Who can go back to work? Employment effects of disability benefit cuts by diagnosis

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

Using administrative data on the universe of disability insurance (DI) claimants in the Netherlands, we estimate effects on employment, earnings and other social assistance of cutting benefits of existing DI claimants. More stringent qualification rules resulted in termination of the benefits of one fifth of claimants and reduction in payments made to an additional 12%. We identify the impact of these cuts on claimants aged 30-44 by comparing the change in their employment (earnings/income) with that of older claimants not subject to the increased stringency, while adjusting for differential trends by age prior to the reform. We estimate that about one half of those forced or induced to leave DI entered employment. The average increase in earnings was 45% of the average loss of DI income. Consistent with the hypothesis that work capacity is higher among DI claimants who qualify through more subjective health conditions, such as back pain and stress-related problems, the employment response was particularly large for claimants with musculoskeletal and mental health diagnoses. However, these claimants were not able to replace a larger fraction of their lost DI income with labor market earnings. On average, 26% of the income lost from DI was replaced through increased dependency on other social assistance programs. This fraction was highest (37%) for claimants with mental health conditions.

JEL: H53, I38, J14, J22

Keywords: Disability Insurance, Social Security, Health, Labor Force Participation

* Garcia Mandico (email: garciamandico@ese.eur.nl), Garcia-Gomez (email: garciagomez@ese.eur.nl), Gielen (email: gielen@ese.nl) , O'Donnell (email: odonnell@ese.eur.nl): Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands and Tinbergen Institute. O'Donnell: School of Economics and Area Studies, University of Macedonia, Thessaloniki, Greece.

I. Introduction

Disability insurance (DI) rolls have often lengthened at the same time that population health indicators have improved. This suggests there is considerable capacity for work among the pool of DI claimants. If this is true, then retrenchment of DI should be possible without jeopardizing the economic wellbeing of individuals with health problems. To some extent, labor market earnings could replace DI income without increasing dependency on other social assistance programs. This logic underpins efforts to rein in DI programs.

This paper provides evidence on the work capacity of the universe of Dutch DI claimants aged 30-44 by using administrative data to estimate the impact on employment, earnings and social assistance income of increased program stringency that terminated the benefits of one fifth of claimants and reduced the benefits of an additional 12%. The Netherlands is an interesting case study in DI policy. The country is infamous for a DI dependency rate that reached 12% of the insured population at the beginning of the 1990s (Koning and Lindeboom 2014), while also being commended for a series of reforms that are claimed to have contributed to a 40% fall in this dependency rate. Countries, such as the US, searching for ways to make the mounting fiscal burden of DI manageable can potentially learn from the Dutch experience (Autor 2015).

Much of the evidence on the employment effects of DI comes from studies that follow Bound (1989) in using the employment of rejected applicants to give an upper bound on the work capacity of claimants (Chen and van der Klaauw 2008; von Wachter, Manchester and Song 2011). Exploitation of plausibly exogenous variation in the award probability can eliminate the upward bias (Maestas, Mullen and Strand 2013; French and Song 2014). While this evidence is pertinent to gauging the impact of policies that tighten entry to DI, it is less relevant to assessing the potential of reforms that aim to release the work capacity of existing DI claimants. Since skills and preferences for work may deteriorate with time spent on DI (Vingard et al. 2004, Bryngelson et al. 2009; Svensson et al. 2010), work capacity at entry is not necessarily indicative of the average capacity in a pool of claimants.

Tightening the screening of DI applications has been shown to be effective in reducing program entry (Autor and Duggan 2003; de Jong, Lindeboom and van der Klaauw 2011).

But without increasing the rate of exit, which tends to be very low (Autor and Duggan 2003; Moore 2015), it is a long route to reducing the stock of claimants, and so program costs. Unfortunately, there is only sparse evidence on the work capacity of claimants that would be targeted by efforts to increase exit. Moore (2015) finds that termination of the benefits of a subset of DI claimants in the US increased the likelihood of moving into employment by 22 percentage points – a large effect relative to the employment histories of those affected. But this curtailment affected only 2% of DI claimants who had qualified, at least in part, through alcohol and drug addiction. Their response is not necessarily indicative of the work capacity, and preferences, of the majority of claimants with more prevalent conditions. Further, the narrow targeting of this measure may have encouraged some of those affected to seek re-qualification through a different health condition (Autor 2015).

Borghans, Gielen and Luttmer (2014) evaluate a more universal scale-back of DI that affected all claimants younger than 45 in the Netherlands. This occurred a decade before the reform we examine. They estimate that measures that cut the benefits of existing claimants by about 10%, on average, reduced the DI participation rate by 3.8 percentage points and raised the employment probability by 2.9 points. Our main contribution relative to this paper is to examine whether and how the work capacity of DI claimants varies with the health condition through which they qualify for DI. Evidence on such heterogeneity is relevant to one of the main explanations proposed for the growth of DI rolls. According to this hypothesis, 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 verify medically (Autor and Duggan 2006, Milligan and Wise 2011). Lower back pain and stress-related problems are the stereotypical examples. In 2010, individuals suffering from mental health disorders constituted one half of existing DI claimants and 27-47% of new awards across all OECD countries (OECD, 2012). Musculoskelatal problems are typically the second most common reason given for a DI claim (SSN, 2013). If this explanation for the growth of DI programs is correct, then claimants with more subjectively defined conditions, such as some mental health and musculoskeletal problems, should have greater capacity to work. Evidence on this is scant and inconsistent.

we use the employment response to more stringent benefit entitlement rules that applied

to all DI beneficiaries irrespective of qualifying health condition to provide the first detailed evidence on variation by diagnosis in the work capacity of a stock of DI claimants. The reform we examine did not change the medical criteria for awarding DI. Rather, it changed the method used to calculate potential earnings (and so benefit entitlement) for any given health condition. This allows us to compare the effects of increased DI stringency on employment across diagnoses.

In addition to providing evidence on work capacity by diagnosis, this is one of the few studies to estimate the work capacity of younger DI claimants from their response to benefit cuts. Borghans et al (2014) identify this effect at the age of 45, while Moore (2015) estimates it for claimants aged 30-61 who qualified through addiction and finds stronger effects at younger ages. We estimate the effect on all DI claimants aged 30-44. This is the age group that would be expected to have the greatest residual work capacity and is likely to be the principal target of any attempt to reduce benefits of existing claimants. [1]

The reform we evaluate involved reassessment of the earnings capacity of the whole stock of DI claimants younger than 45. The reassessment applied stricter rules that could result in complete or partial withdrawal of benefits. We identify the effect by comparing the pre-post change in employment (and other outcomes) with the respective change experienced by older claimants not exposed to the reform. Unlike other studies that rely on difference-in-differences (DID) across age groups to identify the employment impact of more stringent DI (Karlstrom et al 2008; Staubli 2011), we adjust for the difference between the age groups in the employment trend over a period prior to the reform. This trend-adjusted DID (Bell, Blundell and van Reenen 1999) eliminates age-specific trends, as well as period effects. Identification of the treatment effect rests on the assumption that, in the absence of the reform, the age differential in the employment trend would have been that observed in the earlier period. Essentially, this rules out cohort effects. Consistent with this, we demonstrate that trends do not differ by cohort over periods prior to the reform.

[1] Karlstrom et al (2008) find that tightening DI eligibility criteria for those aged 60+ had no impact on employment of this older group in Sweden. At the slightly younger age of 55-56, Staubli (2011) finds a positive impact of reduced DI entitlement on employment in Austria. A few studies that do not use variation arising from benefit cuts find stronger labor supply responses of younger DI claimants (von Wachter 2011; Kostol and Mogstad 2014).

We estimate that application of the more stringent rules cut the benefits received by 14.3%, on average, and raised the rate of transition out of DI by 13 percentage points.[2] The retrenchment had a positive impact on employment. We estimate that the probability of moving into employment was raised 6.3 percentage points. This indicates the work capacity existing in a pool of DI claimants aged 30-44 that can be released through full screening and benefit cuts of the order of 14%, on average. Roughly half of those forced or induced to leave DI entered employment. Work capacity was therefore substantial among the claimants whose degree of disability was revised downward sufficiently to cause them to leave DI. However, these were the minority in the stock of claimants.

We estimate that the cut to DI benefits induced claimants to raise their labor market earnings by 11.5%, on average. In aggregate, increased earnings replaced 44% of the reduction in DI benefits. A further 26% of the loss of DI benefits was compensated through increase reliance on other social assistance programs. This sizable spillover effect is consistent with evidence produced by Borghans et al (2014) from the earlier retrenchment of the Dutch DI program. It signals that the fiscal impact of cutting DI benefits falls well short of the savings to that program alone. The increases in earnings and other social assistance are not sufficient, on average, to fully offset the loss of DI benefits. On average, the income of the claimants who were reassessed under the tighter rules was reduced by 3.45%. This average obscures the much larger fall in income experienced by those whose benefits were cut severely, or even terminated.

While the cuts to DI benefits caused claimants to increase their earnings, it also resulted in lower earnings per day worked. The daily wage fell by around 11%, on average. This suggests that the claimants who were induced to work by the rationalization of the program were less productive than those who would have left DI without the pressure of lower benefits. The cuts caused claimants to settle for lower quality, less well-paying jobs.

Application of the stricter entitlement rules impacted most on DI participation and benefits of claimants with musculoskeletal problems. Their probability of continuing to claim

[2]Benefits were terminated for 20% of the claimants who were reassessed. Our estimated impact on the DI exit rate is smaller because it nets out claimants who would have left DI even if their benefits had not been terminated. In addition, our estimate takes account of re-entry to DI after a termination of benefits.

DI was reduced by 19 percentage points and their benefits were cut by almost 20%, on average. Claimants with mental health problems were also among those who experienced the greatest cuts in benefits. Employment was induced to rise most for those suffering from musculoskeletal conditions. The cuts produced an 8.2 percentage point increase in the rate of employment in this group. For those with mental health disorders, the rate went up by 6.7 points. At the other extreme, the employment rate was raised by only 2.8 percentage points for claimants with cardiovascular disease (CVD). The comparison indicates that claimants with musculoskeletal and mental health conditions do tend to have greater work capacity. The greater impact on the employment of claimants in these two groups is entirely due to the stronger effect on their DI participation. In fact, the proportion of claimants induced to leave DI who entered employment was actually lower in these two groups than it was for claimants with CVD.

Claimants with musculoskeletal and mental health disorders were able to replace 44% and 54% respectively of their lost DI benefits with increased earnings. These fractions are not higher than those achieved by claimants in all of the other diagnostic categories. Claimants with mental health problems make the greatest use of other social assistance to compensate for the loss of DI. For each €1000 of DI benefits lost, claimants in this category received €370 more from other programs. Those with musculoskeletal problems received only €130. Claimants with mental health conditions appear to be highly welfare dependent. When their benefits from one social assistance program is cut, they substitute at a high rate with transfers from another program.

There is indirect and mixed evidence of heterogeneity in the work capacity of DI claimants from studies based on rejected DI applicants in the US (von Watcher et al. 2011; Maestas, Mullen and Strand 2013; French and Song 2014). von Wachter et al (2011) note that rejected DI applicants who reported mental health and musculoskeletal problems had the highest subsequent employment rates. Maestas et al. (2013) correct for unobservable differences between accepted and rejected DI applicants and find no evidence that suggests DI claimants reporting these two types of conditions have greater work capacity. In fact, they estimate that the employment potential of claimants with musculoskeletal problems is almost half that of those with other diagnoses. French and Song (2014), who also attempt to take account of

differences in unobservables, arrive at the opposite conclusion: claimants with musculoskeletal problems have greater work capacity.[3] Moore (2015) provides more direct evidence of variation in work capacity by diagnosis. He finds that those with a primary diagnosis of a mental health or a musculoskeletal condition were more likely to work after being disqualified from DI. The limitation of this evidence, as mentioned before, is that it is obtained only from claimants who partly qualified for DI through addiction.

The next section describes the Dutch DI program and the reform we evaluate. Section III sets out our identification strategy. Section IV describes the Dutch administrative data used, presents descriptives, examines trends in the outcomes and explains the method of estimation. Results are presented in section V. The final section concludes.

II. Dutch Disability Insurance

Even after a decade of belt-tightening, the Dutch DI program remained relatively generous in 2004. Replacement rates, at a maximum of 70% for full disability, were higher than in most countries (OECD, 2009). The program offers compensation for partial disability. A claimant assessed as less than full disabled is permitted to work and supplement her earnings with DI provided she does not earn more than 85% of what she did prior to becoming disabled. This makes the Dutch program a forerunner of return-to-work incentives that have been introduced or contemplated more recently in countries such as the UK and the US (Kostol and Magstad 2014; Autor 2015).

A. Eligibility and benefits

Before 2004, an application for DI could be made if the person considered herself unable to resume her normal work schedule after receiving sick pay for one year.[4] The Social Security Administration (SSA, or UWV using the Dutch acronym) would make a medical assessment to establish whether the applicant was completely incapable of work and to identify work

[3]Maestas et al. (2013) instrument DI status with the award propensity of the case officer. French and Song (2014) concentrate on DI applications that go to appeal and instrument DI status using the record of the presiding judge. The discrepancy between the findings of the studies could lie in the fact that applicants suffering from musculoskeletal disorders are more likely to appeal a rejection. During the appeal process, the plaintiff may keep her earnings low to increase to signal low work capacity.

[4]We describe the disability program for 30-65 year-old workers that operates under the Invalidity Insurance Act (*Wet op de arbeidsongeschiktheidsverzekering,* WAO) and covers around 80% of DI claimants. Separate programs for younger individuals and the self-employed were also changed in 2004 but we do not consider them since the details of the reforms differ and the employment patterns of claimants qualifying for of these programs differ markedly from those of claimants of the main program for prime-aged workers.

activities that she could manage given her medical condition. If the SSA physician judged that the applicant had some residual work capacity, then a vocational expert would assess the applicant to identify specific occupations from a very detailed list that she was considered capable of performing, taking into account her educational attainment.

