numbers of younger benefit recipients in the reform period having raised the rate at which they exited DI relative to older recipients, and with relative increases in the employment and earnings of younger recipients who either left DI or remained on the program despite experiencing a cut in their benefits.

Attribution of the differential trends across periods evident in Figure 1 to the reform

Figure 1: Age group difference-in-differences in outcomes by period



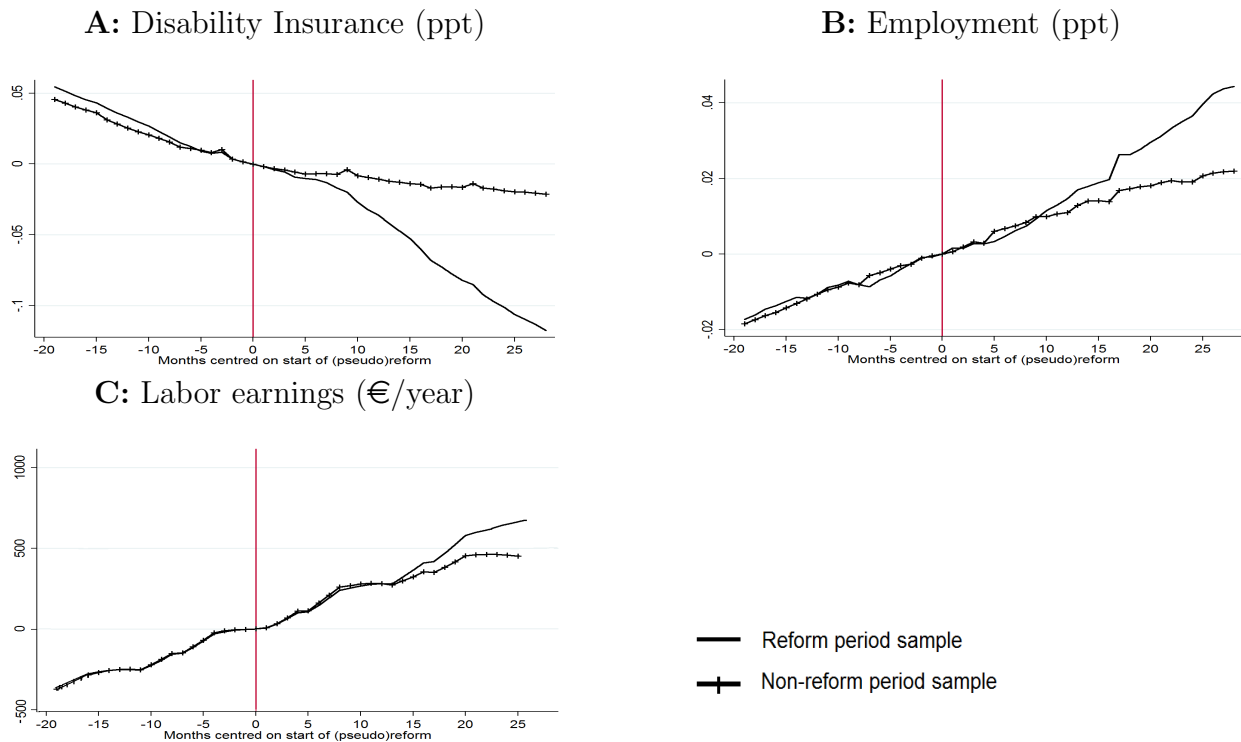
Note: Reform period (Jan. 2004-Dec. 2008) sample consists of individuals aged 30-44 & 50-53 on July 1, 2004 who were claiming DI in January 2004. Non-reform period (Jan. 1999-Dec. 2003) sample consists of individuals aged 30-44 & 50-53 on July 1, 1999 who were claiming DI in January 1999. Month 0 is October 2004 for reform period and October 1999 for non-reform period. Each line traces a period-specific difference-in-differences: the mean outcome at month t minus the mean outcome at month 0 for the 30-44 age group less the respective difference for the 50-53 age group. Group sizes are given in Table 1. ppt = percentage points.

rests on assumption (1) - the age differential in the outcome trend would have been common across periods in the absence of the reform. It is difficult to gauge the plausibility of this assumption from comparison of the outcome trends over two periods of only nine months (Jan.-Sept. 1999 and Jan.-Sept. 2004). To better assess whether the assumption is credible, we examine two different cohorts of DI claimants over a longer duration prior to the start of reassessments in the reform period. One of these cohorts consists of individuals who were: a) claiming DI in January 2003, b) aged 30-44 or 50-53 on July 1, 2004, and c) observable until December 2006. Those in the younger group of this cohort were subject to reassessment from October 2004, provided they were still on DI at that time. They are observed for 21 months prior to this date. The second cohort is defined exactly as the non-reform period groups we use for estimation except that the age criteria are applied on July 1, 2000 (rather than July 1, 1999) and we follow them only until December 2002. The pseudo reform period for this cohort is set as starting in October 2000.

Figure 2 shows the age group differential in the trends in DI participation, employment and labor earnings over the four years that these cohorts are followed. Over the 21 months prior to the start of the reassessments of the reform period sample, the age differentials in the outcome trends do not diverge markedly between the two cohorts. This is slightly less true for DI participation than it is for the other two outcomes. Apparently even before the start of reassessments in the reform period sample, younger claimants in this cohort were exiting DI at a faster rate relative to older claimants than was the case in the earlier period sample. This would be consistent with recipients in the later period leaving the program in anticipation of negative reassessments. This seems unlikely given there is no sign of a similar pre-reform divergence in the employment trends. Someone who anticipated that their DI benefits would be terminated or cut would have no incentive to leave the program before this occurred, unless they had found employment. There is a clear downward kink in the differential trend in DI participation in the reform period sample coincident with the acceleration in the reassessments from around month 5 and no such kink in the non-reform

period sample. The size of this divergence relative to the prior differential trend suggests that while the DADID may overestimate the impact of the reform on the DI exit rate, the upward bias is likely to be small. Further, the similarity of the trends in employment and earnings prior to month 0 across periods supports the validity of the DADID identification assumption for these outcomes.

Figure 2: Age group difference-in-differences in outcomes by period
 - extended duration prior to (pseudo) reform



Note: Reform period (Jan. 2003-Dec. 2006) sample consists of individuals aged 30-44 & 50-53 on July 1, 2004 who were claiming DI in January 2003. Non-reform period (Jan. 1999-Dec. 2002) sample consists of individuals aged 30-44 & 50-53 on July 1, 1999 who were claiming DI in January 2000. Month 0 is October 2004 for reform period and October 2000 for non-reform period.

5 Results

5.1 Main estimates

Table 2 reports estimates of β_4 from regressions (3). Each entry is a DADID estimate of the ATET - the effect of the reform on the respective outcome in 2008 averaged over all individuals who were aged 30-44 and claiming DI in 2004. By 2008, these individuals had been subject to reassessment under the more stringent criteria. We estimate that this reduced the probability of remaining on DI in 2008 by 14.4 percentage points. This includes both the direct effect of claims terminated through application of the stricter rules and any indirect effect that may arise through reduced benefits inducing some to leave DI. Even without the cuts, some claimants would have left the program by 2008. Using the regression estimates, we predict that 84.5% of individuals aged 30-44 who had been claiming DI in 2004 would still have been on the DI roll in 2008 if there had been no tightening of the rules.¹⁵ This implies that reassessment with stricter criteria reduced the DI participation rate by 17% of what it otherwise would have been. It raised the DI exit rate by 93%. On average, reassessment is estimated to have reduced the annual amount of DI benefit received by €1565, or around one fifth of the average amount under the counterfactual. If there had been no tightening of the eligibility criteria, we estimate that the target group of DI recipients would have been paid benefits equivalent to 46% of their pre-disability earnings in 2008. The reform is estimated to have reduced this replacement rate by 7.2 percentage points, on average.¹⁶

These estimates confirm that the 2004 reform substantially reduced DI benefits and dependency. It was about twice as aggressive as the reassessment of Dutch DI claimants a decade earlier that is estimated to have lowered the probability of remaining on DI by 3.8 percentage points, reduced benefits by 10% and decreased the replacement rate by 5.9

¹⁵ This is obtained by subtracting the estimated reform effect on DI participation from the predicted participation rate of the treatment group in 2008, i.e. $\frac{1}{n_T} \sum_i 1(AGE_i \times PERIOD_i \times YEAR_4) \hat{Y}_{it} - \hat{\beta}_4$, where $1()$ is the indicator function and n_T is the number of individuals in the treatment group.

¹⁶The replacement rate is averaged over the whole treatment group and is set to zero for those who had left DI by 2008.

percentage points (for claimants aged 45) (Borghans et al. 2014).

Table 2: Effects of reassessment of DI recipients under more stringent rules

| | Absolute Effect | Mean if no reassessment | Relative Effect |
|-------------------------------|----------------------|----------------------------|-----------------|
| | (1) | (2) | (1)/(2) |
| <i>Disability Insurance</i> | | | |
| Benefit receipt (ppt) | -14.40*** (0.17) | 84.52 | 17.04% |
| Benefit Amount (€/year) | -1,565*** (31.70) | 7,906 | 19.80% |
| Replacement Rate (ppt) | -7.20*** (0.12) | 46.19 | 15.59% |
| <i>Labor Market</i> | | | |
| Employment (ppt) | 6.68*** (0.22) | 33.83 | 19.76% |
| Days worked (year) | 17.03*** (0.58) | 76.26 | 22.33% |
| Earnings (€/year) | 995*** (43.19) | 5,507 | 18.07% |
| Wage (€/day) | 4.33 (2.50) | 65.09 | 6.65% |
| <i>Other social transfers</i> | | | |
| Benefit amount (€/year) | 376*** (17.73) | 877 | 42.90% |
| Number of individuals | 496,586 | | |
| Number of observations | 2,482,930 | | |

Notes: Column (1) gives least squares estimates of β_4 from (3). Standard errors, in parentheses, are adjusted for clustering at the individual level. Column (2) gives predicted mean outcome of 30-44 year olds in 2008 under counterfactual of no reform (see footnote 15). Right-hand column gives the estimate in column (1) as a percentage of the prediction in column (2). The number of individuals is the total across all treatment and comparison groups. For the numbers in each group, see Table 1. ppt = percentage points. *** indicates significance at the 1% level.