The applicants lost earnings capacity, or degree of disability, was calculated as 1 minus the ratio of potential earnings to pre-disability earnings. Earnings potential was the average salary across the three highest paying feasible occupations that each had at least ten jobs currently available.[5] If the estimated loss of earnings capacity was less than 15%, then the claim was rejected. Otherwise, disability benefits were paid at approximately 70% of lost earnings.[6]

Outflow from DI was low and consisted mostly of claimants who found employment paying more than 85% of their earnings prior to entering the program, as well as working claimants who increased their earnings above this threshold. Once the threshold was breached, a claimant continued to receive her benefit during a three-month trial period before entitlement was terminated. Work capacity was reassessed one year after a claim was awarded and every five years thereafter. These reassessments were often based on no more than a postal questionnaire returned by the claimant. It is claimed that they induced less than a fifth of the outflow from DI (Work and Income Inspectorate 2004).

B. The 2004 Reform

From October 2004, a stock of about 275,000 DI claimants younger than 50 (on July 1, 2004) became eligible for reassessment under stricter qualification criteria.[7] The outcome could be complete or partial withdrawal of benefits.

Each claimant below the age threshold was called for reassessment at her local SSA office at some time between October 2004 and May 2009. Many were required to undergo medical assessment by a physician. All had their lost earnings capacity re-evaluated under new rules that could result in upward revision of current potential earnings and downward revision of pre-disability earnings. For any given a health condition and associated functional limita-

[5]The wage attached to each specific occupation was essentially the average wage for that occupation in the applicant's region and at her level of education.

[6]Specifically, the replacement rate was set at 70% of the mid-point of each category of lost earnings capacity defined in Appendix C, Table C3.

[7]The reform was legislated in April 2004 and the initial plan was to start the reassessments from July 2004. Because of strong opposition and lack of consensus about the reassessment criteria, implementation was pushed back to October.

tions, lost earnings capacity under the new rules could not be greater than it was under the old rules. In many cases, it would be lower resulting in a reduction or termination of benefits. The main reason benefit entitlement could be revised downward was that potential earnings were now estimated by averaging over the three highest paying occupations the claimant could perform that each had at least three jobs available. Previously, only occupations with at least ten available jobs had qualified. In addition, jobs requiring Dutch language proficiency and knowledge of IT were considered feasible even if the claimant did not possess those skills. Full-time and night work were also considered even if the claimant had not previously engaged in that type of employment. As a result of expanding the pool of potential work, the average wage in the three highest paying feasible occupations could not be reduced, and may have been raised. In the latter case, benefit entitlement would fall. 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, then there would be a downward revision of lost earnings capacity, and so benefits.

These changes resulted in around one third of all reassessed claimants having their benefits reduced or terminated (UWV 2009). For about one fifth, lost earnings capacity was reassessed to be below the 15% minimum threshold and entitlement to DI was withdrawn. 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 claimants having their benefits raised despite application of the more stringent rules.

Prior to the reform, around three fifths of claimants were assessed as fully disabled (i.e. lost earnings capacity $\geq 80\%$). Many of them remained on full disability benefits, but in the first year of implementation, when younger claimants with higher average work capacity were predominantly called for reassessment, about a quarter had their benefits terminated (UWV 2006).

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 claimant could transfer to Unemployment Insurance (UI) if she was still eligible for that program. If the claimant was not eligible for UI, or had entitlement for less than six months, then she could apply for 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 entitlement was reduced or terminated were granted benefits from this program (see Table C1 in Appendix C).

In 2007, strong criticism of the policy and a change of government resulted in the age threshold for reassessment eligibility being revised from less than 50 to less than 45 on July 1, 2004. As a result, around 25,000 claimants aged 45-49 who had already been reassessed were assessed once more under the old, more lenient rules (Staatsblad 2007, 324).

Initially, the plan was to reassess all younger claimants before moving to older groups. In practice, although younger claimants were more likely to be reassessed earlier, there was a positive probability of reassessment for all ages throughout the 4.5 years it took complete implementation.

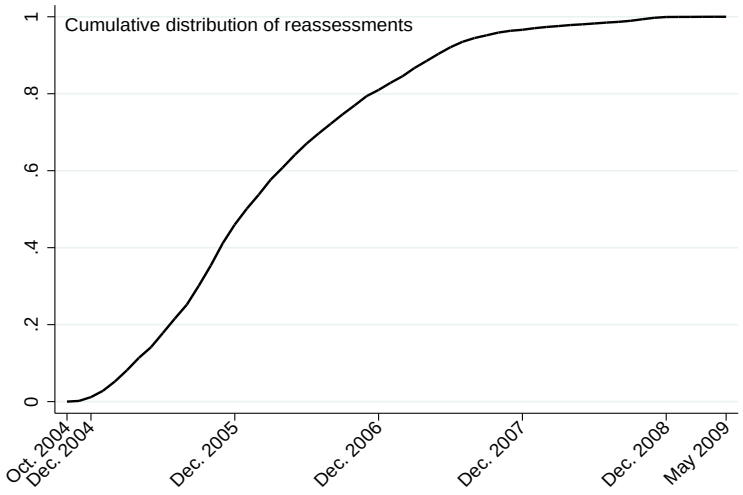
Figure 1 shows the cumulative distribution of reassessments of claimants aged 30-44 in July 2004, which is the cohort for which we estimate effects. Few reassessments were conducted in the last quarter of 2004. The pace quickened in 2005. By the end of that year, 46% of this claimant stock had been reassessed. By the end of 2008, which is the limit of the follow-up period over which we estimate effects, effectively all (99.93%) claimants had been reassessed.

III. Identification

We are interested in the average effect of assessing work capacity using more stringent criteria on DI claimants who were reassessed. We observe the DI and employment status of each individual every month but we cannot separate those who had been reassessed from those who had not. Given the near universal exposure to reassessment by the end of the follow-up, lack of individual level data on who was reassessed does not pose any problem for identification of the impact of reassessment at that time. But it does present a problem for identification of the effect at early points in the implementation period. Monitoring the evolution of the effect over time is interesting since it can reveal how claimants transit through labor market and social assistance states after DI benefits are cut.

We deal with the data limitation by estimating the effect of being in the age group that is

Figure 1. : Cumulative distribution of reassessments of claimants aged 30-44 on July 1, 2004



Source: Authors' calculations using UWV data

eligible for reassessment and scaling this intention-to-treat effect by the proportion of eligible claimants who had been reassessed by each point within the implementation period. Since claimants not in the target age group could not be reassessed under the stricter rules, the scaled estimate corresponds to the average effect of reassessment on those reassessed. [8]

Let Y_{it} be the observed outcome of individual i at time t. We estimate effects on a number of outcomes but employment is of primary interest and we will often refer to this in the exposition of methods. Let Y_{it}^0 and Y_{it}^1 represent potential outcomes without and with reassessment respectively. Consider a simplified set up with one observation point before reassessments begin, *t=0*, and one after reassessments commence but before they have been completed, *t=1*. At *t=0*, $Y_{i0} = Y_{i0}^0, \forall i$. At *t=1*, $Y_{i1} = D_i Y_{i1}^1 + (1 - D_i) Y_{i1}^0$, where D_i is 1 if i has been reassessed by *t=1* and is 0 otherwise.

Let Z_i be a binary indicator equal to 1 if, on the basis of date-of-birth, the individual has a positive probability of being selected for reassessment. If $Z_i = 0$, then $D_i = 0$. Let Y_{it}^E and Y_{it}^N represent potential outcomes if the individual were made eligible and not eligible for reassessment respectively. The intention-to-treat effect on those eligible is,

[8] With random assignment to treatment and one-sided noncompliance, it is well known that the average treatment effect on the treated is the ratio of the intention-to-treat effect (ITT) to the proportion of compliers (Bloom 1984; Angrist and Pischke 2009, p.164). The result holds without random assignment provided the ITT is identified.

$$\begin{aligned} ITTE &= \mathbb{E}[\ Y_{i1}^{E} - Y_{i1}^{N} \mid Z_i = 1\ ] \\ &= \mathbb{E}[\ D_i Y_{i1}^{1} + (1 - D_i) Y_{i1}^{0}) - Y_{i1}^{0} \mid Z_i = 1\ ] \\ &= \mathbb{E}[\ D_i (Y_{i1}^{1} - Y_{i1}^{0}) \mid Z_i = 1\ ] \end{aligned} \tag{1}$$

In general, $\mathbb{E}[\ D_i(Y_{i1}^{1} - Y_{i1}^{0}) \mid Z_i = 1\ ] \neq \mathbb{E}[\ D_i(Y_{i1}^{1} - Y_{i1}^{0})\ ]$ because Z_i is not randomly assigned, and both D_i and $\beta_i = (Y_{i1}^{1} - Y_{i1}^{0})$ could vary with Z_i, which is defined by age. Given those not eligible cannot be reassessed,

$$\begin{aligned} ATET &= \mathbb{E}[\ Y_{i1}^{1} - Y_{i1}^{0} \mid D_i = 1\ ] \\ &= \frac{\mathbb{E}[\ D_i(Y_{i1}^{1} - Y_{i1}^{0}) \mid Z_i = 1\ ]}{P(D_i = 1 \mid Z_i = 1)} = \frac{ITTE}{P(D_i = 1 \mid Z_i = 1)}, \end{aligned} \tag{2}$$

(see Apppendix A).

Since the SSA provides the denominator of (2) in each month from the start of the reassessment period, identification of the ATET rests on identification of the numerator.[9] One strategy that might be considered to identify the *ITTE* would involve a difference-in-differences (DID) comparison across younger claimants who were potentially subject to reassessment and older claimants who were not. This is likely to be problematic, particularly as the age gap widens, since older DI beneficiaries would have a lower probability of returning to work than younger recipients even if the latter were not subject to reassessment.

An alternative comparison group would be DI claimants 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 in the reform period would have changed by the amount observed for the same age group in the earlier period, then the DID across the periods will identify the effect of the reform. Let t_{00} and t_{10} be two points in time prior to t_0 with $t_{10} - t_{00} = t_1 - t_0$. The assumption required for this DID to identify the ITTE is,

$$\mathbb{E}[\ Y_{i1}^{0} - Y_{i0}^{0} \mid Z_i = 1\ ] = \mathbb{E}[\ (Y_{i10}^{0} - Y_{i00}^{0}) \mid Z_i = 1\ ] \tag{3}$$

[9]We observe the proportion of the target group of claimants reassessed each month within sub-groups defined by age and diagnosis. Hence, we can identify heterogeneous effects across these sub-groups.

The threat to this identification strategy comes from period-specific labor market conditions that would invalidate using the earlier period to identify counterfactual employment of the target age group in the absence of the reform.

Our preferred strategy makes use of both comparison groups – older claimants in the same period and claimants 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 or other of the comparisons. We use a period that precedes the introduction of the reform to identify the extent to which the trend in employment of younger DI claimants differs from the trend of older claimants. Effectively, we subtract this differential trend from the DID in employment between the age groups over the period in which the reform is introduced. This differential trend adjusted difference-in-differences (DADID) (Bell et al 1999; Blundell and Costa-Dias 2009) relaxes the assumption of common trends in employment across age groups in the absence of the reform. It also avoids the assumption of common trends in employment for a given age group across periods. The assumption that is required for identification is that the age differential in the employment trends would have been common across periods in the absence of the reform. If this 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 employment of younger claimants.
The assumption required for DADID to identify the ITTE is,

$$\begin{aligned}&\mathbb{E}[\ Y^0_{i1} - Y^0_{i0} \mid Z_i = 1\ ] - \mathbb{E}[\ (Y^0_{i1} - Y^0_{i0}) \mid Z_i = 0\ ] \\ &= \mathbb{E}[\ Y^0_{i10} - Y^0_{i00} \mid Z_i = 1\ ] - \mathbb{E}[\ (Y^0_{i10} - Y^0_{i00}) \mid Z_i = 0\ ]\end{aligned} \tag{4}$$

If this holds, then

$$\begin{aligned}ITTE =& \mathbb{E}[\ Y_{i1} \mid Z_i = 1, t = t_1\ ] - \mathbb{E}[\ Y_{i0} \mid Z_i = 1, t = t_0\ ] \\ &- \big(\mathbb{E}[\ Y_{i1} \mid Z_i = 0, t = t_1\ ] - \mathbb{E}[\ Y_{i0} \mid Z_i = 0, t = t_0\ ]\big) \\ &- \Big\{\big(\mathbb{E}[\ Y_{i10} \mid Z_i = 1, t = t_{10}\ ] - \mathbb{E}[\ Y_{i00} \mid Z_i = 1, t = t_{00}\ ]\big) \\ &- \big(\mathbb{E}[\ Y_{i10} \mid Z_i = 0, t = t_{10}\ ] - \mathbb{E}[\ Y_{i00} \mid Z_i = 0, t = t_{00}\ ]\big)\Big\}\end{aligned} \tag{5}$$

IV. Data and Descriptives

A. Data Sources and Outcomes

We use administrative data maintained by Statistics Netherlands. Social Security records provide information on DI claimants: duration of claim, main diagnosis, assessed loss of earnings capacity and benefit amount paid. We estimate effects the probability of claiming DI, the replacement rate at which benefits are paid and the money (€) amount received. The replacement rate is determined using the information provided on assessed loss of earnings capacity (see Appendix C Table C3).

We use the main diagnosis on entry to DI to distinguish between six major diagnostic categories that each have a prevalence of at least one percent in the stock of claimants. They are cardiovascular, musculoskeletal, neurological, mental health, respiratory and digestive disorders, plus a group labeled 'general diagnoses'. [10]. Cardiovascular, respiratory and digestive disorders each include conditions that are more objectively medically defined. Musculoskeletal and mental health disorders include a greater number of conditions that are more subjectively defined. Neurological conditions range from precisely defined diseases of the nervous system to more difficult to diagnose conditions, such as ME/Chronic Fatigue Syndrome.

The social security files also identify spells on UI and other social assistance programs, and corresponding payments. We aggregate payments from UI, the main means-tested welfare program, sickness pay and smaller social assistance programs to obtain income from social assistance other than DI.

Information on employment and self-employment (days worked and earnings) is obtained from the tax records, as is total individual income from all sources. We only have information on income received by the individual. We cannot examine, for example, the response of spousal earnings or total household welfare transfers to the loss of DI benefits.

Participation in employment and each transfer program, including DI, is identified each month.[11] Earnings, benefit amounts and total income are measured per calendar year.

[10] The latter category includes a hodgepodge of diseases and conditions ranging from tuberculosis and HIV to pain. See Appendix B

[11] We identify an individual as being employed (enrolled in a program) if she worked (claimed benefits) for at least one day in the month.

Month and year of birth are obtained from the municipal registers. Deaths are identified from the death register.

Each administrative file records an individual's identification number (RIN-code) that is issued on compulsory registration with the municipality at birth or after immigration. This is used to merge all records at the individual level. We use data from January 1999, when they first become available, to December 2008.

B. Treatment and comparison groups

We distinguish between a *reform period* from January 2004 to December 2008 that spans the beginning of the reassessments in October 2004 and a *non-reform period* from January 1999 to December 2003 that ends before any reassessments have taken place. Although a small proportion of claimants were reassessed in the first months of 2009, we curtail the follow-up at December 2008 in order that the reform period is the same length as the non-reform period, which is constrained by the unavailability of data prior to January 1999.

The proximity of the non-reform period to the reform period increases the plausibility of the identification assumptions. One potential threat to identification is a major DI reform introduced in 2002. This involved making the employer and employee jointly responsible for taking active measures to enable the latter to continue working during the waiting period for DI. Unless proof of such actions could be offered, a DI claim would not be considered. This reform is credited with substantially reducing the rate of inflow into DI (de Jong et al 2011; Van Sonsbeek & Gradus 2013; Koning & Lindeboom 2015). Any impact on the DI exit rate and the employment of DI claimants, which are relevant in the present context, would be indirect. Most pertinent is that there is no reason to expect this reform to have affected the employment of younger claimants deferentially from older claimants. Only if it did have such a differential effect would it invalidate using the 1999-2003 period to construct the counterfactual of what would have happened in 2004-2008 in the absence of the reform we evaluate.

To construct the reform period sample, we select individuals who were claiming DI in January 2004 and could be followed until the end of 2008. Only the deceased and those exiting

from all benefit programs and from (self-)employment, possibly due to emigration, are lost to follow-up. [12]

Initially, the SSA planned to reassess all DI claimants aged less than 50 on July 1, 2004. We exclude claimants aged 45-49 on that date because the reversal of policy in 2007 meant that claimants in this group either were never reassessed or they had their reassessment reversed. We also exclude those 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.

The target treatment group consists of 160,855 individuals who were claiming DI in January 2004, were aged 30-44 on July 1, 2004 and so were eligible for reassessment some time between October 2004 and May 2009. [13] One of our comparison groups comprises 95,326 individuals who were claiming DI in 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) or 50-53 (as reform period comparison group) on July 1, 1999, and could be followed until 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. January 1999 and 2004, by age group and period. At the time of entry into the respective panel, all observations are DI claimants. The proportion of females is higher in the later, reform period samples. This reflects the increasing female labor participation in the Netherlands. Given the feminization of the claimant stock occurs in both age groups, it is not expected to generate an age differential in the outcome trends between the periods and so should not compromise the DADID identification strategy. The proportion of claimants in the younger group who are native Dutch decreases between periods. This reflects the increase in immigration. We allow for individual fixed effects in all the models estimated, which should be sufficient to deal with this change in sample composition.

[12] Only 3.9% of DI claimants at the beginning of the period cannot be traced to the end. We have confirmed that there the age differential in mortality rates does not differ between the reform and non-reform periods.

[13] The number of claimants from this group who were actually reassessed is 137,419. The others were not reassessed either because they left DI before being called for reassessment (6.08%) or because they had a clear incapacity to work that could be established without assessment.

At entry to the panel, the reform period sample has a slightly shorter duration on DI, has a slightly lower proportion of fully disabled and a lower mean replacement rate. Despite these indicators suggesting that the reform period sample is slightly less disabled, the mean DI payments are higher in the later period. In addition to these differences being marginal, they are also observed in both age groups (with the exception of there being no reduction in the mean replacement rate in the younger group). Consequently, they would not appear to pose a threat to the validity of the identification approach.

At sample entry, the younger group had lower employment and earnings in the reform period than in the earlier non-reform period, while the opposite is true for the older group. Differences in levels within age groups across periods does not invalidate the DADID identification strategy, but it may cause one to question whether the assumption of a constant age differential in the trends under the counterfactual is plausible. We compare trends in the next sub-section.

Income in total and from social assistance programs other than DI are higher for both age groups at the start of the reform period compared with the start of the non-reform period. Since incomes improved for both age groups, the increase is not of particular concern.

For three of the four age-period groups, mental health problems represent the most prevalent category of conditions. For both age groups, the prevalence of these conditions increases within the claimant stock between the non-reform and the reform periods. Musculoskeletal problems form the most prevalent category in the fourth group and is the second most prevalent set of conditions in the three other groups. In the treatment group (30-44 in the reform period) these two categories, which include many of the disabilities that are often considered more subjective, account for around 72% of all claimants. The proportion of claimants qualifying through diagnoses that are labeled 'general' declines for both age groups between the periods.

Table 1—: Characteristics of DI claimants by period and age group - means at sample entry

	Reform period sample		Non-reform period sample	
	Age 30-44	Age 50-53	Age 30-44	Age 50-53
Panel A. Demographics				
Female	0.60	0.46	0.53	0.37
Year of birth	1965.29	1951.93	1960.23	1946.94
Native Dutch	0.75	0.82	0.80	0.82
Panel B. Disability Insurance				
Time on DI (years)	5.44	9.51	5.90	9.95
Proportion fully disabled	0.63	0.64	0.65	0.69
Replacement rate	0.54	0.55	0.54	0.57
Benefit amount (€)	10,074	12,340	8,812	11,081
Panel C. Labor market activity				
Employment participation	32.40	35.62	37.01	34.46
Earnings (€)	4,844	6,465	5,089	5,456
Panel D. Income				
Total (€)	15,869	19,556	14,241	16,826
from other social assistance (€)	956	755	598	519
Panel E. Diagnostic group				
General	9.03	12.24	16.70	17.32
Cardiovascular	1.92	5.09	1.84	5.25
Musculoskeletal	28.80	32.77	26.11	32.63
Neurological	6.54	5.50	5.30	3.82
Mental health	43.09	33.73	36.14	29.35
Respiratory	1.28	2.09	1.31	1.97
Digestive	2.42	2.16	2.14	2.13
Observations	160,768	95,326	140,282	103,355

Note: The left (right) panel refers to DI claimants selected in January 2004 (1999). Columns within each panel are split by age on July 1, 2004 (left panel) and July 1, 1999 (right panel). The extreme left column corresponds to the target treatment group. All others are comparison groups. Income, earnings and benefit levels are measured yearly, and inflated by the CPI to 2015 price levels (Eurostat Netherlands HCPI 2015).Income, earnings and benefit levels are measured yearly, and adjusted by CPI to 2015 price levels (Eurostat Netherlands HCPI 2015).

There are some notable differences by diagnostic group in the characteristics of the treatment group of claimants aged 30-44 in July 2004. As is clear from the means presented in Table 2 females are over-represented in all diagnostic groups except cardiovascular disease. The gender imbalance is greatest among claimants with digestive and mental health conditions. The proportion of fully disabled claimants is highest for those with mental health and neurological conditions and lowest for those with musculoskeletal problems. The ranking is the same for the replacement rate, which is a positive function only of the assessed loss of earnings capacity. The group with musculoskeletal problems has the highest proportion of claimants with the lowest degree of disability. Consequently, for any given tightening of the rules, the proportion pushed below the threshold at which benefits are paid will be largest for this group.

Labor market activity varies a great deal. Only a quarter of claimants with a mental health condition are working, while more than a half of those suffering from a digestive ailment are employed. Reflecting this, earnings are lowest among those with mental health problems. On average, they earn little more than half of what those with a cardiovascular or musculoskeletal condition earn. While claimants with a musculoskeletal problem received the lowest average payment from DI, they are the most dependent on other social assistance.

Table 2—: Characteristics of DI claimants exposed to reform by diagnosis - means at sample entry

	General	Cardiovascular	Musculoskeletal	Neurological	Mental	Respiratory	Digestive
Panel A. Demographics							
Female	0.60	0.49	0.53	0.59	0.63	0.55	0.66
Year of birth	1964.76	1964.08	1965.14	1965.32	1965.52	1964.54	1965.47
Native Dutch	0.76	0.78	0.77	0.83	0.72	0.69	0.77
Panel B. Disability Insurance							
Time on DI (years)	5.84	4.24	4.45	5.85	5.03	5.30	5.72
Proportion fully disabled	0.59	0.62	0.51	0.71	0.72	0.63	0.61
Replacement rate	0.53	0.54	0.47	0.58	0.58	0.54	0.54
Benefit amount (€)	11,165	9,661	8,644	10,512	10,879	9,984	9,714
Panel C. Labor market activity							
Employment participation	0.38	0.39	0.40	0.35	0.25	0.36	0.53
Earnings (€)	6,068	6,418	6,267	5,305	3,390	4,638	5,364
Panel D. Income							
Total (€)	17,893	17,058	16,201	16,557	15,110	15,703	15,802
from other social assistance (€)	664	985	1,296	746	845	1,086	729
Observations	14,510	3,091	46,306	10,522	69,276	2,058	3,891

Note: This table shows the means in January 2004 for the treatment group of claimants aged 30-44 in July 2004. Income, earnings and benefit levels are measured yearly, and are inflated by the CPI to 2015 price levels (Eurostat Netherlands HCPI 2015).

C. Trends

The top panel of Figure 2 shows the proportion of each group on DI and in employment each month over a four-year period. For the reform-period sample, the trends are shown from January 2004 to December 2008. For the non-reform period sample, an analogous four-year period starts in January 1999 and ends in December 2003. Month 0 is October 2004 – when reassessments started – for the reform-period sample and October 1999 for the non-reform period sample. DI participation is 100% for all groups at baseline by construction. The baseline employment rates are above zero because it is possible to work while claiming DI.

In both periods, the younger group has a much more rapid rate of exit from DI and a greater rate of transition into employment. In the reform period sample, these differential trends by age are already evident in the 10 months before reassessments started. This, together with the fact that the rate of both DI exit and employment entry differ by age throughout the non-reform period, confirms that DID across age groups is unlikely to identify the effect of the reform. [14]

Comparing the 30-44 age group across periods, there is little or no difference in the trends at the beginning of the respective periods. Younger claimants were exiting DI slightly quicker between January and October 2004 than were claimants of the same age between January and October 1999. While the difference is marginal, the faster rate of exit in the later period could possibly be due to anticipation of the reform. Claimants expecting a negative reassessment may have taken any opportunity to leave DI in the six months that elapsed between legislation and the start of implementation of the reform. If reform was anticipated and reacted to in this way, then our strategy will deliver lower bound estimates of its impact. But two observations incline against this interpretation. First, those aged 30-44 in 2004 were not entering employment before month 0 at a faster rate than individuals of the same age observed in the earlier period. Without the opportunity of employment, there would be no incentive for a claimant expecting to lose DI benefits to leave the program any earlier than she was forced to do so. Second, before month 0, the DI exit rate was also slightly greater among claimants aged 50-53 in 2004 than it was among those of the same age observed in

[14] The differences by age may reflect underlying differences by diagnosis. As table 1 shows, younger claimants tend to suffer more often from mental health disorders from which there is a higher rate of recovery than there is from some other conditions (Autor and Duggan 2006; Von Wachter et al. 2011). In addition, one expects greater deterioration in the skills and labor market attachment of older claimants who tend to have longer DI spells (Vingard et al. 2004; Bryngelson et al. 2009; Svensson et al. 2010).

1999. Since this age group was not targeted by the reform, the period difference in the trends cannot be attributed to anticipation effects.

A more plausible explanation is that labor market conditions in January-October 2004 differed somewhat from those in January-October 1999 for both older and younger claimants, such that the DI exit rate was slightly higher in the later period for both age groups. Our DADID strategy allows for such period effects. What we need to assume is that the period difference in the DI/employment trends would not have differed by age group in the absence of the reform. The bottom panel of Figure2 lends some support to the plausibility of this assumption. Each graph shows the absolute difference in the probability of claiming DI/being employed across the two age groups within a period. The smooth lines identify the age differential in the trends for the reform period sample, while the crossed lines are for the non-reform period.

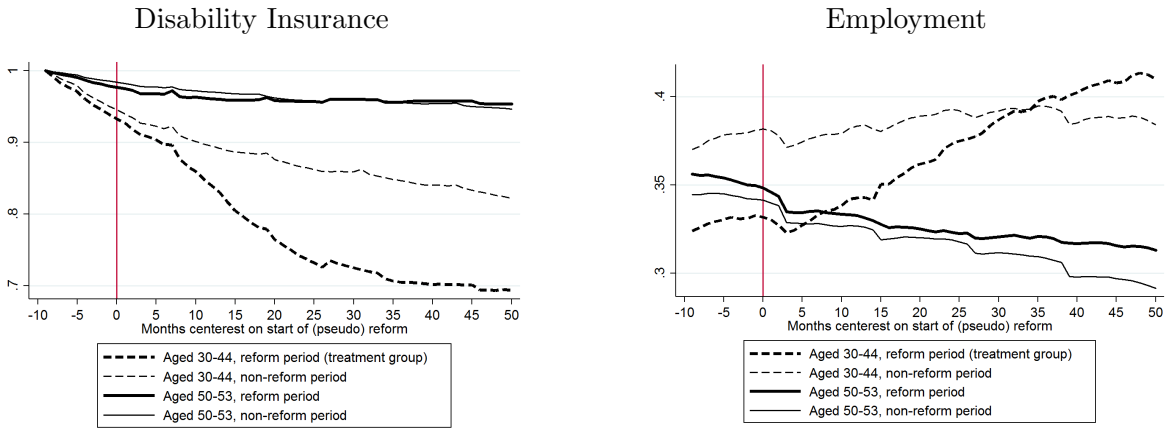
Until about month 5, by which point in the reform period only 8% of those in the treatment group had been reassessed, the age difference in the DI exit trends is quite consistent across periods. In both periods, younger claimants were exiting DI more rapidly. The rate of exit differed slightly across periods for both age groups. However, up to month 5 there is little or no evidence that the age differential in DI exit differed across periods. After that month, corresponding to March 2005 in the reform period, when reassessments in the reform period began to accumulate markedly (see figure 1), the differential between the exit rate of younger and older claimants in that period accelerated rapidly relative to the respective differential in the non-reform period. [15] If the age differential in the non-reform period traces what would have happened in the absence of reassessments, then the difference in these differentials identifies the impact of the reform on DI exit.