Having established that the reform substantially reduced DI benefits, we now turn to the question of central interest: what impact did this increased stringency have on the employment and earnings of claimants? We estimate that reassessment raised the probability

of employment by 6.7 percentage points, which is a 20% increase relative to the predicted employment rate of the treatment group in 2008 in the absence of the reform (Table 2). This is the effect on employment irrespective of whether the person continues to claim DI or not. It cannot therefore be compared with the estimated effect on DI receipt to reveal the fraction of those forced or induced to leave DI who entered employment. In order to calculate that fraction, we need to know the effect on the probability of working and not claiming DI. We estimate that reassessment raised this probability by a significant 8.5 points (SE=0.18, p-value<0.01). This is larger than the effect on the unconditional probability of employment, implying that reassessment reduced the likelihood of working and continuing to claim DI. This is likely due to cuts and terminations of the benefits paid to partially disabled claimants who had been working prior to reassessment. Setting the 8.5 points rise in the probability of working without claiming DI against the 14.4 points reduction in the probability of receiving DI benefits implies that 62% of those forced or induced to leave DI entered employment. This indicates substantial reserves of work capacity among DI recipients whose entitlement was reduced. The less positive interpretation is that almost two fifths of those who were forced or induced to leave DI did not find or look for work.

Borghans et al. (2014) estimate that the less stringent tightening of the Dutch DI program in 1994 increased employment by 2.9 points. In absolute terms, this is less than half the size of the effect we find on employment (unconditional on DI receipt). But it is larger relative to the 3.8 percentage points reduction in DI participation estimated by Borghans et al. The implied lower rate of absorption of displaced claimants from the later reform into employment is consistent with an expected decrease in the work capacity of claimants as the process of DI retrenchment proceeds. Moore (2015) estimates that 22% of US benefit recipients who lost DI entitlement as a result of addictive disorders being excluded from the qualifying conditions gained employment. This lower rate of re-entry to employment may be attributable to the addictive behavior of the targeted group, but it could also be due to the incentive for claimants exiting the US program to stay out of work in order to strengthen

their case should they reapply. There is no such incentive in the Dutch system.

We estimate that greater benefit stringency increased the number of days worked annually by 17; equivalent to 22% of the predicted mean for the treatment group in the absence of the reform. The extensive and intensive margin effects on labor supply produced an estimated €995 average increase in the annual earnings of DI claimants whose entitlement was reassessed. This is an 18% increase relative to predicted earnings under the counterfactual. It is almost two thirds of the estimated average reduction in the benefits received. For each €1 reduction in benefits, 64 cents could be regained through labor market earnings.¹⁷ This is very close to the estimate of Borghans et al. that earnings rose by 61 cents for each €1 reduction in benefits resulting from the 1994 reassessment of Dutch DI claimants. Apparently, even after that tightening of the eligibility criteria and a 2002 reform that is likely to have reduced the rate of entry to DI (Koning and Lindeboom 2015), DI recipients subjected to reassessment in 2004 still had considerable earnings capacity they could draw on to replace a substantial part of the benefits lost due to the increased program stringency. This is even more striking considering that those affected had been claiming DI for more than five years, on average, and 63% were classified as fully disabled (see Table 1). However, it needs to be borne in mind that these are average effects and reassessment did not change the benefit entitlement of 62% of recipients (Social Insurance Benefits Agency (UWV) 2009). Among those whose entitlement was terminated or cut, there are likely to be many who could not increase earnings to an extent anywhere near that sufficient to replace 64% of their lost benefit income.¹⁸

We find no significant effect on the daily wage rate, and the point estimate is positive.

¹⁷The estimated reduction in benefits is the combined effect of cuts and the response to those cuts through claimants leaving DI. Hence, the ratio of the estimated effects on earnings and benefit income cannot be interpreted as the rate at which earnings are crowded out by each €1 of DI benefit. But it suggests that the rate of crowd-out is at least as high as 0.64:1. The average cut in benefits will be less than the average reduction in benefits received.

¹⁸Also note that we are taking the ratio of two averages, not the average of a ratio. However, given that the expected effect on earnings relative to the expected effect on benefits is a lower bound of the expectation of the ratio of the individual earnings effect to the individual benefit effect (Cochran 2007), it is anticipated that we underestimate the average degree to which earnings replace lost benefit income.

This suggests that application of stricter entitlement rules did not cause recipients to leave DI and accept worse paying jobs than they would have settled for if their benefits had not been cut. Earnings rise because claimants are induced to work more and this is not offset by entering less productive jobs.

We estimate that increasing the stringency of DI increased the amount received from other social transfers by €376, on average (Table 2). This is 24% of the average reduction in the income received from DI. The analogous estimate from Borghans et al. is that 30% of the 1994 reduction in DI benefits was compensated through increased claims of other benefits. Opportunities to substitute between programs may have decreased in the decade between the reforms, but apparently not markedly. Around half of the spillover to other programs took the form of increased receipt of unemployment insurance benefits (Appendix B Table B1). This substitution occurs partly by default. Recipients deemed ineligible for DI were automatically transferred to unemployment insurance if they had made sufficient social insurance contributions prior to entering DI. The remainder of the average increase in income from other social programs is split between means-tested welfare payments (28%) and sickness pay (23%). The latter suggests that individuals forced or induced to leave DI continued to have health problems that disrupted their work, at least temporarily.¹⁹ The ability of claimants to switch from DI to other social programs constrains the impact of the retrenchment on household, but also public, finances.

Summing the average effects on earnings and other social transfer income gives a total of €1371, which is about 88% of the estimated average reduction in payments received from DI. Without taking this compensation into account, on average, the cuts to DI benefits lowered income from all sources by 11% relative to what it would have been if there had been no reform.²⁰ Allowing for the compensation through earnings and other transfer programs, the

¹⁹Our estimate of the impact on income from DI is gross of payments made to individuals who initially had their benefits terminated or cut but subsequently experienced health problems that allowed them to re-qualify for DI, or to become entitled to a higher DI benefit.

²⁰Using the estimates from columns (1) and (2) of Table 2, $-1565/(7906+5507+877)=-0.1095$.

relative decrease in mean income from all sources is only 1.3%.²¹ Increased labor market activity and dependence on other social programs greatly cushioned the relative impact of the disability benefit cuts. But it should be kept in mind that the average effect will hide substantially more severe income losses for some. We explore heterogeneity in section 5.3.

The estimates presented in Table 2 are generally robust to using monthly rather than annual data (Appendix B Table B2). This avoids the inaccuracies arising from the small fraction of recipients who were reassessed in the last quarter of 2004 and those reassessed during 2008. The magnitudes of the estimated effects tend to be somewhat smaller with the more disaggregated data, but the differences are not substantial.

5.2 Placebo test

The validity of our empirical strategy rests on the assumption that the age differential in the outcome trends that would have materialized between 2004 and 2008 in the absence of the DI reform is that which occurred between 1999 and 2003. To further assess the plausibility of this assumption, we perform a *placebo test* by estimating the DADID in outcomes of individuals who were not recipients of DI benefits, and so were not exposed to the reform, but who were potentially affected, possibly differentially by age, by differences in labor market conditions across the two periods. Placebo treatment and control groups are defined by age and period analogous to those used to estimate the effect of the reform. The difference is that we only use individuals who did not claim DI at any time between January 2004 and December 2008, or in the non-reform period between January 1999 and December 2003. The placebo treated individuals are the same age as the DI recipients who were reassessed (30-44 on July 1, 2004) and are observed between 2004 and 2008. The placebo control groups are (1) 50-53 on July 1, 2004, (2) 30-44 on July 1, 1999 and (3) 50-53 on July 1, 1999. We exclude individuals who were claiming unemployment insurance in 1999 (for non-reform period groups) or 2004 (for reform period groups) because the DI reform could potentially have affected their labor

²¹ $(-1565+995+376)/(7906+5507+877)=-0.0136$.

market opportunities by increasing the supply of labor from DI claimants. After imposing these exclusion restrictions, there are 6.7 million individuals available for the analysis. We use a random 50% sample of them.²²

The results presented in Table 3 show precisely estimated zero *effects* on three of the four labor market outcomes.²³ There is a very small, but statistically significant, negative *effect* on employment.²⁴ Given the size of the estimate, its significance may simply be attributable to the huge sample. The estimate suggests that employment of individuals aged 30-44 who were not recipients of DI fell by only 0.8% of what it would have been in 2008 if the age differential in the employment trends between 2004 and 2008 had been the same as that observed between 1999 and 2003. Under the same assumption, we estimate that the DI reform raised employment of DI recipients aged 30-44 by 20%. Hence, if anything, we may be slightly underestimating the impact on employment. But the placebo test suggests that any bias is marginal, and it gives no reason to doubt the validity of the identification with respect to the effect on earnings and the other two labor market outcomes.

5.3 Sub-sample estimates

Apparently, DI recipients were able to increase their labor market earnings to replace almost two thirds of the benefit income lost, on average. This suggests that the increased stringency imposed by the reform was warranted. Upward revision of the assessed earnings capacity of many claimants reduced their benefit entitlement and, consistent with the logic of the

²²A memory constraint on the remote server used for the analysis precludes use of all individuals.

²³The zero effect on the daily wage is not quite so precisely estimated. The moderately large standard error is due to a few outliers that greatly increase the variance and skewness of the distribution of this outcome. Excluding these outliers gives a point estimate of -1.58 with a standard error of 2.33.

²⁴The direction of the estimated effect on employment may seem puzzling given that macroeconomic conditions were better in 2004-2008 than they were in 1999-2003. But the effect is not simply a period effect. It is an age difference in the period effect on the trend. To make this explicit, we decompose the estimate by running (1) a DID regression across the two periods using individuals aged 30-44 in each, and (2) a DID regression across the two age groups using individuals observed in the 2004-2008 period. The latter produces an estimated negative effect on the employment of 30-44 year-olds: their employment improved by less than that of 50-53 year-olds between 2004 and 2008. DID (1) produces an estimated positive effect on employment: the employment of 30-44 year-olds increased by more in the 2004-2008 period than it did in the 1999-2003 period. This is consistent with the positive turn of the business cycle from 2004.