From the comparison of employment rates over two periods of only nine months (January-September 1999 and 2004), it is difficult to gauge the plausibility of the identification assumption that the age differential in employment trends would have been common across periods in the absence of the reform. To make a better assessment of whether this is a credible assumption, we examine two different cohorts of DI claimants over a longer period prior to the start of the reassessments. 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 in

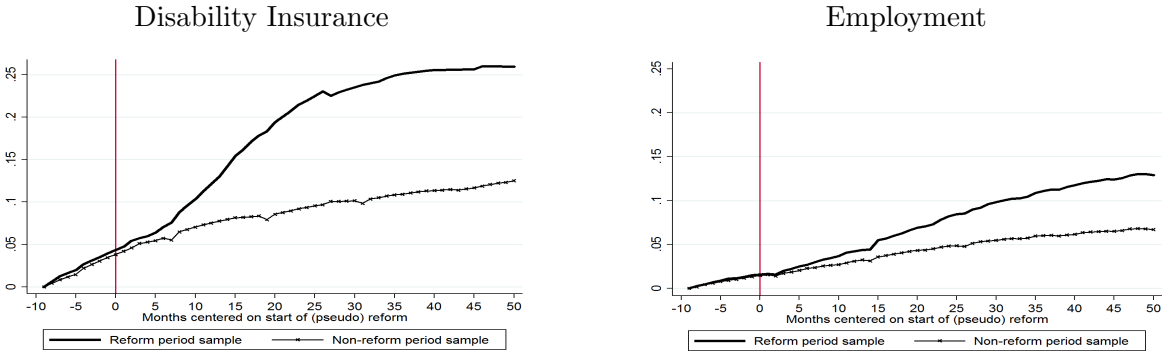
[15] See Appendix C Table C2 for the number of reassessments conducted per month.

Figure 2. : Trends in Disability Insurance dependency and employment by age group and period

Panel A: Proportion in state by age and period



Panel B: Age group difference in DI dependency and employment by period



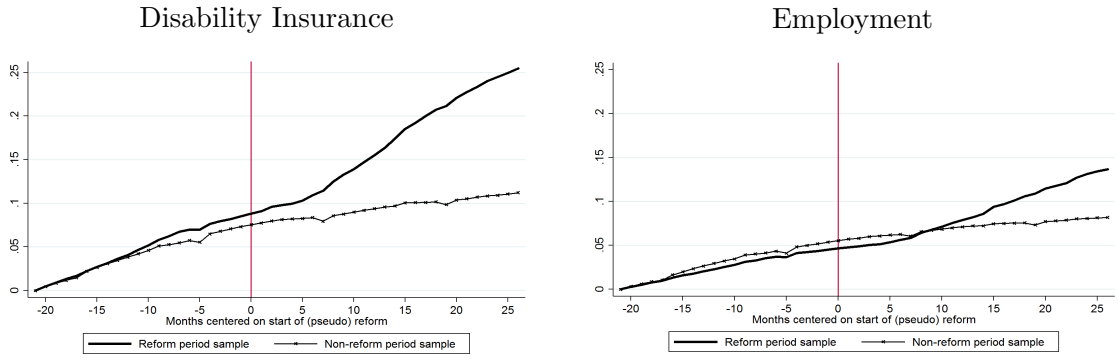
Note: Reform period (Jan. 2004-Dec. 2008) sample consists of individuals aged 30-44 & 50-53 on July 1, 2004 who were claiming DI in January 2004. Non-reform period (Jan. 1999-Dec. 2003) sample consists of individuals aged 30-44 &50-53 on July 1, 1999 who were claiming DI in January 1999. Month 0 is October 2004 for reform period and October 1999 for non-reform period. Top panel shows the proportion of each group in the state in each month. Bottom panel shows the absolute difference between the proportion of the younger group and the proportion of the older group in each state within each period-specific sample. In this panel, the employment trends have been linearly transformed to fix the employment rate of all groups in month -10 at the rate of the treatment group at that point. Hence, the age differential is set to zero for both periods at -10. Group sizes are as given in Table 1.

the data until December 2006. Those in the younger group of this cohort were subject to reassessment from October 2004, if they were still on DI at that time. They constitute an alternative treatment group to the one we use for estimation. But they are observed for 21 months prior to the start of the period of reassessments. 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 3 shows the age group differential in the trends in DI dependency and employment over the four years that these cohorts are followed. For both variables, and particularly for employment, there is little or no difference across the periods in the age differentials over the 21 months prior to the (pseudo) start of the reassessments. As was evident in Figure 2, a period difference in the age differentials in the trends opens up five months after the start of the reassessments. Previous to this, younger claimants were leaving DI and entering employment at faster rates than older claimants in both periods, but the age differences in these rates were roughly constant across periods. Shortly after the start of reassessments, the rates of employment (and DI exit) of younger and older claimants began diverge to a much greater extent in the later period in which the younger group was subjected to more stringent DI qualification criteria.

This figure both lends credence to our identification strategy and suggests that the reform had an important impact on employment. Over a period of almost two years prior to the start of reassessments, the age differential in employment trends was very similar to that observed over an earlier period in which there was no DI reform affecting existing claimants. After the start of the reassessments of the younger claimants, their employment accelerates relative to that of older claimants at a faster rate than is observed in the earlier period in which there were no reassessments.

Figure 3. : Age group difference in DI dependency and employment by period over 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. Figure shows the absolute difference between the proportion of the younger group and the proportion of the older group in each state within each of the period-specific samples. Employment trends have been linearly transformed to fix the employment rate of all groups in month -10 at the rate of the treatment group at that point. Hence, the age differential is set to zero for both periods at -10.

D. Estimation

We estimate the $ITTE$ effects by pooling the balanced panels from the reform and non-reform periods and estimating fixed effects models by least squares. When relying on the DID assumption (3), we use only observations of claimants aged 30-44 on either July 1, 2004 (reform period panel) or July 1, 1999 (non-reform period panel) and estimate the following regression,

$$y_{it} = \sum_{\substack{t=-9,\\ t\neq -1}}^{50} \beta_t P_i M_t + \sum_{\substack{t=-9,\\ t\neq -1}}^{50} \theta_t M_t + \mu_i + \varepsilon_{it} \tag{6}$$

where P_i is an indicator equal to 1 if the observation is from the 2004-2008 panel and is 0 otherwise, M_t is an indicator of the month (t) of the observation, such that $M_0 = 1 \& P_i = 1$ indicates October 2004 and $M_0 = 1 \& P_i = 0$ indicates October 1999, $M_{-9} = 1$ indicates January 2004 or January 1999, and $M_{50} = 1$ indicates December 2008 or December 2003. μ_i is an individual fixed effect and ε_i is an idiosyncratic error, which is mean zero under the identification assumption.

If (3) holds, then β_t corresponds to the $ITTE$ t months after the start of reassessments in October 2004 and this can be consistently estimated by OLS.

To utilize the DADID assumption (4), we add claimants aged 50-53 on either July 1, 2004 or July 1, 1999 to the DID estimation sample and estimate the following regression,

$$y_{it} = \sum_{\substack{t=-9,\\ t\neq -1}}^{50} \beta_t Z_i P_i M_t + \sum_{\substack{t=-9,\\ t\neq -1}}^{50} \theta_t M_t + \sum_{\substack{t=-9,\\ t\neq -1}}^{50} \gamma_t Z_i M_t + \sum_{\substack{t=-9,\\ t\neq -1}}^{50} \delta_t P_i M_t + \mu_i + \varepsilon_{it} \tag{7}$$

Now β_t corresponds to the $ITTE$ t months after the start of reassessments if (4) holds. We are particularly interested in estimates of β_{50} since this corresponds to the effect when effectively all reassessments had been completed in December 2008. After scaling the $ITTE$ by dividing by the proportion reassessed to recover the $ATET$ at each month, we can examine how this impact varies over the period of implementation. We expect the effect on employment to accumulate as UI and temporary income support benefits become exhausted some time after a claimant exits DI.

Income, earnings and benefit amounts are recorded on a yearly basis. We use annual data to estimate the effects on these outcomes. The DADID specification becomes:

$$y_{it} = \sum_{t=-1}^{3} \beta_t Z_i P_i A_t + \sum_{t=-1}^{3} \theta_t A_t + \sum_{t=-1}^{3} \gamma_t Z_i A_t + \sum_{t=-1}^{3} \delta_t P_i A_t + \mu_i + \varepsilon_{it}, \tag{8}$$

where the unit of time is now years and A_t is a year indicator. We are particularly interested in the estimate of β_3, which corresponds to the effect at year 2008. We scale the resulting $ITTE$ estimate by the proportion of reassessments in that year. The DID specification for these outcomes is defined analogously.

We cluster standard errors at the individual level. Since we observe the population proportion that was reassessed in each month, division of the $ITTE$ estimate by this scalar does not introduce any additional sampling variability.

In addition to estimating (6), (7) and their yearly counterparts using all DI claimants, we estimate separate regressions within groups defined by diagnosis and age in order to test the hypothesis that residual work capacity varies with these two characteristics.

V. Results

A. Magnitude of the reform

Before examining the impact of increased DI stringency on labor market activity, we assess the extent to which the reform cut entitlement. Table 3 gives estimates of the average effects of reassessment on: a) the probability of DI participation, b) the DI replacement rate, and c) the amount of DI benefits received. Each entry is an estimate of the ATET, which is derived from the respective ITTE at the end of the follow-up period scaled by the proportion of the target group reassessed by that time. Since this proportion is very close to 1, the scaling has very little effect. The effects on DI participation and the replacement rate are estimated from monthly data and are evaluated at December 2008. The effect on the amount of benefit is estimated from annual data and is the effect in 2008. The effects capture not only the direct impact on entitlement of applying stricter rules but also any behavioral response to the curtailment of benefits. For example, a claimant may leave DI because she was either no

longer eligible according to the new rules or induced by a cut in benefits to seek employment. If application of the new rules reduced the claimant's assessed degree of disability, then her replacement rate would fall correspondingly (see Appendix C C3). If the claimant was forced or induced to leave DI, then we set her replacement rate to zero.

The DID estimates of the effects on participation and the benefit amount are somewhat smaller than the DADID estimates, but the differences are not large. The choice of identification assumption is effectively inconsequential for the estimated effect on the replacement rate. The DADID estimate indicates that reassessment reduced the probability of being on DI by 13 percentage points. Applying stricter criteria removed a substantial fraction of claimants from the DI roll. Not all individuals in the target group would still have been claiming DI at the end of the follow-up period if the reform had not been implemented. In order to give an impression of the relative impact of the reform, we use the estimates of (7) to predict what the DI participation rate would have been if there had been no tightening of the rules. Specifically, we use the predicted DI status of an individual aged 30-44 observed in December 2008 less the treatment effect, i.e. $[\hat{y}_{it}|Z_i = 1, P_i = 1, M_t = 50] - \hat{\beta_{50}}$. This gives a DI participation rate of 56 percent, implying that reassessment with stricter criteria reduced the rate by 23 percent.

The DADID estimate in the second row indicates that reassessment reduced the replacement rate by 7.2 percentage points, on average, which is 15 percent of a predicted replacement rate of 47 percent under the counterfactual of no reform. Using DADID, reassessment is estimated to have reduced the annual amount of DI benefit received by around €1500, which is 20 percent of the average amount under the counterfactual.[16] as $[\hat{y}_{it}|Z_i = 1, P_i = 1, A_t = 3] - \hat{\beta}_3$

These estimates confirm that the reform produced a substantial cut in disability benefits. It appears to have been more aggressive than the earlier tightening of DI eligibility in the Netherlands examined by Borghans et al (2014), which they estimate reduced DI participation by 3.8 percentage points, cut the replacement rate by 11 percent in relative terms and decreased the money value of benefits by 10 percent for claimants just below the age of 45.

Figure 4, which is produced using DADID estimates, shows the evolution of the reform impact over the implementation period. In Panels A and B, it is evident that prior to the reform (¡month 0), the estimates are close to zero and do not display any trend. They are significantly different from zero, but this is due to the very large sample size. This indicates

[16]For this outcome, counterfactual mean is obtained from the estimates of (8)

Table 3—: Estimated effects on DI participation and benefits at end of follow-up

	DID	DADID	Predicted mean if no reform
Effect of reassessment on:			
DI participation (dummy)	-12.01***	-13.06***	83.35%
	(0.15)	(0.17)	
DI replacement rate (percentage points)	-7.30***	-7.20***	46.24%
	(0.09)	(0.12)	
DI benefit amount (€)	-1435.56***	-1507.98***	10,537
	(28.30)	(35.32)	
No. of individuals	301,050	499,949	156,493

Note: Table shows estimated ATET of reassessment obtained from least squares DID (equation (6) *or yearly equivalent) and DADID (equation* (7) *or* (8)*) estimates of ITTE at December 2008 (2008 for DI amount) scaled by the proportion of claimants reassessed by that date. First row gives impact on probability of being on disability insurance (DI) multiplied by 100 to give percentage point effect. Column on right gives predicted outcome for target group at end of follow-up period if there had been no reform, which is obtained from estimates of* (7) *or* (8) *less the DADID estimate of the treatment effect. Sample sizes correspond to the stated number of individuals × number of periods. Standard errors clustered at the individual level in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent levels.*

that before reassessments started, the age differential in the trend in DI participation (and in the replacement rate) differed very little between the two periods. For DI participation, this is precisely what was observed in Figure 2. The ITTE estimates remain close to zero and reasonably flat in the first few months of implementation, when very few reassessments were actually carried out. Thereafter, the magnitude of the estimated ITTE increases as the rate of reassessments picks up. By the end of 2007 (month 38), 83% of claimants had been reassessed and so there is little remaining difference between the ITTE and ATET. By the end of 2008, there is effectively full implementation and so the effects converge.