Table 3: Placebo test - application of empirical strategy with data on non-recipients of DI

| | Absolute effect (1) | Counterfactual mean (2) | Relative effect (1)/(2) |
|------------------------|---------------------------|-------------------------------|-------------------------------|
| Employment (ppt) | -0.57*** (0.01) | 73.43 | 0.78% |
| Days worked (year) | -0.15 (0.15) | 217.18 | 0.07% |
| Earnings (€/year) | -195.90 (125.78) | 34,061 | 0.58% |
| Wage (€/day) | 0.46 (3.78) | 151.81 | 0.30% |
| Number of individuals | 3,345,789 | | |
| Number of observations | 16,728,945 | | |

Notes: Column (1) gives least squares estimates of β_4 from (3) using individuals who did not claim DI at any time in the respective period (1999-2003 or 2004-2008) and were not claiming UI at the beginning of the period. Column (2) gives predicted mean outcome of 30-44 year olds in 2008 assuming the age differential in the outcome trend 2004-2008 equals that observed 1999-2003. (see footnote 15). Right-hand column gives the estimate in column (1) as a percentage of the prediction in column (2). Standard errors, in parentheses, are adjusted for clustering at the individual level. The number of individuals is the total across all placebo treatment and comparison groups. ppt = percentage points. *** indicates significance at the 1% level.

reform, they were able to respond by increasing their earnings. But the average response potentially obscures much variation in the impact of the reform and the earnings capacity it revealed. Such heterogeneity would be relevant for targeting the tightening of DI eligibility criteria elsewhere. We test for heterogeneous effects by splitting the sample by age, gender, cause and degree of disability, and estimating the regression model (3) separately for each sub-sample. We then assess whether the effects vary with the duration for which a recipient had been claiming DI prior to reassessment.

5.3.1 Effect by age and gender

The considerable earnings capacity detected may partly be attributable to the age of the benefit recipients affected by the reform, who, at 30-44, are younger than those targeted by most other DI reforms that have been evaluated.²⁵ The top panel of Table 4 reveals that the work and earnings responses to reassessment are even stronger among claimants aged 30 to 39.²⁶ Their probability of employment increased by twice as much the respective increase among those aged 40-44.²⁷

The employment impact relative to the predicted employment rate under the counterfactual of no reform (given in square brackets) is also twice as large for the younger of the two treatment groups.²⁸ The employment gain is 51% of the fall in DI participation in the younger group. In the 40-44 age group, employment rises by a lower 38% of the reduction in DI dependency. In absolute terms and relative to the counterfactual mean, earnings also rise, on average, by twice as much in the 30-39 group as in the 40-44 group. The younger

²⁵Although the 1994 Dutch reform impacted all claimants below the age of 45, the research design employed by Borghans et al. (2014) identifies the effect at the margin of that age threshold only. Karlström et al. (2008) find no employment effect from withdrawal of special eligibility rules for those aged 60-64 to qualify for the Swedish DI program. At the slightly younger age of 55-56, Staubli (2011) finds a positive impact of reduced DI entitlement on employment in Austria. As mentioned in section 1, Moore (2015) finds that younger (30-39 vs 50-61) US SSDI recipients who qualified through an addiction had a larger employment response to the termination of their benefits.

²⁶Splitting this group into those aged 30-34 and 35-39 reveals little further heterogeneity.

²⁷This difference, like all other heterogeneous effects referred to in the text, is significant.

²⁸The positive effect on days worked is also about twice as large for those aged 30-39 than it is for those aged 40-44. For all heterogeneous effects on days worked, the daily wage rate and other social transfers, see Appendix B Table B3.

group is able to recover 68% of the average reduction in DI income through increased labor market earnings, while the older group can make up only 55% of a smaller average loss. These results consistently indicate greater work and earnings capacity among the youngest DI recipients subjected to reassessment.²⁹

In absolute terms, the employment response of female benefit recipients is almost twice as large as that of male claimants (Table 4). Relative to the counterfactual, the impact on the female employment rate is more than twice that on the male rate. Absolutely and especially relative to the counterfactual mean, the positive impact on market earnings of female recipients is considerably larger than the respective effect on male earnings. Women are able to increase their earnings to replace a larger fraction of their lost DI benefits. The average earnings effect is three quarters of the average reduction in DI income for female recipients, compared with three fifths for male recipients. Female recipients were impacted more by the reform and, judging by their response to it, appear to have had greater earnings capacity.³⁰

²⁹Differences in the sex and disability composition of the age groups (see Appendix A Table A5), do not account for the difference in the earnings response by age. The younger group has a stronger earnings response irrespective of sex (Appendix B Table B4) and cause of disability (Appendix B, Table B5).

³⁰The age difference between male and female DI recipients (Appendix A Table A6) is not responsible for the different response by gender: within each age group, the employment and earnings effects are larger for females than for males (Appendix B Table B4).

Table 4: Effects of reassessment of DI recipients under more stringent rules by age and gender

| | Disability Insurance | | Labor Market | | No. individuals |
|--|---------------------------------|----------------------------------|-------------------------------|---------------------------------|-----------------|
| | Benefit Receipt (ppt) | Benefit Amount (€/year) | Employment (ppt) | Earnings (€/year) | |
| Age | | | | | |
| 30-39 years | -16.68*** (0.23) [20.19%] | -1,823*** (36.09) [24.48%] | 8.55*** (0.27) [25.17%] | 1,248*** (51.01) [23.28%] | 330,042 |
| 40-44 years | -11.35*** (0.22) [12.73%] | -1,225*** (39.47) [14.06%] | 4.30*** (0.27) [12.27%] | 667*** (53.30) [11.29%] | 363,412 |
| $H_0 : \beta_4^{30-39} = \beta_4^{40-44}$, p-value | <0.01 | <0.01 | <0.01 | <0.01 | |
| Gender | | | | | |
| Males | -8.31*** (0.25) [9.76%] | -1,375*** (49.79) [15.55%] | 4.21*** (0.32) [10.80%] | 815*** (72.91) [11.05%] | 244,076 |
| Females | -18.32*** (0.24) [21.38%] | -1,769*** (38.39) [23.43%] | 7.87*** (0.32) [24.68%] | 1,338*** (46.90) [31.71%] | 252,510 |
| $H_0 : \beta_4^{Males} = \beta_4^{Females}$, p-value | <0.01 | <0.01 | <0.01 | <0.01 | |

Notes: Group-specific least squares estimates of β_4 from (3) for the respective outcome. Standard errors adjusted for clustering at the individual level in parentheses. In square brackets is the estimated effect as a percentage of the predicted mean outcome under the counterfactual calculated as in footnote 15. p-values given for tests of equal effects across groups. Number of individuals is across all treatment and comparison groups. Number of observations is the number of individuals multiplied by 5. ppt = percentage points. *** indicates significance at the 1% level.

Table 5: Effects of reassessment of DI recipients under more stringent rules
by cause and degree of disability

| | Disability Insurance | | Labor Market | | No. indivs. |
|---|---------------------------------|----------------------------------|-------------------------------|---------------------------------|-------------|
| | Benefit Receipt (ppt) | Benefit Amount (€/year) | Employment (ppt) | Earnings (€/year) | |
| <i>Cause of disability</i> | | | | | |
| Musculoskeletal | -19.81*** (0.32) [23.03%] | -2,015*** (58.06) [27.83%] | 7.82*** (0.42) [18.84%] | 1,221*** (83.47) [16.93%] | 144,172 |
| Mental | -16.14*** (0.27) [18.34%] | -1,549*** (51.82) [18.49%] | 6.45*** (0.37) [22.22%] | 1,156*** (66.17) [27.45%] | 177,596 |
| Other | -7.80*** (0.31) [9.40%] | -1,111*** (56.36) [13.52%] | 5.48*** (0.37) [15.33%] | 620*** (76.46) [10.38%] | 174,816 |
| $H_0 : \beta_4^{Musculo} = \beta_4^{Other}$, p-value | <0.01 | <0.01 | <0.01 | <0.01 | |
| $H_0 : \beta_4^{Mental} = \beta_4^{Other}$, p-value | <0.01 | <0.01 | 0.0720 | <0.01 | |
| <i>Degree of disability</i> | | | | | |
| Fully disabled | -10.95*** (0.18) [12.22%] | -1,656*** (37.49) [17.05%] | 8.08*** (0.26) [49.93%] | 1,037*** (38.74) [51.03%] | 324,485 |
| Partially disabled | -20.73*** (0.35) [26.24%] | -1,243*** (57.40) [25.30%] | 4.00*** (0.41) [6.06%] | 838*** (99.65) [7.09%] | 172,101 |
| $H_0 : \beta_4^{Full} = \beta_4^{Partial}$, p-value | <0.01 | <0.01 | <0.01 | 0.0670 | |
| <i>Partially disabled</i> | | | | | |
| Not employed | -29.25*** (0.76) [32.48%] | -2,032*** (156.4) [22.72%] | 10.9*** (1.04) [53.93%] | 1,315*** (274.3) [34.50%] | 44,087 |
| Employed | -19.04*** (0.55) [23.45%] | -1,383*** (86.86) [18.83%] | -0.66 (0.41) [0.80%] | 548* (227) [2.94%] | 98,655 |
| $H_0 : \beta_4^{Not} = \beta_4^{Employed}$, p-value | <0.01 | <0.01 | <0.01 | <0.01 | |

Notes as Table 4

5.3.2 Effects by cause and degree of disability

To assess the hypothesis that benefit recipients qualifying through more subjectively defined and difficult to verify conditions have the greatest earnings capacity, we split the sample into three groups according to main diagnosis at entry to DI: musculoskeletal disorders, mental health disorders and a residual category of all other disorders.³¹ Consistent with the hypothesis, the employment response to reassessment is largest for those with musculoskeletal disorders, followed by those with mental health disorders (Table 5). This is entirely due to recipients with these conditions being much more likely to exit DI after reassessment. Their absorption into work appears to have been substantially lower. The employment effect is around 40% of the impact on the DI exit rate for claimants with musculoskeletal and mental health conditions. For those in the residual category of presumed more objective, verifiable disorders, the increase in employment is 70% of the fall in DI participation. This discrepancy could be taken as indicative of overly strict reassessment of recipients with musculoskeletal and mental health problems that produced upwardly biased evaluation of their earnings capacity. Their benefits were terminated and yet they were not able to find work. An alternative interpretation is that the low absorption rate reflects greater distaste for work among those who had qualified for DI through these two types of conditions. This is what one would expect if the work-shy were more likely to take advantage of the difficulty of verifying the validity of a claim of incapacity to work because of a mental health or musculoskeletal problem. Unfortunately, we cannot assess the relative veracity of these two explanations, which obviously have very different implications for the welfare impact of both the reform and the DI program.

In absolute terms, the average earnings response is about twice as large for those with musculoskeletal and mental health problems than it is for those with any other type of dis-

³¹The classification uses the most aggregated level of the International Classification of Diseases version 9. The residual category includes diseases and conditions related to the nervous system, many of which, such as physical pain, are difficult to verify. Disaggregated analysis revealed that the response to the reform in this category did not resemble the responses in the mental health and musculoskeletal categories more than it resembled the response of recipients with the remaining diseases included in the residual category.

order. Relative to predicted mean earnings under the counterfactual, the earnings effect is largest for claimants with mental disorders (28%) and lowest for those in the residual category (10%). This ranking is maintained when the earnings increase is expressed relative to the average drop in DI income (mental health - 75%; all other disorders - 56%). Since a larger proportion of those in the residual category who left DI entered employment, this suggests that claimants in this group were taking lower paid jobs, or they were not increasing labor supply at the intensive margin as much to compensate for the loss of partial disability benefits. Overall, the results indicate that benefit recipients with mental health and musculoskeletal problems did indeed have greater capacity to increase earnings in response to cuts in DI benefits, which were more aggressive for these claimants.³²

The estimates in the middle panel of Table 5 indicate that reassessment reduced the DI participation rate of those initially partially disabled by twice as much as it raised participation of the fully disabled. Despite this, the employment effect on the former group is only half as large as that on the latter group. The differential impact on program participation is because a downward revision in the degree of disability could bring a partially disabled benefit recipient below the 15% threshold necessary to qualify for any DI, while the same reduction might push a fully disabled recipient into a partial disability category. The differential impact on employment is because 68 percent of partially disabled recipients were working initially, compared with 18 percent of the fully disabled (Appendix A Table A8). Obviously, there is greater scope for those in the latter group to enter work. This is confirmed in the bottom panel, which shows that reassessment raised the employment of the partially disabled who were not working initially by almost much as it raised the employment of the fully disabled. Absolutely, and particularly relative to the counterfactual mean, benefit recipients who were initially fully disabled increased their earnings by more than the partially disabled. This is because the benefits of the fully disabled were cut more

³²These differences do not simply reflect sex, age and degree of disability differences by cause of disability (Appendix A Table A7). The earnings responses are largest for those with mental and musculoskeletal conditions within all sex, age and degree of disability groups (Appendix B Tables B6 and B5).

aggressively. Relative to the average loss of benefit income, the average increase in earnings is similar (63-68%). It is notable that recipients who had been classified as fully disabled were able to increase earnings to an extent sufficient to replace more than three fifths of the benefit income they lost.³³

5.3.3 Effects by duration of claim

We motivated this paper by arguing that available evidence indicating that rejected DI applicants have substantial earnings capacity is a poor guide to the effectiveness of measures that raise the exit rate from DI if earnings capacity deteriorates while receiving DI. While we have established that Dutch DI recipients did have considerable earnings capacity that could be tapped to replace a large fraction of the benefit income lost, we have not yet examined whether earnings potential does diminish with the length of a DI spell. We do so in this subsection.

Following Moore (2015), we extend equation 3 to allow the effect of the reform on earnings to vary with the duration of a claim. Let $DURATION_i$ be the number of years an individual had been receiving DI benefits at entry to the sample (January 1999 for the non-reform period and January 2004 for the reform period). We interact a third order polynomial of this duration with the product of the age group, period and year indicators that identifies exposure to the reform and so allow the effect of reassessment to vary nonlinearly with the duration of a claim.

$$\begin{aligned}
 Y_{it} = & \sum_{t=1}^4 \sum_{j=0}^3 \beta_{tj} (AGE_i \times PERIOD_i \times YEAR_t \times DURATION_i^j) \\
 & + \sum_{t=1}^4 \sum_{j=0}^3 \theta_{it} (YEAR_t \times DURATION_i^j) + \sum_{t=1}^4 \sum_{j=0}^3 \gamma_{it} (AGE_i \times YEAR_t \times DURATION_i^j) \\
 & + \sum_{t=1}^4 \sum_{j=0}^3 \delta_{it} (PERIOD_i \times YEAR_t \times DURATION_i^j) + \mu_i + \varepsilon_{it}
 \end{aligned} \tag{4}$$

³³The remainder of the loss of DI income was replaced by increased payments from other social transfer programs (see Appendix B Table B3).

Under the DADID identification assumption (1), the average treatment effect on those exposed to the reform and its nonlinear variation with the duration of a claim are obtained from the β_{4j} parameters. β_{40} corresponds to the effect of reassessment on someone who had just entered DI in January 2004. Adding combinations of the parameters β_{41}, β_{42} and β_{43} gives effects at positive durations. Estimates of these parameters are separately and jointly statistically significant (Appendix B Table B7) and imply that the effects do indeed vary with the duration of a claim. This variation is traced in Figure 3. We show estimates up to a duration of 15 years, which corresponds to the 80th percentile in the treatment group. The 95% confidence intervals (shaded) become wide and uninformative beyond that value. The absolute effect on all four outcomes declines as the time on DI lengthens. The diminishing effects on employment and earnings are consistent with the hypothesis that work and earnings capacity deteriorate while on DI. It is striking, however, that there is a significant earnings response even from individuals who had been claiming DI for 15 years. It is only around €450 per year, but it needs to be remembered that this is an average over all recipients and many experienced no cut in their DI benefits.

The effects on all four outcomes are rather constant up to a claim duration of about two years. Thereafter, they decline rapidly with the length of a claim and begin to flatten out only after about 13 years. This pattern does not resemble the inverted U-shaped employment response to DI terminations that Moore (2015) finds with US data. In that study, the response increases by about 8 percentage points up to a claim duration of three years. Moore suggests this could be due to a rehabilitation effect: time on DI provides an opportunity to recover health and work capacity. We find little or no evidence consistent with this hypothesis. Besides the obvious institutional differences, the discrepancy may be partly due to the different nature of the reforms examined. Moore estimates the response to a reasonably homogeneous treatment: the termination of DI benefits. In our case, while all claimants were reassessed, the consequences for their benefits differed. This complicates assessment of how the employment and earnings responses differ with claim duration. Panels

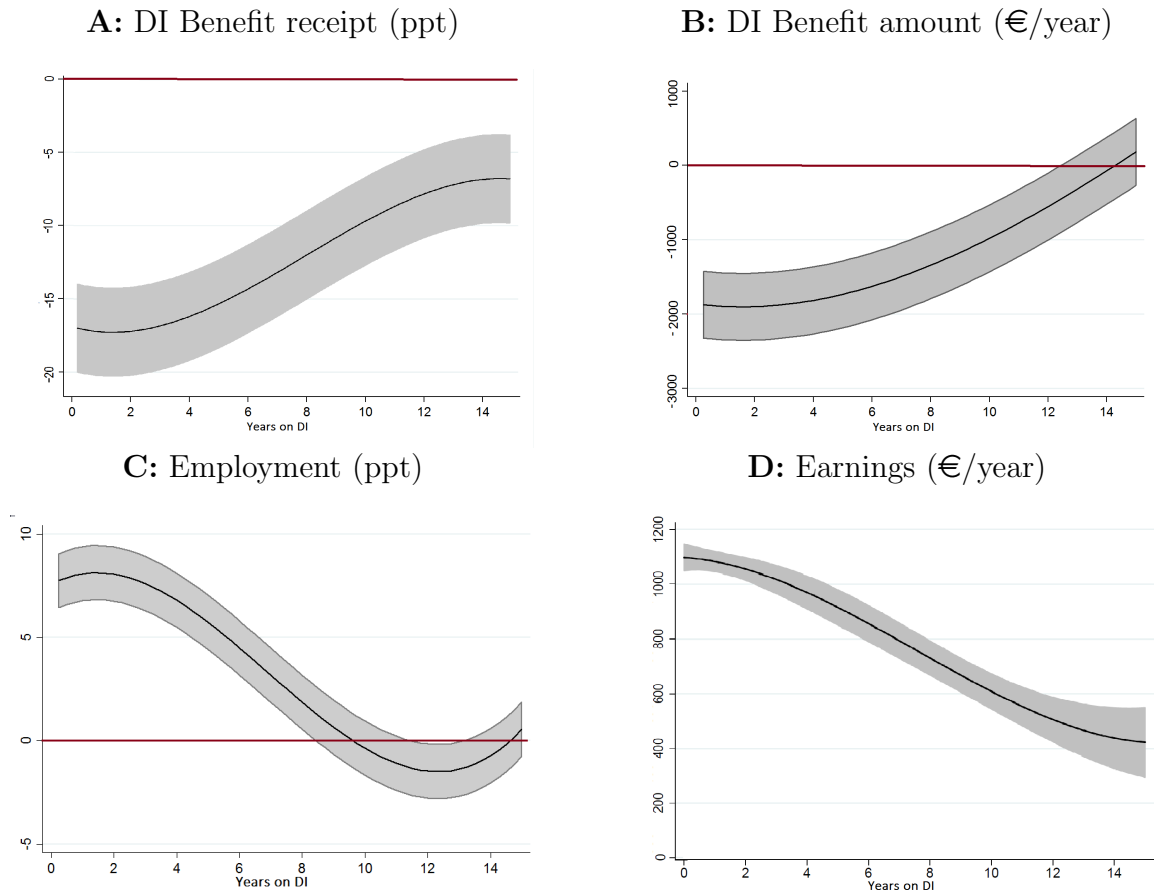
A and B of the figure show that reassessment impacted less on the DI benefits of claimants who had been on the program for longer. Up to a claim duration of about 8 years, the average effect on earnings is approximately half the size of the average effect on DI benefits. Thereafter, the fraction of the benefit cut that was replaced by earnings rises with claim duration. Reassessment provoked a smaller earnings response from DI recipients with longer claim durations because it withdrew less of their benefits. However, it was earnings capacity that was being reassessed. The muted effect on the entitlement of recipients with longer claim durations itself indicates that they had, or were perceived by the assessors to have, less potential to increase their work effort.

The relationship between the earnings effect of reassessment and claim duration differs markedly by age and sex (Figure 4 Panels A and B). The earnings response starts out much larger and declines more steeply with claim duration in the younger group. Earnings capacity is greatest among younger (30-39) beneficiaries who had been claiming DI for fewer years. In the older group (40-44), the earnings response varies less with claim duration, although there is some sign of the inverted U-shape found by Moore (2015), which is consistent with a period on DI providing time to regain earnings capacity before the offsetting effect of labor market detachment begins to dominate. The earnings effect declines steeply with length of time on DI among males but not females. It is remarkable that the average increase in earnings resulting from reassessment is in excess of €1000 per year among women who have been claiming DI for fifteen years.³⁴

The earnings response to reassessment declines more steeply with time on DI among fully disabled recipients than it does for the partially disabled (panel C). Splitting the partially disabled by initial employment status reveals that the earnings response-claim duration relationship mainly differs with the latter characteristic (panel D). The earnings response of the partially disabled who were not initially employed falls steeply with time claiming DI.