We do not show the ATET for DI participation and the replacement rate in the first months of implementation since the low reassessment rate in this period produces volatile estimates. This variability continues until around month 18. Thereafter, the effects increase somewhat in magnitude before steadily decreasing. This pattern could be due to scheduling younger

claimants for reassessment earlier in the implementation period. If, as might be expected, there was a smaller impact on older claimants, then the average effect would decrease as the average age in the pool of reassessed claimants increased. We examine age heterogeneity in the effect below. The fact that the ATET levels off toward the end of the period suggests that the estimates at the end of 2008 presented in Table 3 do capture the full average effects of reassessment.

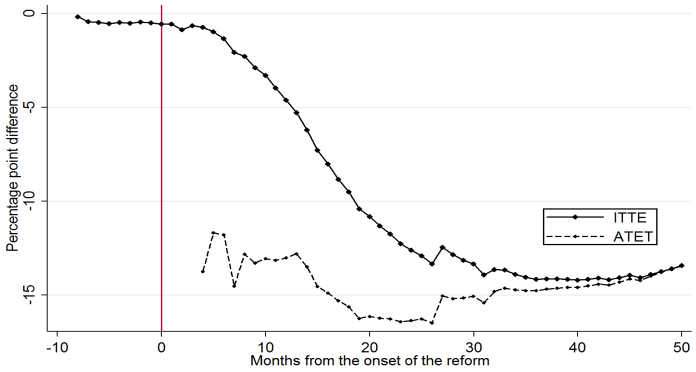
The estimated ATET on the DI benefit amount, which is obtained from annual data, rises sharply over the first two years of implementation as benefits are terminated or reduced following a downward revision in the degree of disability. After two years, the ATET flattens out at around €1500.

Table 4 reveals substantial variation in the impact of increased stringency on DI participation and benefits by diagnosis. Given the concentration of claimants at low initial degrees of disability was greatest for those with musculoskeletal problems, it is not surprising that tightening of the eligibility criteria reduced program participation most for this group. Reassessment reduced their probability of continuing to claim DI by 19 percentage points. It reduced their DI replacement rate by almost 10 points, on average, and it cut the amount of benefits they received by around €1800. The numbers in square brackets show the ATET as a fraction of the predicted outcome for the treatment group at the end of the follow-up period if there had been no tightening of DI eligibility. These reveal that the greater impact on claimants with musculoskeletal problems is even more evident in relative terms. The DI participation and replacement rate of this group were both reduced by about 23 percent of what each would have been if there had been no belt-tightening. The amount of DI benefit received was reduced by around 20 percent for claimants with a musculoskeletal diagnosis.

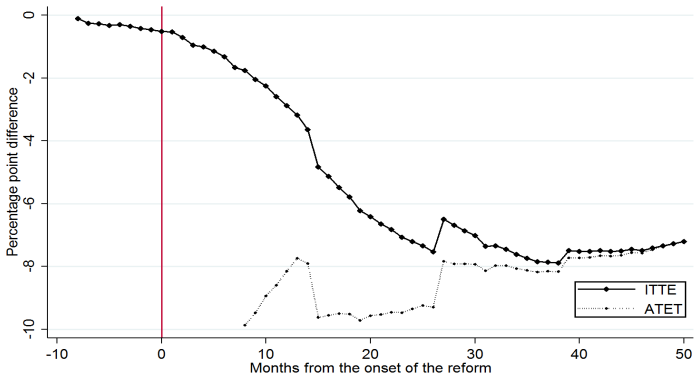
The least affected were claimants with cardiovascular and neurological problems. The impacts on DI participation, replacement rate and benefits of these groups are roughly half as large as those on claimants with musculoskeletal problems. This suggests that the tightening of the DI eligibility criteria did have a greater impact on claimants with conditions that are more subjective in the sense of having a less precise medical diagnosis. Consistent with this, there are also substantial effects on claimants who qualify through mental health problems. Despite claimants in this category initially being more likely than any other group to be assessed as fully disabled, their probability of continuing to claim DI was reduced by 15

Figure 4. : Estimated effects on DI participation and benefits over implementation period

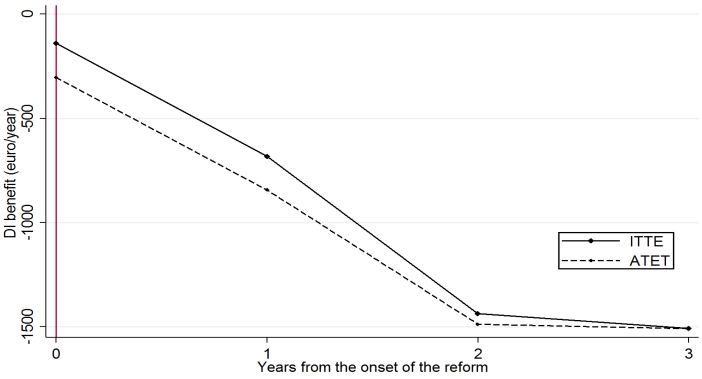
Panel A: DI participation



Panel B: Replacement rate



Panel C: DI amount



Note: The figure shows DADID estimates of the ITTE (equation (7)*), as well as ATET estimates obtained by scaling the ITTE by the proportion of claimants reassessed by each date. Point 0 on the horizontal axis corresponds to October 2004 in Panels A and B, and 2005 in Panel C. All estimates at every date are significantly different from zero at 1% or less.*

percentage points, or around 22% of what it would have been without reassessment under stricter rules. However, the effects on the replacement rate and the benefit amount for those with a mental health condition are not as large as those on groups with more objectively defined respiratory and digestive disorders.

Since these estimates capture not only the cut in benefit entitlement but also the response by opting to leave DI, we cannot conclude that benefit stringency was necessarily tightened for one group of claimants relative to another. However, the impact of the increased stringency on DI participation, if not benefit levels, does appear to have been greatest for the large group qualifying through musculoskeletal problems.

Table 4—: Estimated effects on DI participation and benefits at end of follow-up by diagnosis

	General	Cardiovascular	Musculoskeletal	Neurological	Mental	Respiratory	Digestive
Effect of reassessment on:							
DI participation (dummy)	-11.66***	-8.31***	-19.00***	-9.77***	-14.93***	-15.16***	-13.76***
	(0.49)	(1.08)	(0.35)	(0.62)	(0.28)	(1.39)	(1.19)
	[13.50]	[9.48]	[22.55]	[10.57]	[17.44]	[17.22]	[15.78]
Replacement rate	-5.72***	-4.37***	-9.93***	-5.64***	-8.24***	-9.70***	-7.97***
	(0.30)	(0.69)	(0.21)	(0.42)	(0.19)	(0.92)	(0.78)
	[12.29]	[8.89]	[23.95]	[10.20]	[16.30]	[19.41]	[16.88]
DI amount (€)	-1167.39***	-869.44***	-1805.74***	-690.95***	-1487.07***	-2092.01***	-1823.72***
	(103.64)	(226.96)	(63.16)	(159.09)	(57.64)	(283.24)	(241.95)
	[10.13]	[7.70]	[19.62]	[5.73]	[13.47]	[18.39]	[16.96]
No. of individuals	75,543	15,572	144,434	26,599	178,093	7,711	10,881

Note: Table shows estimated ATET of reassessment from diagnosis-specific least squares DADID (equation (7)*) estimates of ITTE at December 2008 (2008 for DI amount) scaled by the proportion of claimants per diagnosis reassessed by that date. Standard errors clustered at the individual level are given in parentheses. In square brackets is the ATET as a percentage of the predicted outcome for the treatment group under the counterfactual of no reform. Individuals are split by diagnosis recorded at entry to the sample. Notes to Table 3 apply.*

In Table 5 we disaggregate estimates of the effects on DI participation and benefits by age, as well as diagnosis. Two findings are noteworthy. First, the pattern of heterogeneity in the effects by diagnoses is consistent across age groups. Within each age group, the largest impact on DI participation is for claimants with musculoskeletal problems and the biggest effects on the replacement rate and benefit amount are for those with respiratory problems. The smallest effects are often on those with cardiovascular and neurological conditions. Second, the reform impacted most on younger claimants irrespective of diagnosis. The differences by age are substantial. For example, among claimants with musculoskeletal problems, the impact on the DI participation rate is 11.6 percentage points (81%) higher in the group aged 30-34 compared with those aged 40-44, and the effect on the benefit level is 53% larger in the younger group.

Most previous studies have not identified the impact of DI cuts on younger claimants. Borghans et al (2014) provide a rare evaluation of a reform that did not affect only older claimants. But even in this case the effect was identified around the age of 45. Moore (2015) also estimates the impact of terminating DI benefits on a younger sample of claimants, but only for the 2% of US DI claimants who qualified through addiction-related conditions. In the cohort of claimants affected by the reform we evaluate, 56% are below the age of 40. We reveal substantially larger impacts on younger claimants.

Table 5—: Estimated effects on DI participation and benefits at end of follow-up by diagnosis and age

	General	Cardiovascular	Musculoskeletal	Neurological	Mental	Respiratory	Digestive
Aged 30-34							
DI participation (dummy)	-16.86***	-11.21***	-25.93***	-14.18***	-17.85***	-21.53***	-16.23***
	(1.11)	(2.85)	(0.74)	(1.24)	(0.56)	(3.51)	(2.22)
	[21.2]	[13.22]	[33.76]	[15.86]	[22.37]	[26.29]	[19.57]
Replacement rate	-7.46***	-5.57***	-12.44***	-7.84***	-9.29***	-14.39***	-9.09***
	(0.66)	(1.81)	(0.43)	(0.81)	(0.37)	(2.22)	(1.47)
	[17.41]	[11.60]	[32.68]	[14.36]	[19.62]	[29.76]	[19.88]
DI amount (€)	-1,700.32***	-905.08**	-2,215.88***	-1,251.88***	-1,657.75***	-2,957.67***	-1,922.22***
	(181.50)	(470.98)	(103.49)	(223.43)	(86.88)	(517.05)	(355.20)
	[18.78]	[10.11]	[28.81]	[12.39]	[17.46]	[32.22]	[21.69]
No. of individuals	41,581	10,852	80,020	13,013	88,695	4,576	5,789
Aged 35-39							
DI participation (dummy)	-12.32***	-8.57***	-20.41***	-10.57***	-16.00***	-19.87***	-15.54***
	(0.80)	(2.06)	(0.56)	(0.98)	(0.45)	(2.43)	(1.84)
	[14.60]	[10.28]	[24.87]	[11.61]	[18.98]	[22.42]	[18.02]
Replacement rate	-5.94***	-5.09***	-10.65***	-6.18***	-8.70***	-12.47***	-9.23***
	(0.48)	(1.29)	(0.54)	(0.32)	(0.63)	(1.47)	(1.19)
	[12.93]	[10.72]	[25.98]	[11.27]	[17.48]	[25.28]	[19.53]
DI amount (€)	-1231.31***	-1076.70**	-1953.54***	-796.43***	-1614.73***	-2300.68***	-2059.35***
	(147.09)	(364.07)	(88.83)	(205.17)	(77.46)	(404.58)	(319.16)
	[12.31]	[10.87]	[22.99]	[7.04]	[14.88]	[21.11]	[20.75]
No. of individuals	46,830	11,620	89,714	14,879	100,350	5,108	6,393
Aged 40-44							
DI participation (dummy)	-9.67***	-7.44***	-14.29***	-6.94***	-12.05***	-10.19***	-10.72***
	(0.61)	(1.32)	(0.44)	(0.77)	(0.36)	(0.18)	(1.52)
	[10.88]	[8.39]	[16.60]	[7.40]	[14.29]	[11.50]	[12.03]
Replacement rate	-5.21***	-3.70***	-7.95***	-4.13***	-7.02***	-6.47***	-6.23***
	(0.37)	(1.29)	(0.32)	(0.63)	(0.29)	(1.47)	(1.19)
	[10.75]	[7.31]	[18.48]	[7.28]	[13.26]	[12.69]	[12.81]
DI amount (€)	-986.03***	-753.70***	-1451.79***	-327.46*	-1249.42***	-1680.93***	-1569.47***
	(126.36)	(274.67)	(78.70)	(190.09)	(70.98)	(347.27)	(300.92)
	[9.04]	[6.99]	[15.66]	[2.69]	[10.31]	[15.09]	[14.80]
No. of individuals	54,464	13,206	101,126	16,675	110,580	5,889	7,037

Note: Table shows estimated ATET of reassessment from diagnosis-age specific least squares DADID (equation (7)*) estimates of ITTE at December 2008 (2008 for DI amount) scaled by the proportion of claimants per diagnosis and age group reassessed by that date. Individuals are split by diagnosis recorded at entry to the sample and age at July 1 2004 (1999 for non-reform period). Sample size gives the number of individuals per diagnosis and age group in the first period. Notes to Table 3 apply.*

B. Work activity and earnings effects

Table 6 shows estimates of the ATET on the probability of employment in December 2008, and on the number of days worked and on earnings throughout 2008. The DADID estimate indicates that reassessment under stricter rules increased the probability of employment by just over six percentage points. Reliance on the DID assumption gives a somewhat larger effect. The DADID estimate is 18 percent of the prediction of what the employment rate of the treatment group would have been in December 2008 in the absence of the reform (34.6 percent). We also show estimates of the effect on the probability of employment scaled by the effect on DI participation. In this case, we confine attention to being employed without simultaneously claiming DI. The respective DADID estimate suggests that half of those induced to leave DI entered into employment.

Borghans et al (2014) estimate that the earlier tightening of the Dutch DI program reduced the probability of DI participation by 3.8 points and raised employment by 2.9 points, implying that around three quarters of those induced to leave DI entered work. Since the work capacity of claimants would be expected to decrease as the process of rationalizing DI proceeds, it is anticipated that the later reform would produce lower absorption of displaced claimants into employment. Moore (2015) estimates that 22 percent of those losing DI entitlement in the US gain employment. The lower rate of work capacity of DI claimants in the US likely reflects the lower prevalence of DI dependency in that country. It may also be attributable to the incentives those leaving the US DI program have to stay out of work in order to strengthen their case should they reapply. As explained above, in the Netherlands eligibility for DI can be acquired from a state of employment, and workers can keep partial DI benefits. A claimant who moves into employment does not necessarily jeopardize any reapplication for DI submitted at a later date.

If we scale the estimated effect on employment by the average effect on the amount received from DI, then we obtain an estimate of a 4.15 percentage point increase in employment per €1000 reduction in benefits. This is greater than the 2.9 points estimated by Borghans et al. The difference may be because the proportion of fully disabled claimants was higher at the time of the earlier reform. The benefits of many claimants were cut without being terminated. In the later reform, a higher proportion of claimants were close to the threshold

minimum degree of disability and the tightening of the criteria resulted in a larger number of terminations for a given average cut in benefit payments.

Using the DADID assumption, DI benefit cuts are estimated to have resulted in 47 more days of work per year, on average. This is a substantial impact, which is very large relative to the predicted degree of work activity in the absence of the reform. The increase in labor market activity causes earnings to rise. Relying on the DADID assumption, we estimate that reassessment induced DI claimants to raise their earnings by €685, on average. This is 11.9% of mean earnings under the counterfactual of no reform and 45% of the estimated average effect on the amount of DI benefit received. These estimates suggest that DI claimants subjected to reassessment had considerable earnings capacity they could draw on to replace a substantial part of the benefit income lost due to the greater program stringency.

The increase in earnings is entirely due to increased work effort. Estimates of the effect on wages - earnings per day worked - are negative. This suggests that the claimants who were induced to enter employment by the tightening of DI eligibility had lower productivity than those claimants who would have entered employment in any case without the added push from benefit cuts. The DADID estimate of the effect on this outcome is 11.2 percent of the predicted mean daily wage of the treatment group in 2008 if there had been no reform.

Figure 6 shows how the effects on employment and earnings vary over the implementation period. The ITTE again predictably increase over time as more claimants are reassessed. The ATET on employment and earnings both also increase over most of the period after the initial volatility cause by low reassessment rates in the first months. This is different from the pattern observed in Figure 4 for the ATET on DI participation. A potential explanation is that the impact on employment lags that on DI participation because those who leave DI exhaust entitlements to UI and the temporary income support program before moving into employment. Once these entitlements are exhausted, the employment effect grows. Towards the end of our follow-up period, the impact on employment is beginning to level off, while the earnings effect is still rising. The latter suggests that we may not be capturing the full effect on earnings.

Table 7 presents the estimated effects on work activity and earnings by diagnosis. As with DI participation, the largest absolute effects on employment and days worked are among

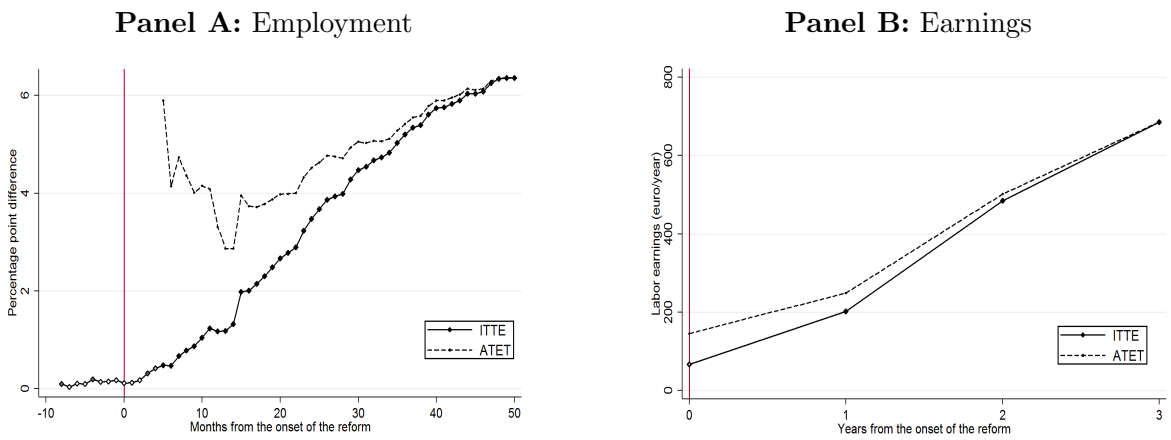
Table 6—: Estimated effects on Work activity and earnings

	DID	DADID	Predicted mean if no reform
Employment (dummy)	7.41***	6.26***	34.6%
	(0.16)	(0.22)	
Days worked	52.34***	46.89***	62.62
	(0.28)	(0.36)	
Earnings (€)	692.85***	685.48***	5,953
	(42.15)	(44.18)	
Earnings per day worked	-6.67***	-9.87***	87.59
	(1.31)	(2.24)	
Employment effect / DI participation effect	0.55	0.51	
Earnings effect / DI amount effect	0.48	0.45	
No. of individuals	301,050	499,949	156,493

Note: Table shows estimated ATET of reassessment obtained from least squares DID (equation (6)*) and DADID (equation* (7)*) estimates of ITTE at December 2008 (2008 for earnings and days worked) scaled by the proportion of claimants reassessed by that date. First row gives impact on probability of being employed presented in percentage points. Second row gives effect on days worked per year. Third row gives the effects on yearly earnings from employment and self-employment. Fourth row gives the effect on earnings per days worked a year. Fifth row gives the probability of being employed and not claiming DI divided by the effect on DI participation from Table 3. Last row gives the earnings effect divided by the effect on the DI amount from Table 3. Column on right gives predicted outcome for target group at end of follow-up period if there had been no reform, which is obtained from estimates of* (7) *or* (8) *less the DADID estimate of the treatment effect. Sample sizes are stated number of individuals × number of periods. Standard errors clustered at the individual level in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent levels.*

claimants who qualified through a musculoskeletal condition. For this group, reassessment increased the probability of employment by around 8 percentage points. The entry in square brackets indicates that this is 16.4 percent of a predicted employment rate of 50.2% for the treatment group in December 2008 under the counterfactual of no reform. The largest impacts on employment and days worked relative to the respective counterfactual rate is for claimants with mental health problems. In this group, reassessment is estimated to have increased employment by 18.8 percent and days worked by 96 percent relative to what each outcome would have been without increased stringency of the DI program. The large effects on those with musculoskeletal and mental health conditions combined with the high prevalence of these diagnoses means that these groups contributed by far the most to the overall

Figure 5. : Estimated effects on employment and earnings over implementation period



Note: The figure shows DADID estimates of the ITTE (equation (7)*), as well as ATET estimates obtained by scaling the ITTE by the proportion of claimants reassessed by each date. Outcomes are defined in the notes to Table 6. Point 0 on the horizontal axis corresponds to October 2004 in Panel A and 2005 in Panel B. Black markers indicate significance at the 1% level.*

impact on the rate of employment.

The ranking of the employment effects by diagnosis is different when each is scaled by the respective effect on DI participation. Evidently, the proportion of claimants induced or forced to leave DI who subsequently entered employment is highest among those with cardiovascular and neurological conditions. The tighter eligibility criteria were less likely to cause claimants with these conditions to leave DI. But those who exited the program did not have greater difficulty finding work than those with more subjectively defined conditions. The revised eligibility criteria do not appear to have been biased against claimants in a particular diagnostic category for whom it became relatively more difficult to find work after being forced or induced to leave DI. A caveat is that employment depends on both capacity and preferences for work. It may be that those exiting DI with more subjectively defined conditions were more capable of finding employment, but were also were less inclined to seek it.

The largest absolute effect on earnings is for claimants with respiratory diseases and mental health conditions. The effect on the latter group is also large relative to the counterfactual of no reform, at 19 percent. In comparison, stricter eligibility criteria are estimated to have

raised earnings for those with a musculoskeletal condition by 10.6 percent relative to the mean under the counterfactual. Although the reform increased the employment of claimants with musculoskeletal problems greatly, it had a larger impact on the earnings of those with mental health problems.

The estimated effect on earnings per day worked is negative for all diagnostic groups, and is largest in absolute terms for those with a respiration-related condition.

Table 7—: Estimated effects on work activity and earnings by diagnosis

	General	Cardiovascular	Musculoskeletal	Neurological	Mental	Respiratory	Digestive
Employment (dummy)	3.16***	2.80***	8.22***	5.90***	6.70***	6.54***	7.11***
	(0.60)	(1.44)	(0.44)	(0.99)	(0.38)	(1.90)	(1.59)
	[7.50]	[6.65]	[16.37]	[16.59]	[19.42]	[18.00]	[15.89]
Days worked	32.80***	30.14***	49.81***	26.60***	38.99***	38.50***	39.07***
	(1.02)	(2.12)	(0.66)	(1.25)	(0.54)	(2.79)	(2.34)
	[54.72]	[50.99]	[77.95]	[55.46]	[94.28]	[88.76]	[68.97]
Earnings	456.65***	379.38	796.45***	786.87***	809.63***	1076.58***	664.42**
	(118.69)	(292.39)	(82.52)	(204.93)	(70.56)	(351.05)	(336.08)
	[6.87]	[5.54]	[10.66]	[25.97]	[18.90]	[21.14]	[10.60]
Earnings per day worked	-9.55**	-5.98	-10.17***	-3.60	-5.53	-20.83	-10.54
	(4.63)	(8.60)	(3.90)	(4.30)	(3.88)	(13.13)	(12.99)
	[10.79]	[6.52]	[11.13]	[7.35]	[7.51]	[30.57]	[15.43]
Employment effect / DI participation effect	0.55	0.64	0.56	0.59	0.51	0.54	0.51
Earnings effect / DI amount effect	0.39	0.44	0.44	1.14	0.54	0.51	0.36
No. of individuals	75,543	15,572	144,434	26,599	178,093	7,711	10,881

Note: Table shows effects given in Table 6 allowing for heterogeneity by diagnosis. They are obtained from diagnosis-specific least squares DADID (equation (7)*) estimates of ITTE at December 2008 (2008 for earnings) scaled by the proportion of claimants per diagnosis reassessed by that date. Entries in square brackets give the ATET as a percentage of the predicted outcome of the treatment group under the counterfactual of no reform. Individuals are split by diagnosis recorded at entry to the sample. Sample size per diagnosis is stated number of individuals × number of periods. Notes to Table 6 apply.*

Estimated effects on employment and earnings within age and diagnostic groups are given in Appendix D, Table D1. For all diagnoses, the effects are substantially larger for younger claimants. But as for the DI outcomes, the pattern of heterogeneity by diagnosis does not vary much with age. One exception is that among the youngest claimants, the employment and earnings effects are large for those with neurological disorders.

C. Spillover to other social assistance programs and the net effect on income

We estimate that claimants subjected to stricter eligibility rules were able to recover 45% of the average loss of DI benefits by increasing their labor market earnings. We now examine the extent to which the remaining deficit was filled by increased reliance on other social assistance programs. Table 8 gives estimates of the ATET on income from all social assistance programs other than DI, as well as the effect on the total individual income in 2008. Estimates are smaller using DID as opposed to the DADID strategy. Using the latter, the estimate of a €398.5 increase in income received from other social assistance corresponds to more than a quarter of the average loss of DI income. Borghans et al (2014) estimate that 30% of the DI benefit lost from the earlier tightening of eligibility was replaced by income from other social assistance. Opportunities to substitute between programs decreased only marginally in the decade between the reforms. Almost half (48.6%) of the spillover to other social assistance is through increased reliance on UI (see Appendix D, Table D2). A little less than a quarter (23.1%) is through increased claims of sickness pay and the remainder (28.3%) is due to greater dependence on means-tested welfare, as well as smaller programs. [17]

Using DADID, we estimate that reassessment resulted in a fall of €556 in the total income of a claimant, on average. This is 3.4% of the predicted 2008 mean income of the affected cohort under the counterfactual of no reform. While this may seem a rather modest loss, the average impact is obviously not indicative of the much large income reductions experienced by the 13% of DI claimants we estimate were forced or induced to leave DI. The average loss of total income is 37% (DADID estimate) of the average reduction in DI payments. In aggregate, 63% of the income lost from DI was replaced by increased income from all other

[17] We do not have information on payments received from the temporary income replacement program put in place to smooth the impact of cuts in DI income and so we underestimate the extent to which other social assistance substitutes for lost DI income. This temporary income support is included in the measure of total income. Table D3 in Appendix D shows how the estimated 13 percentage point reduction in the probability of claiming DI is accounted for by transitions to mutually exclusive labor market and social assistance states.

sources.

Table 8—: Estimated effects on other social assistance and total income

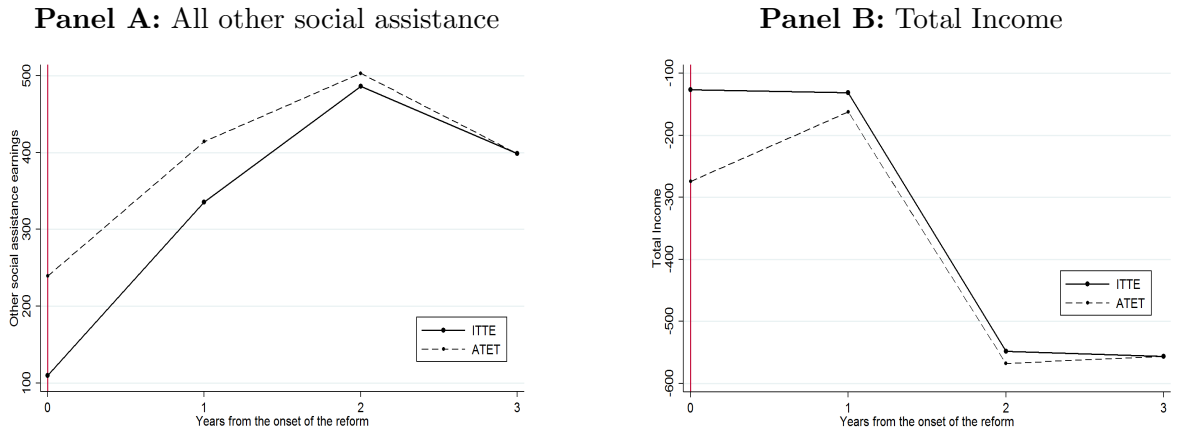
	DID	DADID	Predicted mean if no reform
All other social assistance income	339.98***	398.54***	875.32
	(13.57)	(20.73)	
Social assistance effect /DI amount effect	0.24	0.26	
Total income	-446.20***	-556.45***	16,096
	(42.46)	(44.24)	
Total income effect /DI amount effect	0.31	0.37	
No. of individuals	301,050	499.949	156,493

Note: Table shows estimated ATET of reassessment obtained from least squares DID (equation (6)*) and DADID (equation* (7)*) estimates of ITTE at 2008 scaled by the proportion of claimants reassessed by that date. First row gives effect on income from UI, welfare, sick pay and all other social assistance programs. All incomes are measured at the individual, not household, level. Column on right gives predicted incomes for target group in 2008 if there had been no reform. They are obtained from estimates of* (8) *less the DADID estimate of the treatment effect. Sample sizes are stated number of individuals × number of periods. Standard errors clustered at the individual level in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent levels.*

Figure 6 reveals that the spillover effect on transfers from other social assistance programs increased over the first three years of implementation (2005-2007), before falling back in the fourth year. This pattern presumably reflects the accumulating exit from DI and its substitution with other transfers followed by exhaustion of entitlement to UI. The ITTE on total income is rather muted in the first year of implementation, reflecting both the limited number of reassessments in this period and the availability of alternative social assistance to substitute for DI. The steep increase in the magnitude of the total income loss between years 1 and 2 indicates that rising compensation from other social assistance and increases in earnings over this period were insufficient to offset the steep loss in DI. The flattening out of the total income effect between years 2 and 3 reflects the same pattern in effect on DI income (Figure 4).