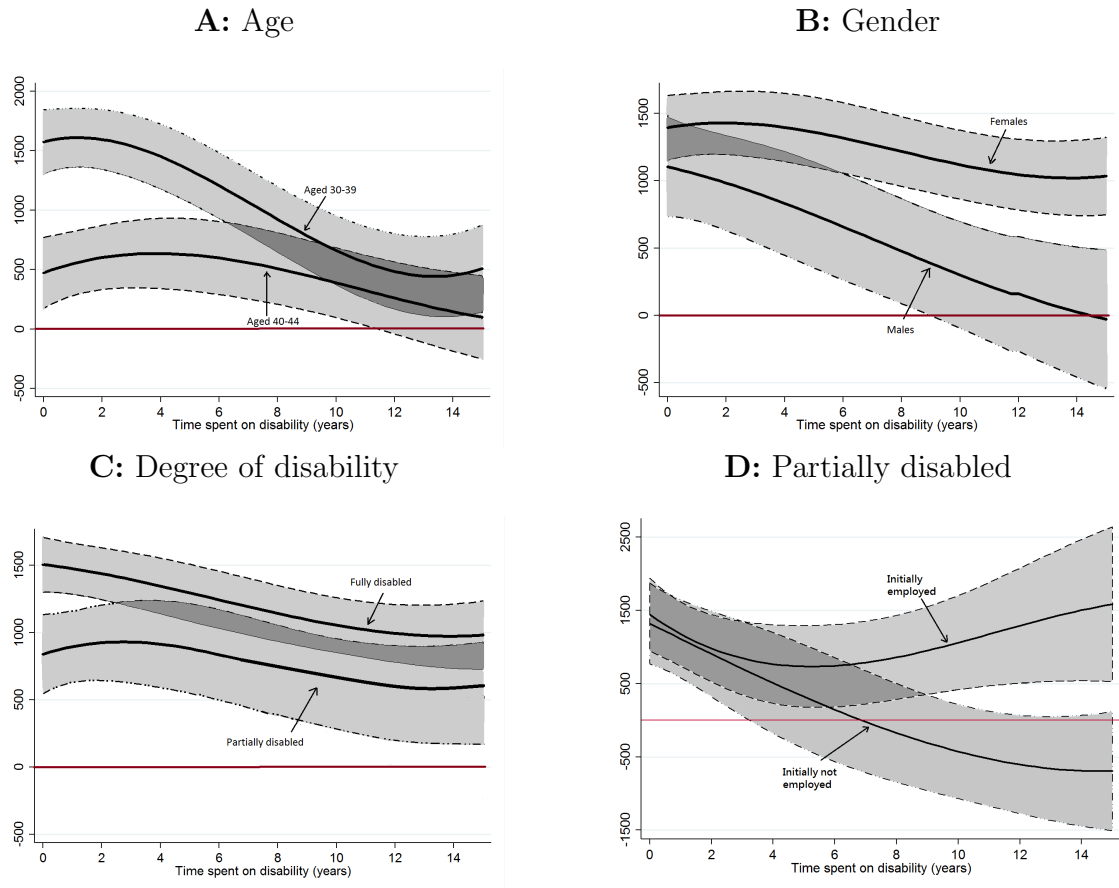
³⁴Partly due to the difference by sex, the earnings response declines more steeply with claim duration among recipients with musculoskeletal conditions, who are more likely to be male (Appendix B Figure B1). The earnings response curve is flatter for those with mental health problems, who are more likely to be female.

Figure 3: Effects of reassessment of DI recipients under more stringent rules by duration of claim



Notes: Estimates of average effects of reassessment by length of time on DI at entry to the sample. Derived from least squares estimates of $\beta_{4j}, j = 0, 1, 2, 3$ from (4) ((Appendix B Table B7). Shading indicates 95% confidence interval computed from delta method standard errors adjusted for clustering at the individual level. ppt = percentage points. Sample size is 496,586 individuals and 2,482,930 observations.

Figure 4: Earnings effect of reassessment by duration of claim - heterogeneity



Notes: Estimates of average effects of reassessment under stricter DI criteria by length of time on DI at entry to the sample. Derived from least squares estimates of $\beta_{4j}, j = 0, 1, 2, 3$ from (4) estimated separately for each sub-sample. Shading indicates 95% confidence interval computed from delta method standard errors adjusted for clustering at the individual level. Sample sizes is given in Tables 4 and 5.

After a duration of 3.5 years, there is no significant earnings response from this group. In contrast, the earnings response of the partially disabled who were initially working is always significant, and it is as strong for those who had been on DI for 15 years as it is for those who had just entered DI when the reassessments started. This is consistent with partial disability combined with employment being effective in maintaining earnings capacity as the time claiming DI lengthens.

6 Conclusion

To design policies that can relieve the fiscal strain caused by swelling disability insurance programs while protecting the wellbeing of those claimants who have little latitude to raise their labor market earnings, it is critical to know the earnings capacity of benefit recipients and how this varies with characteristics that can potentially be used to target interventions. This paper is one of only a few that deliver evidence on the employment and earnings responses of DI recipients to benefit cuts. Using a reform that affected Dutch DI recipients aged 30-44, we find that reassessment of entitlement under more stringent rules reduced benefits by around 20 percent, on average, and reduced dependency on the program by 14.4 percentage points. These cuts produced a 6.7 percentage points increase in employment and a 8.5 point increase in the probability of working and not claiming DI. Setting the latter against the fall in program dependency implies that three fifths of those forced or induced to leave DI entered employment. We estimate that the benefit cuts induced an 18 percent average increase in earnings. Each euro of DI income lost was replaced by 64 cents of labor earnings. We find no significant effect on the daily wage. Benefits recipients who were induced or forced to leave DI do not appear to have been pushed into less productive jobs.

These findings confirm that prior to the 2004 reform, some recipients of disability benefits in the Netherlands had substantial capacity to substitute earnings for benefits. Since the earnings response could come from an income effect as well as being due to disqualified

claimants no longer facing work disincentives that are built into DI rules, we do not know the extent to which it is indicative of inefficiency in the form of transfers crowding out earnings by reducing the return to work. It could also be that claimants impacted by the reform choose to enter work, or to work longer hours, because they have less non-labor income. Recent evidence suggests that much, perhaps even all, of the earnings crowd-out from US SSDI is an income effect (Gelber et al. 2017). In the Dutch context, it seems likely that a substantial part of the earnings response is indicative of inefficient moral hazard. One reason is that the response is so large. It is unlikely that this could be generated by an income effect alone. A second reason is that even a decade into retrenchment of the Dutch DI program, the dependency rate was still substantially higher than it is in the US. This reflected laxer qualification criteria and higher replacement rates that made the system vulnerable to moral hazard. The 2004 reform was intended to correct this. A third reason is that we find a stronger earnings response from benefit recipients who had qualified through mental health and musculoskeletal conditions that are the most difficult to verify. Without doubt, many, probably most, claimants who qualify for DI through these conditions have little capacity to earn sufficiently to support themselves and their families. However, we find that, on average, these claimants have greater potential to increase their earnings in response to benefit cuts than do claimants with other conditions.

Our finding of large employment and earnings responses to disability benefit cuts does not mean that all recipients of DI benefits in the Netherlands, or even a majority of them, have abundant earnings capacity. While all benefit recipients younger than 45 had their entitlement reassessed under stricter criteria, benefits were left unchanged for around two thirds of them. The strong earnings response is from the third who lost benefits. Across-the-board cuts in DI benefits would reduce the gains from insurance and cause hardship among many benefit recipients who would be unable to raise their earnings to make up the loss of income. The difficulty lies in discriminating between claimants who do and do not have the capacity to replace benefits with earnings. Our analysis reveals that those who do

are disproportionately younger and female, and that earnings potential is also higher among those who have been claiming DI for a shorter period, are diagnosed with a more subjectively defined health condition and are claiming partial disability benefits while working.

References

- Autor, D. H. (2015), The unsustainable rise of the disability rolls in the united states: Causes, consequences, and policy options, *in* J. K. Scholz, H. Moon and S.-H. Lee, eds, ‘Social Policies in an Age of Austerity: A Comparative Analysis of the US and Korea’, Edward Elgar, Northampton, MA, pp. 107–136.
- Autor, D. H. and Duggan, M. G. (2003), ‘The rise in the disability rolls and the decline in unemployment’, *The Quarterly Journal of Economics* **118**(1), 157–206.
- Autor, D. H. and Duggan, M. G. (2006), ‘The growth in the Social Security disability rolls: a fiscal crisis unfolding’, *Journal of Economic perspectives* **20**(3), 71–96.
- Autor, D. H., Maestas, N., Mullen, K. J. and Strand, A. (2015), Does delay cause decay? The effect of administrative decision time on the labor force participation and earnings of disability applicants, Working Paper 20840, National Bureau of Economic Research.
- Autor, D., Kostol, A. R., Mogstad, M. and Setzler, B. (2017), Disability benefits, consumption insurance, and household labor supply, Working Paper 23466, National Bureau of Economic Research.
- Bell, B., Blundell, R. and Van Reenen, J. (1999), ‘Getting the unemployed back to work: the role of targeted wage subsidies’, *International Tax and Public Finance* **6**(3), 339–360.
- Blundell, R. and Costa Dias, M. (2002), ‘Alternative approaches to evaluation in empirical microeconomics’, *Portuguese Economic Journal* **1**(2), 91–115.
- Borghans, L., Gielen, A. C. and 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.
- Bound, J. (1989), ‘The health and earnings of rejected Disability Insurance applicants’, *American Economic Review* **79**(3), 482–503.
- Bryngelson, A. (2009), ‘Long-term sickness absence and social exclusion’, *Scandinavian Journal of Public Health* **37**(8), 839–845.
- Burkhauser, R. V., Daly, M. C., McVicar, D. and Wilkins, R. (2014), ‘Disability benefit growth and disability reform in the us: lessons from other oecd nations’, *IZA Journal of Labor Policy* **3**(1), 4.
- Chen, S. and Van der Klaauw, W. (2008), ‘The work disincentive effects of the Disability Insurance program in the 1990s’, *Journal of Econometrics* **142**(2), 757–784.
- Cochran, W. G. (2007), *Sampling techniques*, John Wiley & Sons, New York.
- De Jong, P., Lindeboom, M. and Van der Klaauw, B. (2011), ‘Screening Disability Insurance applications’, *Journal of the European Economic Association* **9**(1), 106–129.