From Table 9 it is apparent that the spillover effect on other social assistance is largest in both absolute and relative terms for claimants with mental health conditions. This may be

Figure 6. : Estimated effects on other social assistance and total income over implementation period



Note: The figure shows DADID estimates of the ITTE (equation (7)*), as well as ATET estimates obtained by scaling the ITTE by the proportion of claimants reassessed by each date. Outcomes are defined in the notes to Table 8. Point 0 on the horizontal axis corresponds to 2005. Black markers indicate significance at the 1% level.*

attributable to the low labor force attachment of this group. At baseline, it had the lowest rate of employment, mean earnings and mean income (see Table 2. Inconsistent with this, this group was the second most successful in replacing lost DI income with earnings (see Table 7). As a consequence of its gains in earnings and other social assistance, the average loss of total income relative to the average loss of DI income is not particularly high for the group of claimants with mental health conditions. Claimants that qualified through musculoskeletal problems are able replace a much lower fraction of their large average losses in DI benefits. For this group, the increased payments received from other social programs is only 13 percent of the loss in DI transfers. Its average loss of total income is 44 percent of its average loss of DI income. The net income loss relative to the reduction in DI income is even greater for claimants with cardiovaslcular disease. For this group, only 34 percent of the loss of DI income is replaced by income from all other sources. This due to low replacement of DI income with income from both employment and other transfer programs. The former may be due to limited work capacity of this group. The lack of substitution with other social assistance may be due to the initial relatively high income of this group (see Table 2).

Table 9—: Estimated effects on other social assistance and total income by diagnosis

	ATET						
	General	Cardiovascular	Musculoskeletal	Neurological	Mental	Respiratory	Digestive
Other social assistance income	353.34***	206.91**	232.72***	176.77***	549.69***	367.65***	477.41***
	(46.03)	(86.10)	(29.43)	(57.60)	(27.78)	(126.47)	(94.22)
	[48.12]	[23.79]	[20.76]	[25.44]	[68.43]	[40.06]	[69.71]
Social assistance effect /DI amount effect	0.30	0.23	0.13	0.26	0.37	0.18	0.26
Income	-509.41***	-569.75***	-769.49***	43.21	-391.90***	-294.81	-492.83**
	(98.81)	(193.67)	(56.13)	(127.01)	(48.52)	(252.68)	(245.33)
	[2.84]	[3.17]	[4.81]	[0.24]	[2.55]	[1.82]	[3.03]
Total income effect /DI amount effect	0.44	0.66	0.43	0	0.26	0.14	0.27
Individuals (Cell)	75,543	15,572	144,434	26,599	178,093	7,711	10,881

Note: Table shows estimated ATET of reassessment obtained from least squares DADID (equation (7)*) estimates of ITTE at 2008 scaled by the proportion of claimants reassessed by that date. First panel gives effect on income from UI, welfare, sick pay and smaller social assistance programs. Second panel gives impact on individual income from all sources. Sample size corresponds to stated number of individuals × number of periods. Standard errors clustered at the individual level in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent levels.*

VI. Conclusion

Identification of policies that can relieve the fiscal strain of disability insurance requires evidence not only on the capacity of claimants to work but also on the extent to which earnings potential can be realized through curtailment of entitlements. This paper is one of the few to deliver evidence on the employment and earnings responses to cuts in the benefits of existing DI claimants. More stringent assessment of the degree of disability of Dutch claimants that reduced benefits by 14.3%, on average, and caused a 13 point decrease program participation produced a 6.3 point increase in the employment rate of claimants. This is a 15.7% increase relative to what employment would have been without the benefit cut. Earnings of claimants increased by 11.5%, on average. This is insufficient to fully replace the cut in DI benefits. Claimants partially made up the difference by increasing dependence on other social assistance programs. This program spillover effect, which amounts to 26% of the reduction in DI benefits, substantially reduces the public expenditure savings from reining in DI.

Claimants with musculoskeletal disorders were affected most by using stricter criteria to assess disability. Almost one fifth of claimants in this diagnostic category were forced or induced to leave DI. Their employment rate increased by 8.2 percentage points. This category of claimants appears to have the greatest work capacity. Those who qualified through mental health problems were also strongly affected, but not more so than claimants with some more easily verifiable medical conditions. They were reasonably successful in replacing lost DI income with earnings, but were also the group that turned most to other social assistance programs when DI was cut. Although claimants with musculoskeletal conditions had the strongest employment response, they were able to replace only 44 percent of lost DI income with earnings and they substituted least with other social assistance. The muted earnings effect in this group relative to the employment response is due to a 10% decrease in the daily wage. It appears that claimants with musculoskeletal problems were forced to leave DI to take up low quality employment.

References

Autor, David H. (2015). The Unsustainable Rise of the Disability Rolls in the United States: Causes, Consequences, and Policy Options in John Karl Scholz, Hyunpyo Moon, and Sang-Hyop Lee (eds.) Social Policies in an Age of Austerity: A Comparative Analysis of the US and Korea, pp. 107-136. Northampton, MA: Edward Elgar Publishing.

Autor, David H. and Duggan Mark G. (2003). The Rise in the Disability Rolls and the Decline in Unemployment Quarterly Journal of Economics, 118(1), February 2003, 157 - 206.

Autor, David H. and Duggan Mark G.(2006). "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding." Journal of Economic Perspectives, 20(3): 71-96.

Bell, B., Blundell, R., & 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.

Bloom, H.S. (1984). Accounting for No-Shows in Experimental Evaluation Designs. Evaluation Review, 8(2): 225–46.

Blundell, R. & Costa-Dias M. (2009). Alternative Approaches to Evaluation in Empirical Microeconomics. Journal of Human Resources, 44(3): 565-640.

Borghans, L., Gielen A., & Luttmer E. (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". (No. w2816). National Bureau of Economic Research.

Burkhauser, R. V., and Daly, M. C. (2011). Testing educational tools to demonstrate returns to work for children aging out of the SSI-disabled children program.

Bryngelson, A. (2009). "Long-term sickness absence and social exclusion." Scand J Public Health 2009, 37(8), 839-845.

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

Charles, K. K. (2003). The longitudinal structure of earnings losses among work-limited disabled workers. Journal of Human Resources, 38(3), 618-646.

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.

de Jong, P., Lindeboom, M. and van der Klaauw, B. (2011). Screening disability insurance applications. Journal of the European Economic Association, 9, 106–129.

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. The Review of Economic Studies, 72(2), 395-427.

French, E., & Song, J. (2014). The effect of disability insurance receipt on labor supply. American Economic Journal: Economic Policy, 6(2), 291-337.

Duggan, M., & Imberman, S. A. (2009). Why are the disability rolls skyrocketing? The contribution of population characteristics, economic conditions, and program generosity. In Health at older ages: The causes and consequences of declining disability among the elderly (pp. 337-379). University of Chicago Press.

Duggan, M. (2015). Senate testimony. http://www.budget.senate.gov/democratic/public/_cache/files/3ceeaf1b-9427-48c2-a51b-e422c36eb802/mark-duggan-testimony.pdf

Garcia-Gomez, P. & Gielen, A. C. (2014). Health Effects of Containing Moral Hazard: Evidence from Disability Insurance Reform, Tinbergen Institute Discussion Papers 14-102

Hassink W., Koning P., Zwinkels W. (2015). Employers opting out of public disability

insurance: selection or incentive effects?. IZA discussion paper, IZA DP No. 9181.

Imbens, G.W. & Rubin, D.B. (1997). Bayesian inference for causal effects in randomized experiments with noncompliance. The Annals of Statistics, 25(1), 305-327.

Kapteyn, A., Smith, J.P., and Van Soest, A. (2007). Vignettes and Self Reports of Work Disability in the United States and the Netherlands. American Economic Review, 97(1), 461-473.

Karlstrom A., Marten P., Svensson I. (2008). The employment effect of stricter rules for eligibility for DI: evidence from a natural experiment in Sweden. Journal of Public Economics, April 2008, pages 2071-2082.

Koning, Pierre, 2009, Experience Rating and the Inflow into Disability Insurance, De Economist, 157(3), 315-335.

Koning, Pierre, 2012, Anticipated and Unanticipated Incentives of Disability Insurance Experience Rating: The Case of The Netherlands, Policy and Practice in Health and Safety, 10(2), 63-76..

Koning, P., & Lindeboom, M. (2015). The Rise and Fall of Disability Insurance Enrollment in the Netherlands. The Journal of Economic Perspectives, 29(2), 151-172.

Kostol, A. R., & Mogstad, M. (2014). How financial incentives induce disability insurance recipients to return to work. American Economic Review, 104(2), 624-55.

Lahiri, K., Song, J., & Wixon, B. (2008). A model of Social Security Disability Insurance using matched SIPP/Administrative data. Journal of Econometrics, 145(1), 4-20.

Li, J., Hurley, J., DeCicca, P. and Buckley, G. (2014), Physician Response to Pay-for-performance: Evidence from a Natural Experiment. Health Econ., 23: 962–978. doi: 10.1002/hec.2971

Maestas, N., Mullen, K. J., & Strand, A. (2013). Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt. The American Economic Review, 103(5), 1797-1829.

Milligan, K. S., & Wise, D. A. (2011). Social Security and Retirement around the World: Historical Trends in Mortality and Health, Employment, and Disability Insurance Participation and Reforms-Introduction and Summary (No. w16719). National Bureau of Economic Research.

Moore, T. (2015), The employment effects of terminating disability benefits, Journal of Public Economics, Volume 124, April 2015, Pages 30-43

Mora Villarrubia, R., & Reggio, I. (2012). Treatment effect identification using alternative parallel assumptions. Universidad Carlos III, Departamento de Economia.

Newey, W.K. (2007). Course materials for 14.386 New Econometric Methods, Spring 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on 21 June, 2016.

O'Donnell, O., van Doorslaer, E., & Van Ourti, T. (2013). Health and Inequality, Netspar Discussion Paper No. 10/2013-060.

Staubli S. (2011). The impact of stricter criteria for disability insurance on labor force participation. Journal of Public Economics, Volume 95, May 2011, 1233-1235.

OECD. (2010). Sickness, Disability and Work: Breaking the Barriers: A Synthesis of Findings across OECD Countries, OECD Publishing, Paris.

OECD. (2012), Sick on the Job?: Myths and Realities about Mental Health and Work, Mental Health and Work, OECD Publishing, Paris.

OECD. (2015), Fit Mind, Fit Job: From Evidence to Practice in Mental Health and Work, Mental Health and Work, OECD Publishing, Paris.

Svensson T, Müssener U, 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

Staubli, S. (2011). The impact of stricter criteria for disability insurance on labor force participation. Journal of Public Economics, 95(9), 1223-1235.

UWV (Uitvoeringsinsituut Werknemersverzekeringen), 2008. "Staatsblad 2007", Amsterdam, The Netherlands.

UWV (Uitvoeringsinsituut Werknemersverzekeringen), 2006. "UWV, 4e kwartaal 2005", Amsterdam, The Netherlands.

UWV (Uitvoeringsinsituut Werknemersverzekeringen), 2009. "Rapportage afronding eenmalige herbeoordelingsoperatie", Amsterdam, The Netherlands.

Van Sonsbeek, J.M. & Gradus, R.H.J.M. (2013). Estimating the Effects of Recent Disability Reforms in the Netherlands. Oxford Economic Papers 65(4), 832–55.

Vingard E, Alexanderson K, Norlund A. (2009). "Consequences of being on sick leave." Scand J Public Health, 32 (63 suppl), 207-215.

Von Hippel, Paul T. (2005). "Mean, Median, and Skew: Correcting a Textbook Rule". Journal of Statistics Education 13

Von Wachter, T., Song, J., & Manchester, J. (2011). Trends in employment and earnings of allowed and rejected applicants to the social security disability insurance program. The American Economic Review, 3308-3329.

Appendix A: Identification of average treatment effect on the treated

It is well known that with random assignment to treatment, non-compliance among those assigned to treatment and the absence of always-takers, the average treatment effect on the treated is the intention-to-treat effect divided by the probability of treatment given assignment to treatment (Bloom 1984). Equivalently, with one-sided non-compliance, using random assignment as an instrument, the IV estimate given by the ratio of the reduced form to the first stage effects identifies the ATET (Imbens and Rubin 1997). We merely demonstrate that without random assignment to treatment but with one-sided non-compliance, the ratio of the average effects of assignment on the outcome and on treatment identifies the ATET. Following Newey (2007), use iterative expectations to write

$$\mathbb{E}[D_i\beta_i \mid Z_i = 1\ ] = \mathbb{E}[D_i\,\mathbb{E}[\beta_i \mid D_i, Z_i = 1] \mid Z_i = 1\ ].$$

Using

$$D_i\,\mathbb{E}[\beta_i \mid D_i\ ] = \left\{ \begin{array}{r} 0, D_i = 0 \\ \mathbb{E}[\beta_i \mid D_i = 1\ ], D_i = 1 \end{array} \right\} = D_i\,\mathbb{E}[\beta_i \mid D_i = 1\ ],$$

we can write,

$$\mathbb{E}[D_i\beta_i \mid Z_i = 1\ ] = \mathbb{E}[D_i\,\mathbb{E}[\beta_i \mid D_i = 1, Z_i = 1\ ] \mid Z_i = 1\ ].$$

Since $D_i = 1 \Rightarrow Z_i = 1$,

$$\begin{aligned} \mathbb{E}[D_i\beta_i \mid Z_i = 1\ ] &= \mathbb{E}[D_i\,\mathbb{E}[\beta_i \mid D_i = 1\ ] \mid Z_i = 1\ ] \\ &= \mathbb{E}[\beta_i \mid D_i = 1\ ]\,\mathbb{E}[D_i \mid Z_i = 1\ ]. \\ &= ATET \times P(D_i = 1 \mid Z_i = 1) \end{aligned}$$

Therefore,

$$\begin{aligned} ATET &= \frac{\mathbb{E}[\ D_i(Y_{i1}^1 - Y_{i1}^0) \mid Z_i = 1\ ]}{P(D_i = 1 \mid Z_i = 1)} \\ &= \frac{AEEE}{P(D_i = 1 \mid Z_i = 1)} \end{aligned}$$

Appendix B: Health conditions included in diagnostic group labeled 'general'

- Fever
- Pain
- Malaise and fatigue
- Other symptoms, no elsewhere classified
- Kaposi's sarcoma
- Other neoplasms
- Chromosome Abnormalities appointed in number and shape
- Congenital abnormality, not elsewhere classified
- Tuberculosis
- Status HN-positive
- Disease by HIV
- Disease Pfeiffer
- Tetanus
- Brucellosis
- Other zoonoses
- Late effects of certain infectious diseases
- Other specified infectious diseases
- Freezing
- Sarcoidosis [Disease Besnier Boeck]
- Other diseases
- Systemic disease of the connective tissue
- Metabolism of lipoproteins
- Other metabolic disorders
- Nutritional Deficiency
- Obesity [adiposity / obesity]
- Other forms of overfeeding
- Acute poisoning
- Chronic poisoning
- Complication of surgical and medical treatment
- Toxic inhalation fever
- Radiation Sickness
- Because of heat and light
- Effect of reduced temperature
- Caisson Disease
- Other effects of air pressure and water pressure
- Asphyxia
- Allergic reaction
- Other undesirable consequences of external causes
- Sickness unspecified

Appendix C: Additional descriptives and counterfactual outcomes

Table C1—: Temporary income benefits (TRI) granted

	2005	2006	2007	2008	2009
Decreased or terminated benefits	33,269	26,376	15,809	7,363	2,350
TRI requests	6,280	8.581	10.443	2.137	554
TRI granted (%)	69%	70.6%	77.3%	63.5%	70.4%

Source: UWV (2007, 2008, 2009)

Table C2—: Number of reassessments conducted per month for the age group 30-44

	Jan.	Febr.	March	April	May	June	July	August	Sept	Oct.	Nov.	Dec.
2004										3	243	1362
2005	2257	3336	3994	4445	3753	5100	5168	4962	6768	7222	7960	6647
2006	5636	5029	5389	4306	4481	4102	3597	3412	3408	3188	3310	2215
2007	2550	2340	2953	2494	2548	2377	1917	1342	972	991	630	366
2008	565	424	325	379	251	304	308	266	365	569	518	281
2009	31	14	28	16	2							
TOTAL 137,419												

Note: This table shows the number of reassessments per month for the group 30-44. Source UWV.

Table C3—: DI replacement rate assigned per degree of disability

Level of Disability	Benefit as a % of the pre-sickness Daily Wage
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: This table shows the degree of disability and corresponding replacement rate of the Dutch DI scheme prior to 2006. Source UWV.

Appendix D: Additional results

Table D1—: Estimates of the impact of the reform on the employment and earnings by diagnosis and age group

	ATET						
	General	Cardiovascular	Musculoskeletal	Neurological	Mental	Respiratory	Digestive
Aged 30-34							
Employment (dummy)	5.43***	7.00*	11.27***	11.05***	9.26***	10.35***	6.04***
	(1.20)	(3.72)	(0.84)	(1.70)	(0.66)	(4.03)	(2.80)
Employment effect/ DI participation effect	0.55	0.83	0.53	0.62	0.48	0.46	0.42
Earnings (€)	661.79***	1425.58**	1374.42***	1654.15***	946.22***	2006.31***	1042.07**
	(215.63)	(642.34)	(153.50)	(297.92)	(106.48)	(756.97)	(521.76)
Earning effect /DI amount effect	0.38	1.57	0.62	1.32	0.57	0.68	0.54
Earnings effect/ Worked days effect	8.92	8.29	10.95	18.55	14.84	17.85	15.47
Sample size	41,581	10,852	80,020	13,013	88,695	4,576	5,789
Aged 35-39							
Employment (dummy)	3.73***	3.71	9.38***	5.16***	8.01***	9.00***	11.14***
	(0.90)	(2.42)	(0.66)	(1.36)	(0.54)	(2.93)	(2.25)
Employment effect /DI participation effect	0.56	0.68	0.57	0.58	0.53	0.53	0.53
Earnings (€)	701.52***	674.91	842.42***	735.46***	946.21***	1504.33***	1356.14***
	(168.33)	(473.75)	(123.43)	(282.00)	(93.91)	(515.64)	(424.01)
Earning effect /DI amount effect	0.57***	0.63	0.43	0.92	0.60	0.65	0.66
Earnings effect/ Worked days effect	11.51	10.38	13.66	25.64	17.84	23.90	14.90
Sample size	46,830	11,620	89,714	14,879	100,350	5,108	6,393
Aged 40-44							
Employment (dummy)	2.03***	1.29	5.82***	3.81***	4.07***	3.80	4.52**
	(0.72)	(1.37)	(0.54)	(1.18)	(0.46)	(2.32)	(1.99)
Employment effect /DI participation effect	0.55	0.55	0.56	0.57	0.52	0.62	0.56
Earnings (€)	222.30	-34.16	480.07***	375.93	593.21***	513.85	-75.94
	(143.81)	(350.89)	(102.45)	(241.34)	(86.32)	(415.47)	(408.87)
Earning effect /DI amount effect	0.22	0.04	0.33	1.15	0.47	0.31	0.05
Earnings effect/ Worked days effect	18.07	15.09	20.38	42.90	26.53	33.95	21.55
Sample size	54,464	13,206	101,126	16,675	110,580	5,889	7,037

Note: Table shows estimated ATET of reassessment obtained from least squares DADID (equation (7)*) estimates of ITTE at December 2008 (2008 for earnings) scaled by the proportion of claimants reassessed by that date. First row gives impact on probability of being employed. The estimates are multiplied by 100 to give effects in percentage points. Second row gives effect on probability of being employed and not claiming DI divided by the effect on DI participation from Table 3. Third row gives the effect on earnings from (self-) employment. Fourth row gives the earnings effect divided by the effect on the DI amount from Table 3. Sample sizes correspond to the number of individuals (not observations=individuals × period). Standard errors clustered at the individual level in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent levels.*

Table D2—: Estimates of the impact of the reform on income and other social assistance income

	Other social assistance income	
	DID	DADID
Unemployment benefits	176.45***	182.22***
	(10.08)	(13.25)
Welfare benefits	98.96***	106.26***
	(9.83)	(14.28)
Sick Pay	83.58***	86.42***
	(6.09)	(6.85)
No. of individuals	301,050	499.949

Note: This table shows the estimates of the DADID specified in equation (7) *for the average treated DI claimant. Significance stars indicate significance of the estimate at the 1 percent level ***, 5 percent level **, 10 percent level *. Standard errors are in parentheses, clustered at the individual level. The observations correspond to the number of individuals in the first period observed.*

Table D3—: Estimates of the impact of the reform on other social assistance

	ATET	
	DID	DADID
Employment only (dummy)	7.89***	7.09***
	(0.12)	(0.14)
Other social assistance (dummy)	2.80***	2.75***
	(0.06)	(0.07)
Unemployment (dummy)	0.84***	0.94***
	(0.05)	(0.05)
Welfare (dummy)	1.08***	1.04***
	(0.04)	(0.05)
Sick Pay (dummy)	0.42***	0.42***
	(0.03)	(0.03)
Others (dummy)	2.42***	2.76***
	(0.07)	(0.08)
No. of individuals	301,050	499,949

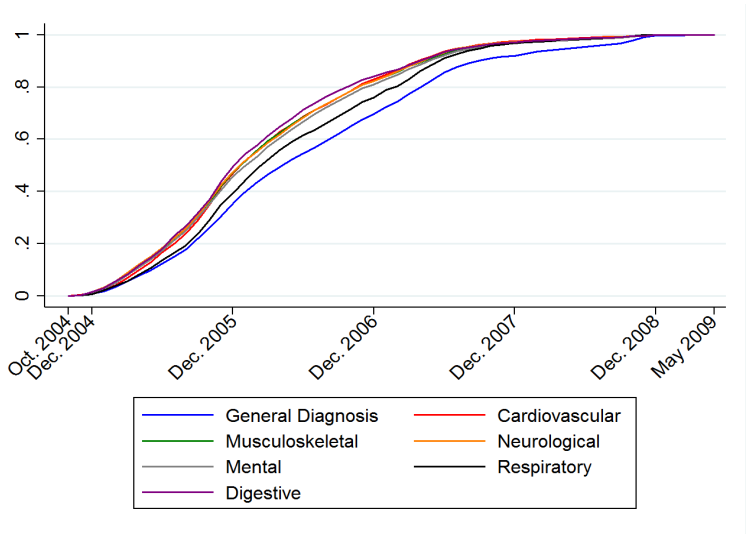
Note: Table shows estimated ATET of reassessment obtained from least squares DID (equation (6)*) and DADID (equation* (7)*) estimates of ITTE at December 2008 scaled by the proportion of claimants reassessed by that date. First row gives impact on probability of being employed and not claiming DI. Second row gives impact on probability of receiving unemployment benefits and not claiming DI or being employed. Third row gives effect on probability of receiving welfare benefits and not claiming DI, UI or being employed. The fourth row estimates represents the probability of claiming sick pay and not claiming DI, UI, welfare or being in employment. The last row estimates the probability of being in none of the other categories. The estimates are multiplied by 100 to give effects in percentage points. Sample sizes correspond to the number of individuals (not observations=individuals × period). Standard errors clustered at the individual level in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent levels.*

Table D4—: Estimates of the impact of the reform on the income and transfer income by diagnosis and age

	ATET						
	General	Cardiovascular	Musculoskeletal	Neurological	Mental	Respiratory	Digestive
Aged 30-34							
Income	-886.74***	961.65*	-552.28***	527.56**	-340.91***	-637.32	-381.28
	(206.29)	(558.98)	(139.72)	(254.72)	(94.90)	(651.32)	(461.58)
Unemployment benefits	-9.78	25.16	103.69**	71.94	104.27***	77.02	188.70*
	(59.13)	(130.39)	(45.94)	(68.56)	(30.76)	(221.71)	(114.56)
Welfare benefits	93.61	476.23**	62.32	-24.55	116.44***	151.75	288.97**
	(66.10)	(223.28)	(41.51)	(93.35)	(36.32)	(208.76)	(118.22)
Sick Pay	69.79*	-50.69	125.99***	77.58**	151.86***	86.02	22.73
	(39.11)	(91.09)	(28.37)	(34.03)	(19.34)	(155.69)	(74.34)
Individuals (Cell)	41,581	10,852	80,020	13,013	88,695	4,576	5,789
Aged 35-39							
Income	-208.11	-214.052	-893.46***	68.23	-210.68**	-306.43	-260.85
	(163.18)	(427.66)	(113.84)	(251.30)	(84.92)	(469.13)	(388.23)
Unemployment benefits	85.21*	-67.43	129.02***	74.55	231.11***	171.82	248.15**
	(48.87)	(109.19)	(39.39)	(56.32)	(29.11)	(171.82)	(111.90)
Welfare benefits	173.66***	46.28	23.58	-32.39	90.70***	282.50*	73.35
	(51.74)	(143.97)	(33.12)	(82.78)	(33.16)	(152.82)	(103.05)
Sick Pay	67.02	215.51***	66.73***	87.00***	120.18***	116.10	124.60
	(26.12)	(76.38)	(21.21)	(31.74)	(15.44)	(86.43)	(72.43)
Individuals (Cell)	46,830	11,620	89,714	14,879	100,350	5,108	6,393
Aged 40-44							
Income	-292.11**	-665.644**	-626.33***	139.87	122.84	-860.93**	-1327.28***
	(136.65)	(324.16)	(95.05)	(213.45)	(77.85)	(386.80)	(360.50)
Unemployment benefits	159.39***	52.38	213.20***	61.40	313.58***	196.33	135.11
	(41.57)	(95.07)	(34.10)	(56.10)	(27.60)	(132.73)	(101.47)
Welfare benefits	246.77***	-13.25	94.85***	-0.42	135.31***	-17.50	138.91
	(47.78)	(141.73)	(28.66)	(86.61)	(31.41)	(120.92)	(104.68)
Sick Pay	53.23***	89.92	40.14**	30.44	88.03***	131.16*	47.36
	(20.15)	(51.45)	(18.53)	(20.78)	(12.95)	(71.25)	(59.82)
Individuals (Cell)	54,464	13,206	101,126	16,675	110,580	5,889	7,037

Note: This table shows the estimates of the DADID specified in equation (7) *per initial disability condition. Significance stars indicate significance of the estimate at the 1 percent level ***, 5 percent level **, 10 percent level *. Standard errors are in parentheses, clustered at the individual level. The observations correspond to the number of individuals suffering from a specific diagnosis in the first period observed.*

Figure D.1. : Cumulative distribution of reassessments of claimants aged 30-44 in July 2004 per disability cause



Source: Authors' calculations using UWV data