- French, E. and Song, J. (2014), ‘The effect of Disability Insurance receipt on labor supply’, *American Economic Journal: Economic Policy* **6**(2), 291–337.
- Gelber, A., Moore, T. J. and Strand, A. (2017), ‘The effect of Disability Insurance payments on beneficiaries’ earnings’, *American Economic Journal: Economic Policy* **9**(3), 229–61.
- Gruber, J. and Kubik, J. D. (1997), ‘Disability Insurance rejection rates and the labor supply of older workers’, *Journal of Public Economics* **64**(1), 1–23.
- Karlström, A., Palme, M. and Svensson, I. (2008), ‘The employment effect of stricter rules for eligibility for DI: Evidence from a natural experiment in Sweden’, *Journal of Public Economics* **92**(10-11), 2071–2082.
- Koning, P. and Lindeboom, M. (2015), ‘The rise and fall of Disability Insurance enrollment in the Netherlands’, *Journal of Economic Perspectives* **29**(2), 151–72.
- Kostol, A. R. and Mogstad, M. (2014), ‘How financial incentives induce Disability Insurance recipients to return to work’, *American Economic Review* **104**(2), 624–55.
- Maestas, N., Mullen, K. J. and Strand, A. (2013), ‘Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt’, *American Economic Review* **103**(5), 1797–1829.
- Ministry of Justice (2007), *Staatsblad van het Koninkrijk der Nederlanden*, Vol. 192, Sdu Uitgevers, The Hague.
- Moore, T. J. (2015), ‘The employment effects of terminating disability benefits’, *Journal of Public Economics* **124**, 30–43.
- OECD (2012), *Sick on the job? Myths and realities about mental health and work*, OECD Publishing, Paris.
- OECD (2014), *Self-reported health and disability*, OECD Publishing, Paris.
- Social Insurance Benefits Agency (UWV) (2009), *Rapportage afronding eenmalige herbeoordelingsoperatie*, Sdu Uitgevers, The Hague.
- Staubli, S. (2011), ‘The impact of stricter criteria for Disability Insurance on labor force participation’, *Journal of Public Economics* **95**(9-10), 1223–1235.
- Svensson, T., Müssener, U. and Alexanderson, K. (2010), ‘Sickness absence, social relations, and self-esteem: a qualitative study of the importance of relationships with family, workmates, and friends among persons initially long-term sickness absent due to back diagnoses’, *Work* **37**(2), 187–197.
- Van Sonsbeek, J.-M. and Gradus, R. H. (2012), ‘Estimating the effects of recent disability reforms in the Netherlands’, *Oxford Economic Papers* **65**(4), 832–855.
- Vingård, E., Alexanderson, K. and Norlund, A. (2004), ‘Consequences of being on sick leave’,

Scandinavian Journal of Public Health **32**, 207–215.

Von Wachter, T., Song, J. and Manchester, J. (2011), ‘Trends in employment and earnings of allowed and rejected applicants to the Social Security Disability Insurance program’, *American Economic Review* **101**(7), 3308–29.

Appendices

A Additional Descriptives

Table A1: Disability Insurance replacement rate per degree of disability

| Degree of disability | Benefit paid as % of pre-disability earnings |
|----------------------|--|
| 0-15% | 0 |
| 15-25% | 14% |
| 25-35% | 21% |
| 35-45% | 28% |
| 45-55% | 35% |
| 55-65% | 42% |
| 65-80% | 50.75% |
| 80-100% | 70% |

Note: Replacement rates prior to 2006. Source Social Insurance Benefits Agency (UWV).

Table A2: Reclassification of degree of disability following reassessments, recipients aged 30-44 on July 1, 2004

| | | Degree of disability after reassessment | | | | | | | |
|---|----------------|---|--------|--------|--------|--------|--------|--------|---------|
| | | <15% | 15-25% | 25-35% | 35-45% | 45-55% | 55-65% | 65-80% | 80-100% |
| Degree of disability before reassessment | 15-25% | 47.93 | 35.47 | 6.27 | 1.82 | 0.92 | 0.75 | 0.44 | 6.40 |
| | 25-35% | 29.67 | 18.71 | 34.86 | 5.56 | 2.04 | 1.02 | 0.75 | 7.39 |
| | 35-45% | 23.77 | 9.81 | 16.37 | 31.81 | 5.77 | 1.57 | 1.41 | 9.49 |
| | 45-55% | 25.75 | 5.27 | 6.41 | 9.67 | 32.34 | 4.22 | 2.09 | 14.25 |
| | 55-65% | 23.74 | 6.91 | 5.57 | 5.58 | 10.61 | 25.89 | 4.59 | 17.10 |
| | 65-80% | 16.60 | 5.51 | 6.86 | 7.20 | 7.07 | 8.44 | 26.71 | 21.61 |
| | 80-100% | 17.05 | 2.49 | 2.09 | 1.76 | 1.92 | 1.37 | 1.45 | 71.87 |

Notes: Authors' calculations using Social Insurance Benefits Agency (UWV) data.

Table A3: Reassessments of DI recipients aged 30-44 on July 1, 2004 by year

| | Number | Cumulative % |
|----------------|---------|--------------|
| 2004 (Oct-Dec) | 1,608 | 1.17 |
| 2005 | 61,612 | 46.01 |
| 2006 | 48,073 | 80.99 |
| 2007 | 21,480 | 96.62 |
| 2008 | 4,555 | 99.93 |
| 2009 | 91 | 100 |
| TOTAL | 137,419 | |

Notes: Authors' calculations using Social Insurance Benefits Agency (UWV) data. The table shows the timing of the reassessments of the 137,419 recipients aged 30-44 on July 1, 2004 who were called to have their entitlement reviewed under the new rules. An additional 22,775 recipients in this cohort were not reassessed. Most of them (94%) had a condition that was considered, without being subject to the full reassessment process, to render them completely incapable of work. The remainder left DI before being called for reassessment).

Table A4: Data sources and variable definitions

| Variable | Source | Definition | Timing and Units |
|------------------------------------|-----------|--|--|
| <i>I. Disability Insurance</i> | | | |
| Benefit receipt | SSB Files | Recorded claiming DI any time in year | 0/1 dummy |
| Benefit amount | SSB Files | DI benefit income after tax | €/year; 2015 prices |
| Degree of disability | SSB Files | Category of disability (0-15%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-80%, 80-100%). Partially disabled if DD (< 80%) Fully disabled if DD (\geq 80%) | At entry to sample January 1999 or 2004 |
| Cause of disability | SSB Files | Initial main diagnosis in 3 categories: i) mental, ii) musculoskeletal, iii) other | At entry to sample January 1999 or 2004 |
| Duration | SSB Files | Years claiming DI | At entry to sample January 1999 or 2004 |
| <i>II. Labor market</i> | | | |
| Employment | Tax Files | If have income from paid employment in year | 0/1 dummy |
| Days worked | Tax Files | Annual days worked | Days |
| Earnings | Tax Files | Net annual earnings from paid employment | €/year 2015 prices |
| Wage | Tax Files | Earnings/Day worked | €/day 2015 prices |
| <i>III. Other social transfers</i> | | | |
| Benefit amount | SSB Files | Net income from Unemployment Insurance, means-tested programs, Sickness Pay | €/year 2015 prices |

Note: SSB = Social Statistical Database. When using monthly data, we construct a monthly indicator of employment from the start and end dates of employment spells recorded in the SSB.

Table A5: Characteristics of DI recipients exposed to reassessment
Means at sample entry by age

| | Aged 40-44 | Aged 30-39 |
|--|------------|------------|
| <i>Panel A: Demographics</i> | | |
| Females | 47.39% | 54.79% |
| Year of Birth | 1956.38 | 1961.41 |
| <i>Panel B: Disability insurance</i> | | |
| Claim duration (years) | 7.35 | 6.57 |
| Benefit amount (euros/year) | 8918 | 8811 |
| Fully disabled | 63.48% | 62.71% |
| <i>Panel C: Labor market</i> | | |
| Employment | 40.24% | 36.64% |
| Earnings (€/year) | 5232 | 4638 |
| <i>Panel D: Other social transfers</i> | | |
| Benefit amount (€/year) | 688 | 971 |
| <i>Panel D: Diagnosis</i> | | |
| Mental disorders | 33.21% | 40.94% |
| Muskuloskeletal disorders | 26.83% | 29.85% |
| Other disabilities | 39.96% | 29.20% |
| Number of Individuals | 62,688 | 70,486 |

Notes as Table A6

Table A6: Characteristics of DI recipients exposed to reassessment
Means at sample entry by gender

| | Males | Females |
|--|--------------|----------------|
| <i>Panel A: Demographics</i> | | |
| Year of Birth | 1964.71 | 1965.68 |
| <i>Panel B: Disability insurance</i> | | |
| Claim duration (years) | 5.32 | 5.52 |
| Benefit amount (€/year) | 8783 | 8184 |
| Fully disabled | 59.78% | 65.91% |
| <i>Panel C: Labor market</i> | | |
| Employment | 40.71% | 32.78% |
| Earnings (€/year) | 5520 | 3342 |
| <i>Panel D: Other social transfers</i> | | |
| Benefit amount (€/year) | 1344 | 845 |
| <i>Panel D: Diagnosis</i> | | |
| Mental disorders | 40.03% | 45.14% |
| Muskuloskeletal disorders | 33.87% | 25.55% |
| Other disabilities | 26.10% | 29.31% |
| Number of Individuals | 63,614 | 96,580 |

Note: DI recipients are selected in January 2004. Money amounts inflated to 2015 price levels (Eurostat Netherlands HCPI 2015).

Table A7: Characteristics of DI recipients exposed to reassessment
Means at sample entry by cause of disability

| | Mental Disorders | Musculoskeletal Disorders | Other Disabilities |
|--|-----------------------------|--------------------------------------|-------------------------------|
| <i>Panel A: Demographics</i> | | | |
| Females | 63.13% | 53.39% | 63.03% |
| Year of Birth | 1965.53 | 1965.14 | 1965.09 |
| <i>Panel B: Disability insurance</i> | | | |
| Claim duration (years) | 5.42 | 5.06 | 5.86 |
| Benefit amount (€/year) | 9116 | 7264 | 8546 |
| Fully disabled | 72.53% | 50.52% | 62.87% |
| <i>Panel C: Labor market</i> | | | |
| Employment | 29.34% | 43.22% | 38.56% |
| Earnings (€/year) | 3074 | 5344 | 4780 |
| <i>Panel D: Other social transfers</i> | | | |
| Benefit amount (€/year) | 955 | 1376 | 838 |
| Number of Individuals | 69,057 | 46,225 | 44,912 |

Notes as Table A6

Table A8: Characteristics of DI recipients exposed to reassessment
Means at sample entry by degree of disability

| | Fully Disabled | Partially Disabled |
|--|----------------|--------------------|
| <i>Panel A: Demographics</i> | | |
| Females | 62.60% | 56.27% |
| Year of Birth | 1965.36 | 1965.17 |
| <i>Panel B: Disability insurance</i> | | |
| Claim duration (years) | 5.95 | 4.55 |
| Benefit amount (€/year) | 10528 | 4761 |
| <i>Panel C: Labor market</i> | | |
| Employment | 17.55% | 67.87% |
| Earnings (€/year) | 1435 | 9025 |
| <i>Panel D: Other social transfers</i> | | |
| Benefit amount (€/year) | 539 | 1921 |
| <i>Panel D: Diagnosis</i> | | |
| Mental disorders | 49.26% | 32.42% |
| Muskuloskeletal disorders | 22.97% | 39.09% |
| Other disabilities | 27.77% | 28.50% |
| Number of Individuals | 101,679 | 58,515 |

Notes as Table A6

B Additional Results

Table B1: Effects of reassessment of DI recipients under more stringent rules on income from other social transfers

| | Absolute effect | Mean if no re-assessment | Relative effect |
|---------------------------------|----------------------|--------------------------|-----------------|
| | (1) | (2) | (1)/(2) |
| Unemployment Insurance (€/year) | 182.22*** (13.25) | 396.34 | 45.60% |
| Means-tested welfare (€/year) | 106.26*** (14.28) | 345.54 | 30.72% |
| Sick Pay (€/year) | 86.42*** (6.85) | 185.36 | 46.56% |
| <i>Number of individuals</i> | <i>496,586</i> | | |
| <i>Number of observations</i> | <i>2,482,930</i> | | |

Notes: Column (1) gives least squares estimates of β_4 from (3). Column (2) gives predicted mean outcome of 30-44 year olds in 2008 under the counterfactual of no reassessments. Right-hand column gives the estimate in column (1) as a percentage of the prediction in column (2). Standard errors, in parentheses, are adjusted for clustering at the individual level. Sick pay refers to publicly provided sickness benefits. *** indicates significance at the 1% level.

Table B2: Effects of reassessment of DI recipients under more stringent rules - Estimates using annual and monthly data

| | Annual data (1) | Monthly data (2) |
|------------------------------|----------------------|----------------------|
| <i>Disability Insurance</i> | | |
| Benefit receipt (ppt) | -14.40*** (0.17) | -13.57*** (0.37) |
| Benefit amount (€/year) | -1,565*** (31.70) | -1,521*** (65.17) |
| Replacement rate (ppt) | -7.20*** (0.12) | -7.78*** (0.25) |
| <i>Labor Market</i> | | |
| Employment (ppt) | 6.68*** (0.22) | 5.67*** (0.46) |
| Days worked (year) | 17.03*** (0.58) | 15.46*** (1.23) |
| Earnings (€/year) | 995*** (43.19) | 784*** (93.12) |
| Wage (€/day) | 4.33 (2.50) | 3.98 (1.50) |
| <i>Number of individuals</i> | 496,586 | |

Notes: Column (1) gives least squares estimates of β_4 from (3) using annual data. Column (2) gives equivalent estimates obtained using monthly data. In this case, differences are taken relative to September 2004 (in reform period) and estimated effects at December 2008 are presented. Standard errors, in parentheses, are adjusted for clustering at the individual level. To get number of observations, multiply number of individuals by 5 for annual data and by 60 for monthly data. ppt = percentage points. Significance levels: *** = 1%.

Table B3: Heterogeneous effects of reassessment of DI recipients under more stringent rules on additional outcomes

| | Labor Market | | Other Social Transfers | <i>No. individuals</i> |
|---|--------------------------------|-----------------------------|-----------------------------------|------------------------|
| | Days worked in year | Wage (€/day) | Benefit amount (€/year) | |
| <i>Panel A: Age</i> | | | | |
| Aged 30-39 | 21.25*** (0.70) [28.41%] | 4.51 (2.49) [7.05%] | 331.53*** (21.17) [33.13%] | 330,042 |
| Aged 40-44 | 11.60*** (0.71) [14.17%] | 4.03 (3.40) [6.04%] | 432.28*** (22.31) [55.96%] | 363,412 |
| <i>Panel B: Gender</i> | | | | |
| Males | 12.40*** (0.91) [12.45%] | 4.78 (2.95) [7.09%] | 65.57*** (27.78) [5.53%] | 244,076 |
| Females | 21.95*** (0.72) [35.50%] | 7.16 (6.71) [11.87%] | 575.99*** (22.03) [80.39%] | 252,510 |
| <i>Panel C: Cause of disability</i> | | | | |
| Mental disorders | 18.48*** (0.89) [30.55%] | 4.55 (5.64) [7.20%] | 443.89*** (29.81) [48.96%] | 177,596 |
| Musculoskeletal disorders | 20.88*** (1.15) [21.19%] | 2.68 (3.89) [3.94%] | 318.52*** (33.94) [30.51%] | 144,172 |
| Other disabilities | 12.04*** (0.99) [14.70%] | 5.58 (3.88) [8.65%] | 324.08*** (29.48) [43.43%] | 174,816 |
| <i>Panel D: Degree of disability</i> | | | | |
| Fully disabled | 18.32*** (0.57) [54.73%] | 0.41 (4.68) [0.71%] | 679.97*** (18.07) [94.15%] | 324,485 |
| Partially disabled | 13.21*** (1.25) [8.49%] | 6.36** (2.98) [9.17%] | -141.89*** (37.73) [11.83%] | 98,655 |

Least squares estimates of β_4 from (3). Standard errors, in parentheses, are adjusted for clustering at the individual level. In brackets is effect as percentage of the predicted mean outcome of 30-44 year olds in 2008 under the counterfactual of no reassessment. *** indicates significance at the 1% level.

Table B4: Effects of reassessment of DI recipients under more stringent rules by age and degree of disability within gender

| | Males | | | | Females | | | |
|-------------------------------|----------------------|-------------------|-----------------------|-------------------|----------------------|-------------------|-----------------------|-------------------|
| | Estimate | Mean if no reform | Estimate | Mean if no reform | Estimate | Mean if no reform | Estimate | Mean if no reform |
| | Aged 30-39 | | Aged 40-44 | | Aged 30-39 | | Aged 40-44 | |
| <i>Disability Insurance</i> | | | | | | | | |
| Benefit receipt (ppt) | -9.57 (0.35) | 81.49% | -6.97 (0.30) | 88.84% | -20.53 (0.30) | 83.19% | -15.00 (0.32) | 89.48% |
| Benefit Amount (€/year) | -1,536.89 (59.54) | 8,045.69 | -1,202.49 (61.48) | 9,627.01 | -1,905.87 (43.25) | 7,047.52 | -1,212.89 (48.02) | 7,927.82 |
| <i>Labor Market</i> | | | | | | | | |
| Employment (ppt) | 5.60 (0.42) | 38.36% | 2.81 (0.38) | 39.59% | 9.72 (0.37) | 31.98% | 5.07 (0.39) | 31.71% |
| Days worked (year) | 16.79 (1.18) | 96.49 | 7.94 (1.08) | 102.66 | 26.18 (0.85) | 60.51 | 15.56 (0.90) | 63.78 |
| Earnings (€/year) | 1,156.49 (92.58) | 7,156.52 | 465.28 (87.15) | 7,589.20 | 1,584.13 (55.03) | 4,098.25 | 968.59 (60.14) | 4,404.13 |
| Wage (€/day) | 4.64 (3.23) | 66.89 | 4.84 (4.05) | 67.98 | 8.36 (6.62) | 58.25 | 5.05 (7.74) | 63.89 |
| <i>Other social transfers</i> | | | | | | | | |
| Benefit Amount (€/year) | -41.69 (35.04) | 1,353.31 | 172.37 (34.28) | 1,019.65 | 527.88 (25.84) | 815.70 | 649.83 (28.05) | 565.40 |
| <i>Number of Individuals</i> | | 180,884 | | 182,041 | | 185,100 | | 150,745 |
| | Fully Disabled | | Partially disabled | | Fully Disabled | | Partially Disabled | |
| <i>Disability Insurance</i> | | | | | | | | |
| Benefit receipt (ppt) | -4.85 (0.26) | 90.23% | -14.53 (0.46) | 79.14% | -14.52 (0.25) | 89.25% | -25.89 (0.59) | 79.28% |
| Benefit Amount (€/year) | -1,327.75 (84.09) | 13,685.00 | -1,423.55 (104.72) | 66,89.84 | -1,656.27 (55.73) | 11,170.75 | -1,886.19 (118.24) | 5,434.19 |
| <i>Labor Market</i> | | | | | | | | |
| Employment (ppt) | 4.57 (0.40) | 20.15% | 3.50 (0.52) | 66.64% | 9.73 (0.35) | 14.29% | 3.13 (0.71) | 66.69% |
| Days worked (year) | 11.76 (0.94) | 46.04 | 13.82 (1.74) | 177.16 | 22.17 (0.71) | 26.10 | 14.83 (1.88) | 136.84 |
| Earnings (€/year) | 779.38 (98.63) | 3,724.74 | 1,026.74 (206.79) | 19,172.48 | 1,471.23 (57.98) | 1,962.48 | 1,286.11 (181.87) | 12,775.26 |
| Wage (€/day) | -0.45 (9.11) | 82.90 | 12.65 (5.87) | 94.92 | 3.59 (14.39) | 69.15 | 13.56 (14.34) | 82.51 |
| <i>Other public benefits</i> | | | | | | | | |
| Benefit Amount (€/year) | 403.12 (40.67) | 1,191.48 | -757.02 (71.39) | 2,240.90 | 937.88 (28.26) | 793.10 | 78.55 (78.17) | 1,047.52 |
| <i>Number of Individuals</i> | | 197,909 | | 145,494 | | 232,406 | | 104,020 |

Least squares estimates of β_4 from (3). Standard errors, in parentheses, are adjusted for clustering at the individual level.

Table B5: Effects of reassessment of DI recipients under more stringent rules by age and cause of disability

| | Aged 30-39 | | | | | | Aged 40-44 | | | | | |
|-------------------------------|--------------------|-------------------------|--------------------|-------------------------|--------------------|-------------------------|------------------|-------------------------|------------------|-------------------------|--------------------|-------------------------|
| | Estimate | Mean if no reform | Estimate | Mean if no reform | Estimate | Mean if no reform | Estimate | Mean if no reform | Estimate | Mean if no reform | Estimate | Mean if no reform |
| | Musculoskeletal | | Mental disorders | | Other disabilities | | Musculoskeletal | | Mental disorders | | Other disabilities | |
| <i>Disability Insurance</i> | | | | | | | | | | | | |
| Disability (ppt) | -20.02 (0.28) | 79.72% | -19.89 (0.31) | 87.24% | -3.86 (0.86) | 73.96% | -11.06 (0.71) | 85.08% | -12.14 (0.21) | 0.90 | -9.12 (0.12) | 81.27% |
| Benefit Amount (€/year) | -2,526.33 (378) | 7,653.34 | -1,790.62 (545) | 9,805.52 | -1,138.82 (372) | 9,044 | -1,214 (306) | 9,215 | -1,338 (599) | 11,300 | -1,046 (459) | 10,670 |
| <i>Labor Market</i> | | | | | | | | | | | | |
| Employment (ppt) | 8.86 (0.71) | 45.57% | 9.18 (0.36) | 28.07% | 6.31 (0.34) | 37.57% | 3.99 (0.23) | 44.55% | 4.89 (0.71) | 0.30 | 2.15 (0.20) | 43.39% |
| Days worked (year) | 28.49 (1.85) | 80.21 | 21.13 (1.86) | 50.80 | 16.91 (2.76) | 38.91 | 10.03 (2.44) | 19.90 | 9.93 (1.54) | 64.96 | 8.13 (3.74) | 91.80 |
| Earnings (€/year) | 1,653 (116) | 9,681 | 1,312 (110) | 5,747 | 708 (107) | 8,777 | 1,015 (171) | 9,832 | 698 (78.3) | 6,452 | 464 (230) | 7,613 |
| Wage (€/year) | 4.63 (10.77) | 88.41 | -4.25 (3.38) | 89.96 | 1.83 (6.69) | 89.34 | 3.69 (1.89) | 81.83 | 8.89 (7.23) | 75.23 | 4.82 (3.80) | 95.24 |
| <i>Other social transfers</i> | | | | | | | | | | | | |
| Benefit Amount (€/year) | 540 (108) | 1,159 | 173 (33.41) | 1,513 | 122 (181) | 1,207 | -156 (27.04) | 1,721 | 981 (835) | 544 | 1,818 (1,864) | 1,261 |
| <i>Number of Individuals</i> | | 109,246 | | 149,833 | | 106,867 | | 89,286 | | 110,518 | | 132,981 |

Least squares estimates of β_4 from (3). Standard errors, in parentheses, are adjusted for clustering at the individual level.

Table B6: Effects of reassessment of DI recipients under more stringent rules by cause of disability and gender

| | Musculoskeletal | | Mental | | Other disabilities | | Musculoskeletal | | Mental | | Other disabilities | |
|-------------------------------|-----------------------|-------------------|-----------------------|-------------------|----------------------|-------------------|---------------------------|-------------------|-----------------------|-------------------|-----------------------|-------------------|
| | Estimate | Mean if no reform | Estimate | Mean if no reform | Estimate | Mean if no reform | Estimate | Mean if no reform | Estimate | Mean if no reform | Estimate | Mean if no reform |
| | Males | | | | | | Females | | | | | |
| <i>Disability Insurance</i> | | | | | | | | | | | | |
| Disability (ppt) | -13.14 (0.44) | 85.10% | -9.29 (0.39) | 88.17% | -2.14 (0.45) | 82.51% | -25.94 (0.48) | 87.14% | -20.03 (0.37) | 87.80% | -11.06 (0.43) | 83.26% |
| Benefit Amount (€/year) | -252.67 (67.17) | 2,030.29 | 126.83 (67.32) | 1,379.70 | -34.93 (66.16) | 1,348.38 | 681.21 (58.32) | 835.23 | 678.29 (48.20) | 1005.24 | 498.30 (46.54) | 790.64 |
| | -1,883.30 (107.03) | 9,464.59 | -1,381.61 (113.53) | 12,001.82 | -548.08 (123.50) | 11,368.09 | -2,203.01 (92.38) | 8,043.72 | -1,972.94 (80.55) | 9,537.65 | -1,450.15 (92.72) | 9,660.77 |
| <i>Labor Market</i> | | | | | | | | | | | | |
| Employment (ppt) | 5.41 (0.55) | 48.71% | 3.53 (0.54) | 29.73% | 3.10 (0.55) | 40.98% | 8.91 (0.65) | 36.30% | 7.71 (0.52) | 29.05% | 6.85 (0.52) | 32.68% |
| Days worked (year) | 18.04 (1.71) | 127.93 | 12.51 (1.40) | 69.57 | 6.62 (1.61) | 106.57 | 25.60 (1.47) | 70.66 | 21.94 (1.14) | 55.28 | 18.19 (1.21) | 64.71 |
| Earnings (€/year) | 1,743.90 (186.36) | 12,238.23 | 1,121.46 (165.64) | 6671.89 | 454.43 (239.80) | 10,726.13 | 1,978.59 (121.81) | 5,887.98 | 1,510.24 (102.78) | 4,691.40 | 1,163.77 (113.53) | 5,654.61 |
| Wage (€/year) | 9.75 (6.74) | 93.28 | 13.52 (16.88) | 81.50 | 6.06 (5.50) | 91.09 | -4.56 (14.87) | 89.43 | 2.38 (21.29) | 83.40 | 19.67 (21.29) | 70.55 |
| <i>Other social transfers</i> | | | | | | | | | | | | |
| Benefit Amount (€/year) | -252.67 (67.17) | 2,030.29 | 126.83 (67.32) | 1,379.70 | -34.93 (66.16) | 1,348.38 | 681.21 (58.32) | 835.23 | 678.29 (48.20) | 1,005.24 | 498.30 (46.54) | 790.64 |
| <i>Number of Individuals</i> | | 116,546 | | 137,743 | | 89,810 | | 85,980 | | 151,903 | | 98,632 |
| | Fully Disabled | | | | | | Partially Disabled | | | | | |
| <i>Disability Insurance</i> | | | | | | | | | | | | |
| Disability (ppt) | -16.79 (0.38) | 90.64% | -12.70 (0.26) | 91.45% | -4.66 (0.32) | 87.27% | -24.04 (0.53) | 82.46% | -23.87 (0.72) | 77.57% | -13.51 (0.61) | 76.88% |
| Benefit Amount (€/year) | -2,350.08 (99.48) | 11,419.04 | -1,956.68 (73.32) | 12,187.92 | -1,150.74 (89.76) | 12,488.20 | -1,399.51 (113.23) | 5,588.70 | -1,357.09 (168.99) | 5,974.24 | -1,213.94 (139.68) | 6,841.23 |
| <i>Labor Market</i> | | | | | | | | | | | | |
| Employment (ppt) | 11.28 (0.57) | 16.62% | 7.05 (0.40) | 16.17% | 6.81 (0.45) | 16.27% | 3.93 (0.63) | 67.23% | 4.66 (0.90) | 62.93% | 2.96 (0.66) | 68.02% |
| Days worked (year) | 24.68 (1.25) | 35.55 | 17.77 (0.85) | 31.59 | 13.85 (1.00) | 35.16 | 12.85 (2.03) | 166.66 | 19.17 (2.52) | 137.51 | 7.11 (2.09) | 160.82 |
| Earnings (€/year) | 1,856.38 (109.92) | 2,762.82 | 1,309.48 (81.26) | 2,500.11 | 961.04 (99.91) | 2,870.39 | 1,199.67 (217.39) | 16,955.76 | 1,447.35 (293.38) | 14,321.14 | 507.11 (240.73) | 16,552.71 |
| Wage (€/day) | -12.68 (22.54) | 90.48 | 9.34 (20.44) | 66.19 | 0.79 (8.43) | 75.73 | 9.36 (6.49) | 91.29 | 10.40 (11.90) | 90.66 | 12.33 (8.15) | 89.44 |
| <i>Other public benefits</i> | | | | | | | | | | | | |
| Benefit Amount (€/year) | 841.46 (50.81) | 1,152.88 | 698.29 (38.42) | 975.26 | 586.24 (41.47) | 839.01 | -312.17 (78.89) | 1,589.46 | -21.52 (120.50) | 1,502.57 | -80.92 (82.25) | 1,142.16 |
| <i>Number of Individuals</i> | | 98,843 | | 211,973 | | 119,498 | | 97,535 | | 80,892 | | 71,111 |

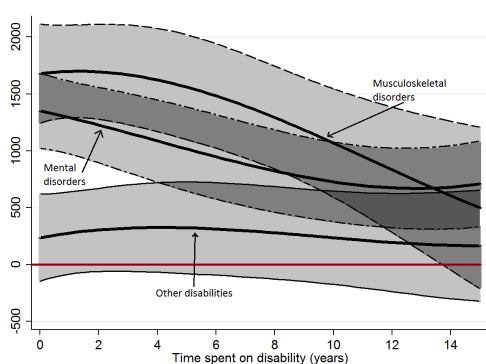
Least squares estimates of β_4 from (3). Standard errors, in parentheses, are adjusted for clustering at the individual level.

Table B7: Coefficients on interaction between reform exposure and third-order polynomial of DI claim duration

| | Disability Insurance | | Labor Market | |
|--|------------------------|-------------------------|-----------------------|-----------------------|
| | Benefit Receipt (ppt) | Benefit Amount (€/year) | Employment (ppt) | Earnings (€/year) |
| β_{40} | -16.87*** (1.31) | -1865.14*** (498.45) | 7.56*** (0.94) | 921.53*** (215.29) |
| β_{41} | -0.5838** (0.2326) | -51.47*** (18.54) | 0.0193*** (0.0073) | 31.98* (17.20) |
| β_{42} | 0.2237*** (0.0214) | 16.82*** (4.24) | -0.1623 (0.1824) | -3.06 (3.87) |
| β_{43} | -0.0078*** (0.0003) | -0.2872*** (0.0008) | 0.0096*** (0.0001) | 0.0567** (0.0248) |
| $H_0 : \beta_{40} = \beta_{41} = \beta_{42} = 0,$ p-value | <0.01 | <0.01 | <0.01 | <0.01 |
| Number of individuals | 496,586 | 496,586 | 496,586 | 496,586 |

Notes: Least squares estimates of $\beta_{4j}, j = 0, 1, 2, 3$ from (4). Standard errors adjusted for clustering at the individual level in parentheses. ppt = percentage points. The number of individuals is the total across all treatment and comparison groups. For the numbers in each group, see Table 1. Number of observations is the number of individuals multiplied by 5. ***, ** & * indicate significance at the 1%, 5% and 10% levels respectively.

Figure B1: Relationship between earnings response to reassessment and duration of DI claim by cause of disability



Notes: Estimates of average effects of reassessment under stricter DI criteria by length of time on DI at entry to the sample. Derived from least squares estimates of $\beta_{4j}, j = 0, 1, 2, 3$ from (4) estimated separately for each sub-sample. Shading indicates 95% confidence interval computed from delta method standard errors adjusted for clustering at the individual level.